Quantifying P -wave secondary microseisms events: a comparison of observed and modelled backprojection

R Zhang, P Boué, Michel Campillo, J Ma

Geophysical Journal International

Geophys. J. Int. (2023) **234,** 933–947
Advance Access publication 2023 March 09
GJI Seismology

https://doi.org/10.1093/gji/ggad103

Quantifying *P*-wave secondary microseisms events: a comparison of observed and modelled backprojection

R. Zhang,[1,2] P. Boué,[2] M. Campillo[2] and J. Ma[1,3]

[1] *School of Mathematics and Center of Geophysics, Harbin Institute of Technology,* 150001 *Harbin, China. E-mail: jma@hit.edu.cn*
[2] *Université Grenoble Alpes, Université Savoie Mont Blanc, CNRS, IRD, IFSTTAR, ISTerre, Grenoble, France*
[3] *School of Earth and Space Sciences, Center of Artificial Intelligence Geoscience, Peking University, Beijing* 100871, *China*

Accepted 2023 March 1. Received 2023 March 1; in original form 2022 September 5

SUMMARY

Secondary microseisms are caused by nonlinear interactions between ocean waves of approximately equal wavelengths and opposite propagation directions. This seismic forcing is evaluated using ocean sea-state hindcast data and further modulated by the bathymetric effect. The numerical ocean model provides a global activity representation of the secondary microseisms, from which we isolate major events. We backprojected teleseismic *P*-wave propagation into the Earth's mantle to validate these events as effective seismic sources. The ocean model provides spectral amplitude information for modelling microseisms generated seismic wavefield. A comparison of the backprojection for *P* and *PP* phases from observed and synthetic microseisms forcing indicates high reliability in the ocean model, at least for major sources. A combination of *P* and *PP* phases detected across a global network of stations enables global ocean coverage. We improve backprojection images even further by introducing a two-step stacking for the *P* phase to address the problem of unbalanced station distribution. Thresholds of microseisms events forces valuable for seismic imaging are determined by comparing backprojections and ocean models for the years 2015 and 2020. Finally, we extracted a catalogue of microseisms events every 3-hr from 1994 to 2020 from the ocean hindcast data set. This catalogue is an intriguing resource for future applications of interferometric imaging at large scale.

Key words: Numerical modelling; Body waves; Seismic noise; Wave propagation.

1 INTRODUCTION

Microseisms are permanent and ubiquitous vibration of the ground as part of ambient noise (e.g. Gutenberg 1958; Ardhuin *et al.* 2015). It is usually divided into primary and secondary microseisms, with corresponding periods of about 10–20 and 1–10 s, respectively. The primary microseisms have relatively low energetic peak in the frequency spectrum (Peterson 1993; Berger *et al.* 2004). It is generated by the direct interaction of offshore ocean gravity waves with the seafloor and therefore has the same frequency as the ocean waves (Hasselmann 1963; Ardhuin *et al.* 2015). Secondary microseisms have more energy than primary microseisms (Peterson 1993; Berger *et al.* 2004). Its mechanism is well understood and it is caused by nonlinear interactions of ocean gravity waves with approximately equal frequencies and opposite propagation directions; it is twice the frequency of ocean waves (Longuet-Higgins 1950; Hasselmann 1963; Ardhuin *et al.* 2011). Direct teleseismic *P*-wave observations from background ambient noise analysis were reported along with the early development of array seismology (e.g. Backus *et al.* 1964) and quickly associated with ocean microseisms (e.g. Vinnik 1973; Toksöz & Lacoss 1968). More recent detection analyses completed these early studies with array-based backprojections from various types of seismic phases (e.g. Gerstoft *et al.* 2008; Landès *et al.* 2010), highlighting also the good agreement with model-based microseisms sources maps computed from ocean hindcast models (e.g. Nishida & Takagi 2016; Retailleau & Gualtieri 2021).

While such observations may be useful for constraining ocean-state, no direct use of such wavefields was considered for imaging applications before the rise of interferometric approaches in the last 20 yr. The spatial and temporal complexity of the source time function (STF) necessitates reliance on relative observation between station pairs, which is the purpose of the cross-correlation operator. In the context of large-scale imaging applications, it has been demonstrated that teleseismic *P*-wave phases emerge from the cross-correlation of continuous seismic recordings in the secondary microseisms frequency band, which is an excellent match compared to 1-D Earth response expectations (Poli *et al.* 2012; Boué *et al.* 2013; Xia *et al.* 2016). In addition to these expected seismic phases, cross-correlation wavefields reveal spurious seismic phases (phases with no counterparts in the Earth's natural response), highlighting a significant break in the underlying hypothesis of evenly distributed sources at the surface, which should guarantee at least

a partial convergence of such a correlation function toward the expected Green's function (e.g. Ruigrok *et al.* 2008 for a theoretical derivation of interferometric relations at the global scale). For instance, Li *et al.* (2020a), clearly demonstrated the impact of a particular secondary microseisms source in a highly active region near New Zealand. Some researchers proposed using only one source to focus on specific interferences and avoid unwanted ones (Pedersen & Colombi 2018; Xie *et al.* 2021). Nishida & Takagi (2022) recently created a catalogue of global *P*-wave microseisms using a dedicated auto-focusing beamformer using years of continuous data on a dense regional array (Hi-net stations, Japan). We propose to expand on this concept by also proposing a catalogue of significant secondary microseisms sources of *P* waves using seismological data and ocean models. We expect that such a catalogue will help explore the potential of seismic interferometry on a large scale.

Based on the microseisms excitation theory and the development of both the ocean observation and modelling techniques, the ocean pressure to solid Earth seismic coupling can be computed (e.g. Ardhuin *et al.* 2011; Rascle & Ardhuin 2013). The oceanic wave-wave interaction model provides the power spectral density of the pressure field at the ocean surface that is responsible for the source of seismic waves in the solid Earth for the secondary microseisms frequency band (Fig. 1a). Bathymetry affects the coupling of the ocean-generated pressure field to seismic waves; scalar coefficients can be determined as a function of wave types (surface, *P*, *S*) and considered slowness (Longuet-Higgins 1950; Stutzmann *et al.* 2012; Gualtieri *et al.* 2013, 2014). Gualtieri *et al.* (2014), in particular, defined the wavefield as the superposition of plane waves, deduced the bathymetric effect coefficients of *P* and *S* waves, and demonstrated that the bathymetric effect is stronger for *P* waves than *S* waves. After being modulated by the appropriate bathymetric coefficient, the source model was proved to be a good representation of oceanic forcing in this frequency range, being close to observed seismic data (Hillers *et al.* 2012; Obrebski *et al.* 2013; Farra *et al.* 2016; Li *et al.* 2020b; Nishida & Takagi 2022). Still, from a seismic imaging perspective, it would be very useful and convenient to define a set of parameters to reduce this continuum (although numerically discrete) sources in both time and space to a discrete catalogue of significant events. Such parameters are difficult to estimate without a strategy for comparing the source model with seismic data. Our research combines these ocean hindcast data with seismic observations to quantify valuable microseisms events criteria. Our objective is to extract a reliable catalogue of significant secondary microseisms sources directly from the model, which is much more efficient than exploring years of global continuous seismic data. Nevertheless, the quality of the model prediction is confronted to some seismic data based on mantle *P* waves.

In this study, we define a list of microseisms events based on their mantle *P* waves detectability at the global scale in a period band ranging from 3 to 10 s. This catalogue will be generated from the multiyear ocean model after it has been partially validated with seismic data at the global scale and a selection of 60 time-windows (from 0:00 to 3:00 UTC time between 2011 and 2020). First, synthetic seismograms are generated using the secondary microseisms pressure field sea-state hindcast data as input amplitude. Second, a strategy of seismic data and synthetic backprojection for *P* and *PP* phases is established for a fair comparison. This comparison allowed us to set a threshold value for the model above which we can expect to seismically isolate an event, without the need to explore a tremendous continuous global seismic data set for years. Finally, we discuss the main parameters of the event catalogue for the entire years 2015 and 2020.

2 METHODS AND DATA

In this section, we describe how we model secondary microseisms waveforms and how the backprojection approaches for *P* and *PP* phases are applied. The modelled waveforms are in the 3–10 s period bands, which are the most energetic part of the secondary microseisms spectrum. Traditional delay-and-sum backprojection aligns and stacks phases simply (hereafter referred to as simple stacking), but because the station distribution is highly non-uniform, the coverage difference for different locations on the Earth causes strong distortions in the backprojection maps. A two-step stacking method for the *P* phase is proposed and tested to mitigate the impact of uneven-stacking order.

2.1 Secondary microseisms modelling of *P* wave

Bathymetry modulated ocean models provide amplitude information for pressure acting on the sea floor. The resulting source map can be convolved with global synthetic waveforms to model the microseisms records at any seismic station. The microseisms source provided by ocean hindcast data is the PSD of the continuous non-stationary pressure field on the ocean surface ($F_p(\mathbf{K} \simeq 0, f_2)$ in the following formula), it cannot be used directly as a source. We convert it to a vertical force according to Gualtieri *et al.* (2014):

$$F(f_2) = 2\pi C_P(f_2)\sqrt{F_p(\mathbf{K} \simeq 0, f_2)\,df\,dS} \tag{1}$$

where f is ocean wave frequency, $f_2 = 2f$ is seismic wave frequency. C_P is the bathymetric effect coefficient of *P* waves (phase of interest) and dS is the grid area ($dS = R^2 \sin(\phi) d\lambda d\phi$—where R is the radius of the Earth, ϕ is the colatitude and λ is the longitude). The ocean hindcast data set has a temporal resolution of 3-hr and a spatial resolution of 0.5° for both latitude and longitude. In our application, we adopted oceanic models with reflection, where the reflection coefficient $R^2 = 0.1$ for continents and large islands, $R^2 = 0.2$ for smaller islands, $R^2 = 0.5$ for icebergs (Ardhuin *et al.* 2011). Fig. 1 shows the ocean power spectral density and vertical force after modulation by bathymetric coefficient C_P (Gualtieri *et al.* 2014). This frequency-dependent bathymetric coefficient, which can be seen as a coupling coefficient at the bottom of the water column for a given seismic waves (here diving *P* waves) directly relates to the local resonance for a given bathymetry. Some microseism events are amplified (Fig. 1b, event II) while other are attenuated (Fig. 1b, event I).

Since we expect microseisms sources to be uncorrelated spatially, we apply a random phase to the vertical force provided by the oceanic wave model and we use it as the STF. Green's function of *P*-wave propagation at the global scale is calculated using AxiSEM (Nissen-Meyer *et al.* 2014; van Driel *et al.* 2015) using the 1-D Earth model AK135f (Kennett *et al.* 1995; Montagner & Kennett 1996). The source in this case is the unitary vertical point force. *P* wave propagates from ocean sources to stations and reciprocity is not used. The sampling rate is 20 Hz. We only keep the waveforms from *P* arrival to *PP* arrival (*PP* phases are included) and we mute later arrivals like surface waves to focus on the *P*-only wavefield. We expect this choice to differ significantly from actual continuous seismic data, which contain mostly surface waves in this frequency band, as discussed later. Fig. 2 summarizes our workflow of simulating *P*-only wavefields from the ocean forcing. For a given receiver k, the modelled secondary microseisms ($u_k(t)$) is calculated by summing the convolution of the STF_{ij} and the corresponding Green's

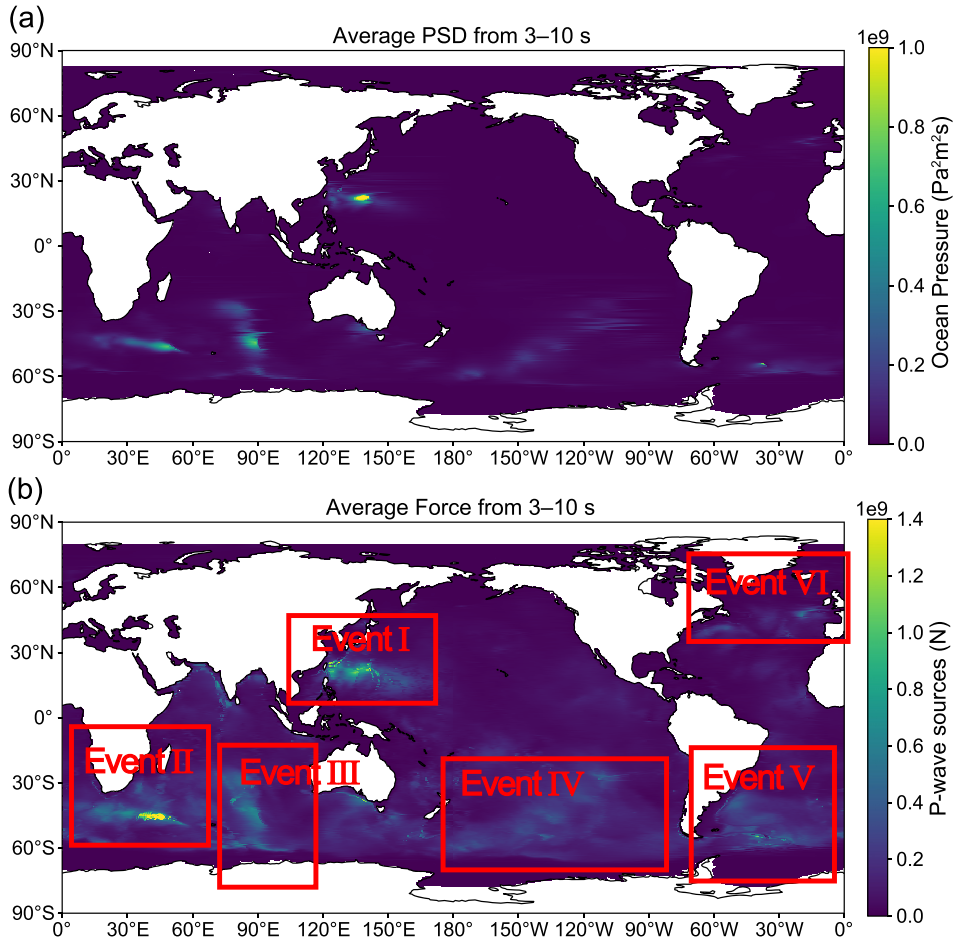

Figure 1. On 2019 August 9, from 0:00 to 3:00, (a) the average PSD of the ocean pressure field from 3 to 10 s, and (b) vertical force after modulation by bathymetric coefficients. The locations of some storms discussed in the following are shown as red boxes.

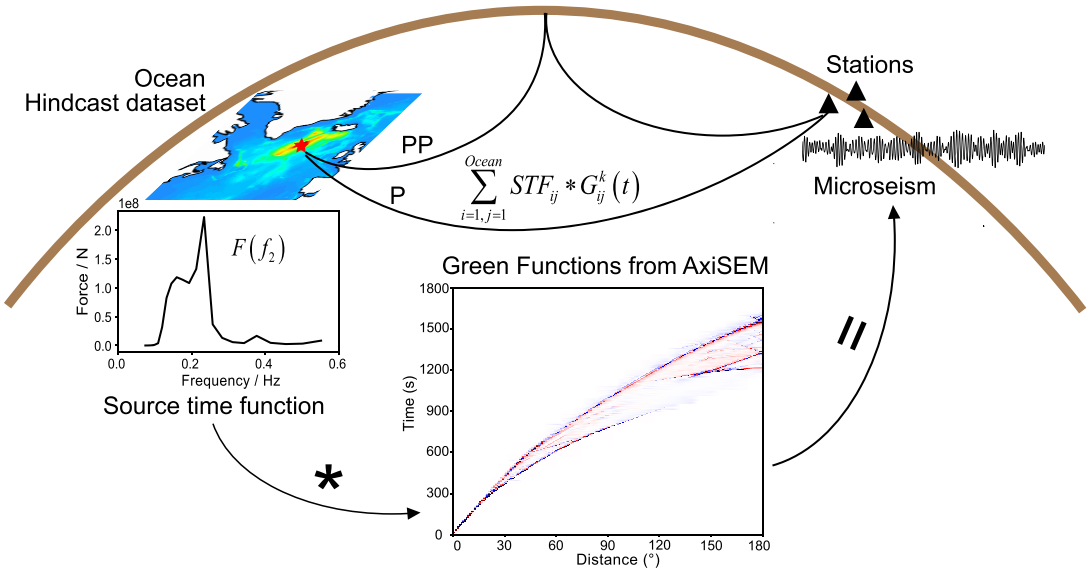

Figure 2. Depicts a sketch of our microseisms modelling strategy. Every point (red star) in the ocean is a source, and its force is a function of frequency; then, impose a random phase to the force as the STF, and convolve it with the Green's function from this point to a station. The modelled microseisms waveforms are obtained by summing the contributions of all ocean surface points to this station.

function ($G_{ij}^{k}(t)$) at each point (i, j) of the ocean:

$$u_k(t) = \sum_{i=1,j=1}^{\text{Ocean}} \text{STF}_{ij} * G_{ij}^{k}(t) \tag{2}$$

We simulate a 30-min recording. On the one hand, the time limit is imposed because the simulation is time-consuming and consumes numerous computing resources. Alternatively, a 30-min Green's function ($G_{ij}^{k}(t)$, as shown in Fig. 2) is sufficient for *P* and *PP* propagation.

2.2 Backprojection

Backprojection is a classical array-based method for searching for source locations and dynamics in seismology (e.g. Ishii *et al.* 2005; Gerstoft *et al.* 2008; Reading *et al.* 2014; Gal *et al.* 2015; Liu *et al.* 2016; Wang *et al.* 2018). The underlying concept is simply to enhance coherent signals while suppressing incoherent noises using delay-and-sum seismic traces (e.g. Rost & Thomas 2002). Although initially developed to detect impulsive sources, it can be used to detect distant microseisms events (e.g. Vinnik 1973). Some comparisons for body waves based on regional arrays revealed a good agreement between seismic observations and ocean models (e.g. Hillers *et al.* 2012; Euler *et al.* 2014; Farra *et al.* 2016; Liu *et al.* 2016). Since our target is the entire ocean surface, we would like a regular coverage of seismic data to backproject for any ocean location. A single small aperture array that would be used for *P*-phase backprojection only is not sufficient to solve for all ocean locations, due to the propagation distance of seismic phases that includes physical shadow zones (e.g. Ishii *et al.* 2005; Gerstoft *et al.* 2008; Hillers *et al.* 2012; Euler *et al.* 2014). Therefore, a global distribution of observation points is necessary. It is also a method for overcoming ambiguous source locations caused by overlapping ray parameters from *P*-waves multiples (Pyle *et al.* 2015; Meschede *et al.* 2017; Ward Neale *et al.* 2018). A combination of multiple phases has also been used to achieve coverage of global sources (e.g. Gerstoft *et al.* 2008; Retailleau *et al.* 2018; Nishida & Takagi 2022). We extend the multi-array and multiphase approaches to use the global stations as a single array to detect *P*-wave sources (Walker & Shearer 2009; Xu *et al.* 2009; Koper *et al.* 2012).

2.2.1 *Simple stacking*

The backprojection method used here is based on a simple delay-and-sum beamforming algorithm in which time delays are parametrized based on the theoretical traveltime computed for a specific seismic phase and source location. The beam amplitude resulting from the stacking procedure for a given set of parameters highlights the existence of microseisms sources. For the source location in latitude i and longitude j, the seismograms are summed:

$$s_{ij}(t) = \sum_{k=1}^{N} w_{ij}^{k} u_k\left(t - t_{ij}^{k}\right) \tag{3}$$

where $u_k(t)$is the vertical component seismogram recorded at the kth station and t_{ij}^{k} is the theoretical *P* or *PP* phase traveltime from the source (i, j) to the kth station, which is calculated using the global velocity model AK135 (Kennett *et al.* 1995). w_{ij}^{k} is a weighting factor for station k and source point (i, j). We refer to it as a simple stacking procedure when $w_{ij}^{k} = 1$. We compute the beam amplitude $s_{ij}(t)$ by an ocean-wide gridpoint search approach. The grid spacing in latitude and longitude was 2.5°.

Previous *P*-waves studies mostly used regional arrays or a combination of multi-array to investigate some local sources (Gerstoft *et al.* 2008; Liu *et al.* 2016; Ward Neale *et al.* 2018; Euler *et al.* 2014; Wang *et al.* 2018). For multi-array combinations, the backprojection area can only be the intersection of each array's coverage area (Ward Neale *et al.* 2018). Even detecting global sources would be limited by space resolution and *P*-wave propagation distance (Euler *et al.* 2014; Wang *et al.* 2018). Furthermore, as previously stated, since *P* and *PP* have the same slowness at different distances, there is ambiguity between *P* and *PP* backprojection, which can lead to wrong source localization. To search the entire ocean surface and reduce *P*/*PP* backprojection ambiguity, we extend the multi-array approach using all available broadband stations around the world.

To be consistent with the time interval of the ocean model available (Ardhuin *et al.* 2011; Rascle & Ardhuin 2013), we download the corresponding 3-hr seismic data. We collected all available high-gain seismometers for the target period. The vertical component data were downloaded through the FDSN web services. A list of the networks is given in the Supporting Information. The sensor responses were removed from original waveforms through deconvolution. Waveforms were further downsampled to 2 Hz, and band-pass filtered from 3 to 10 s. For earthquake signals, we clipped the waveforms with amplitudes larger than three times the standard deviation in a 3-hr window. However, the impact of large earthquakes ($M_w > 5.5$, impact duration > 30-min) will still be significant on our backprojection maps. We thus decided to avoid periods with significant earthquakes to focus on ocean forcing. Backprojection for the *P* phase has a distance range of 30°–98°, which is the typical range of teleseismic *P* wave and avoids upper-mantle triplication of *P*-wave traveltime before 30°. The distance range for the *PP* phase is 30°–180°.

An example of backprojection for *P* and *PP* for 2019 August 9, is presented in Fig. 3, while the source distribution predicted by the ocean model is depicted in Fig. 1(b). The beam amplitude is normalized to one for the representation, and the zoom panel is normalized by values in the zoom range. The stacking amplifies possible sources of coherent *P* (Fig. 3a) or *PP* (Fig. 3b) excitation while reducing the impact of incoherent phases, which are supposed to be mostly surface waves in this frequency range. Stacking order refers to the number of stations used for a given ocean point. For example, for *P* phase it is the number of stations within 30°–98° from source point (i, j). A theoretically perfectly coherent signal hidden in an incoherent noise should stack with a signal-to-noise ratio (SNR) increase $\sqrt{M}$ (M is stacking order) compared to a single station. Alternatively, the beam amplitude also decreases with the increase of stacking order; see Figs 3(a) and (c) for the Antarctic Ocean, and Figs 3(b) and (d) for the North-East Pacific Ocean. Although counterintuitive, this observation is directly related to the large amplitude of the Rayleigh wave in the secondary microseisms that are stacked incoherently when aligned with *P* and *PP* traveltime. In other words, this effect is particularly problematic when numerous stations are sensitive to local (and strong) surface wave excitation. As seismic stations in Europe and the US constitute a large part of the overall available stations, they significantly unbalance the stacking order and thus produce unwanted patterns in the backprojection results. In the case of the *P* backprojection, in locations where the number of available stations is greatly reduced due to the limited distance range considered, the beam amplitude increases significantly due to the contribution of regional noise in the vicinity of the remaining

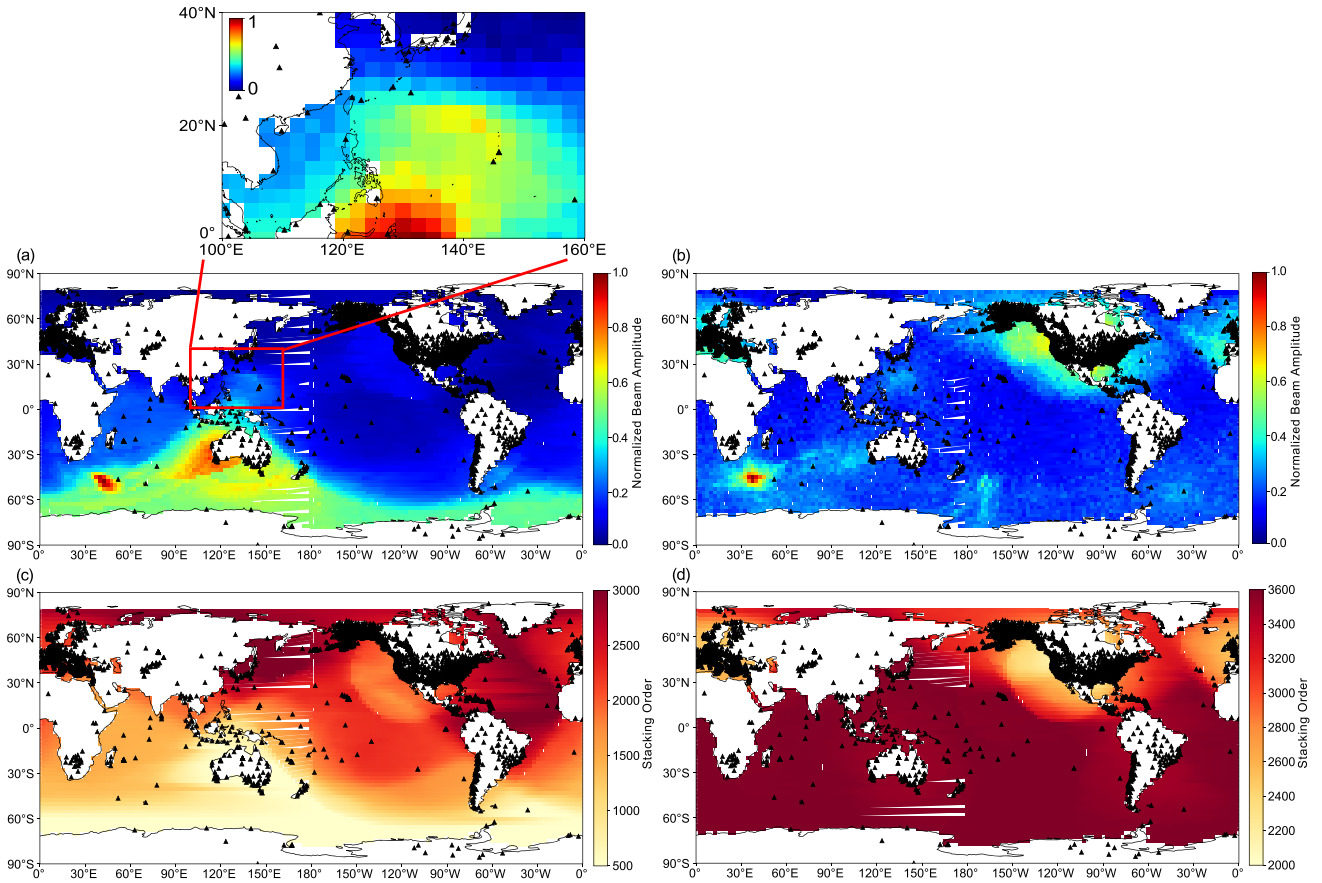

Figure 3. Backprojection map of 2019 August 9, for (a) *P* phase, (b) *PP* phase, and corresponding stacking order of (c) *P* phase, (d) *PP* phase. The beam amplitude is normalized to 0–1 and the zoom panel in (a) is normalized by values in the zoom range.

stations. The comparison with the ocean activity depicted in Fig. 1 indicates that this high amplitude pattern is a spurious feature. For the same reason, a shadow zone is formed around the United States and Europe for the *PP* backprojection.

When the stacking order is low, the coherent but weak *P*-wave signal is drowned out by the incoherent but strong surface waves. There is a microseisms source (Fig. 1) south of Japan, as shown in Fig. 3(a) zoom, but due to the low stacking order in this area, the source cannot be effectively resolved. Due to the limited available distance for backprojection, it is impossible to keep the stacking order consistent on a global scale without significantly decreasing the gain of the coherent stacking.

The stacking order for *PP* simple stacking backprojection is high in the Southern Hemisphere, while its relative spatial variations are small in this area. Consequently, the sources in the Southern Hemisphere are more constrained compared with Fig. 1. For *P*-wave backprojection, the stacking order varies significantly and the distribution is not uniform. To address this problem, which is particularly problematic for *P* waves, we propose a two-step stacking procedure to artificially balance the station distribution with spatial normalization. Finally, we should note that for *PP* backprojection, the ambiguity caused by the slowness of the coincidence of *P* and *PP* persists (see example in Fig. S1, Supporting Information). In Fig. 3(b), there is no source at (60° S, 165° W), but we observe an artefact caused by a strong source south of Africa.

2.2.2 *Two-step stacking*

The idea of the two-step stacking is to weight seismic signals according to the spatial distribution of the corresponding stations. The goal is to reduce the relative importance of the large number of stations that are close to each other. Conversely, a station that is far from the others will be given higher priority in the stack. The main objective of this empirical workflow is to balance the stacking order and enable better and more homogeneous detectability of microseisms events across the entire ocean. As shown in Fig. 4, we triangulate the Earth (Moresi & Mather 2019), and grouped the stations accordingly. Every station group was considered a subarray. All traces of the group are shifted and stacked according to the theoretical traveltime relative to the centre of the corresponding triangular patch. The resulting beamed traces of all groups were normalized. Beams are then sorted by distance from the source and binned in 100 km increments for every possible source location. The second step of stacking is finally applied after a final traveltime correction. In this case, weighting factor in eq. (3) is $w_{ij}^k = 1/(P^k Q)$, P^k is the number of stations within the triangle where station k is located, Q is the number of distance bins.

The area of the triangle and the distance interval between the bins are two parameters of this two-step stacking procedure. We tested various combinations of triangle area and distance intervals (the areas are 540 000, 134 000 and 33 500 km^2, and the distance intervals are 50, 100, 200 and 300 km). The distance interval plays

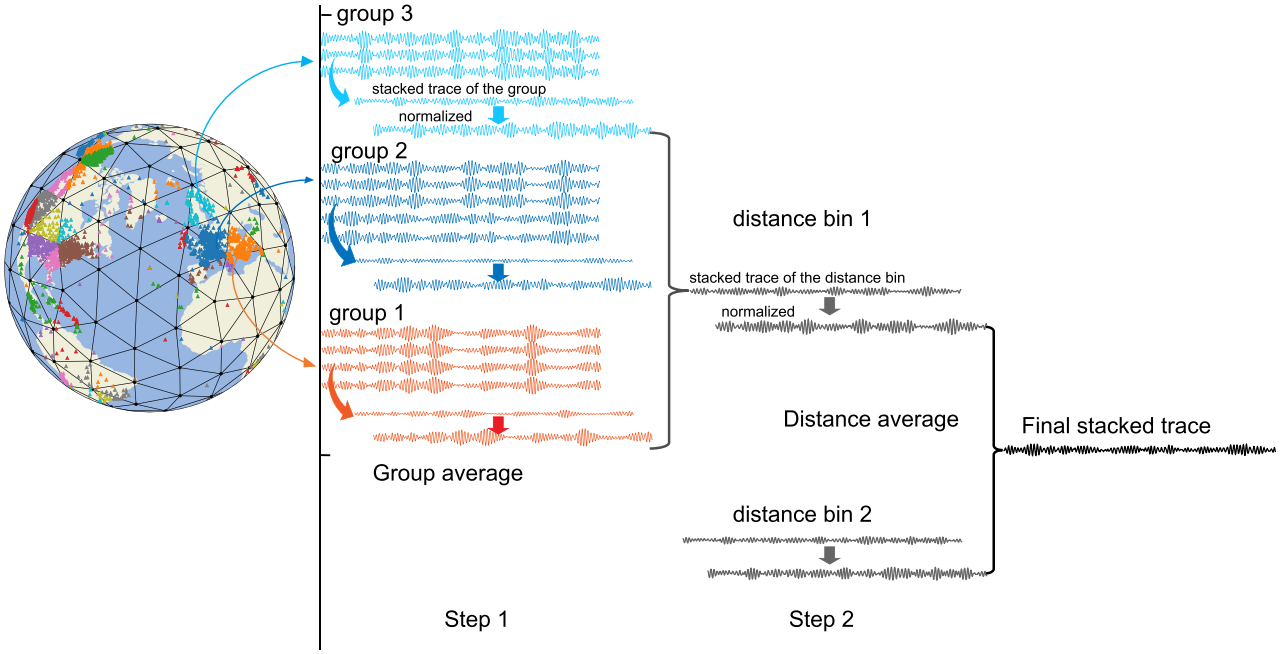

Figure 4. The scheme of two-step stacking. For the sake of clarity, the triangle on the Earth in the figure is much larger than the actual application in the text.

a greater role in homogenizing stacking order than the triangle area (see Fig. S2, Supporting Information). We sacrifice the SNR in some regions to have a little more homogeneous stacking order. Note that the final stacking order is also influenced by the distance interval. A larger distance interval makes the spatial variability of the stacking order smoother but reduces its average value, thus impacting the final image's SNR. Group stacking (step 1 in Fig. 4) suppresses regional incoherent waves, such as Rayleigh waves. For a possible source, groups from different azimuths reduce directivity artefacts in the backprojection. The combination we finally chose is a triangle with a surface area of 134 000 km^2 and a distance interval of 100 km.

Fig. 5 depicts the two-step backprojection map and stacking order distribution on 2019 August 9. The spatial variability of the stacking order is now small compared to Fig. 3(c), especially in the Northern Hemisphere. Note that the significant decrease in the order is counterbalanced by the first step (group) stacking. Comparing Figs 3(a) and 5(a) as illustrated, the smaller stacking order in the Southern Hemisphere still creates a higher noise level on the backprojection map but is less problematic for detection. The events in southern Africa are powerful, and the energy levels exceed the noise, making them well resolved. Simple stacking cannot reliably resolve the event in the western Pacific, but its location and shape are well highlighted by the two-step stacking process.

When using the two-step stacking for the *P* phase and the simple stacking for the *PP* phase, there is still a form of 'shadow zones' on the backprojection maps, which corresponding to low stacking order, that is, low SNR. Events in the shadow zones can only be detected if their energy is sufficiently large, thus emerging from incoherent noise. For simple stacking of *PP* backprojection, the most significant shadow zones are around the United States and Europe, and affect the Bering Sea and Greenland and Iceland (Figs 3b and d). Events around the US and Europe are completely invisible (Fig. S3a, Supporting Information). Very strong events usually occur in south of Greenland and Iceland. These events were still indistinguishable on the *PP* backprojection map (Fig. S3b, Supporting Information). However, these events can be well constrained by the two-step stacking of *P* phase. For the two-step backprojection of the *P* phase, the shadow zones are located widely in the Southern Hemisphere. When the event energy is large enough, it can also be observed in areas with low stacking order (Figs S3c and d, Supporting Information), while the events with weaker energy cannot be observed (Fig. S3e, Supporting Information). Likewise, events in the shadow zones of *P* phase in the Southern Hemisphere can also be well constrained by simple stacking of *PP* phase (Figs S3c–e, Supporting Information). Another point to note is that, stacking order maps are dynamic in time as the global network evolves.

Consequently, to qualify our ability to seismically detect a microseism event wherever in the ocean, we adopt simple stacking for the *PP* phase to constrain Southern Hemisphere events, and two-step stacking for the *P* phase to constrain Northern Hemisphere events. We reject the results for the regions with the lowest stacking orders, that is, the Southern Hemisphere for the *P* backprojection and the Northern Hemisphere for the *PP* backprojection (i.e. the regions where the stacking order is coded in yellow in Figs 3d and 5b).

3 RESULTS

3.1 Correlation coefficient of modelled and observed backprojection

Our objective is now to evaluate the validity of the ocean model based on a comparison with seismic data. Following Section 2.1, synthetic seismograms are derived from the model and compared to actual data through the backprojection introduced in Section 2.2. Given the computational effort required for the modelling, we limit our comparison to 60 time-windows ranging from 2011 to 2020 (approximately six time-windows per year, mostly in January and August, and a few in September and October), all of which include at least one significant microseisms event, expecting that to be representative. Our conclusions from this comparison are then extrapolated to include all duration of the available ocean model. January and August are periods of active ocean activity in the Northern and

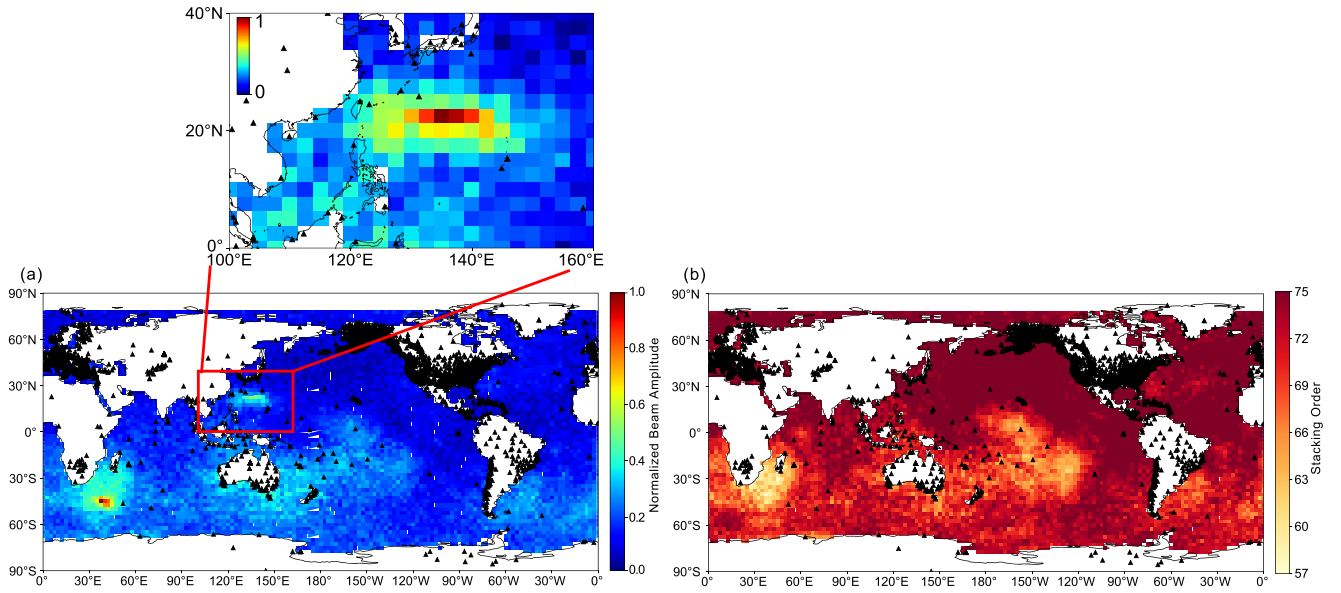

Figure 5. Backprojection map of 2019 August 9, for (a) *P* phase by two-step stacking, and corresponding stacking order (b). The beam amplitude is normalized to 0–1 and the zoom panel in (a) is normalized by values in the zoom range.

Southern Hemispheres (see Section 3.2 below), respectively, and tropical cyclones form in the Western Pacific and North Atlantic in September and October. Active oceans and tropical cyclones are potential microseisms sources. The ocean model used ocean hindcast data from 0:00 to 3:00 UTC every day, but modelled seismic data are limited to 30-min for computational reasons. As mentioned in Section 2.1, 30-min is sufficient for *P* and *PP* propagation. Corresponding seismic data for comparison are in the same period, but the duration time is 3-hr, to ensure the same temporal resolution as the ocean hindcast data set.

We first use the example of 2019 August 9, to present the results, as shown in Fig. 6. This is an excellent illustration, of significant events in both the Northern and Southern Hemispheres. Overall, the observed and modelled backprojections exhibit a good agreement. We identified six specific events, as illustrated in Fig. 1(b). For the *P* phase, we compare the observed backprojection in Fig. 6(a) to the modelled backprojection in Fig. 6(c). Events I and II are visible in both the observed and modelled backprojection maps, with the modelled map having a higher amplitude. Despite their large distribution area, events III and V have weak imprints in the observed and modelled backprojection maps, which are due to the weak force at each gridpoint. Event IV has a relatively high magnitude in both the observed and modelled backprojection maps, which is due not only to its presence but also to its location, which corresponds to a lower station order. The higher amplitudes observed throughout Australia are due to the same reason. Event VI has a weak imprint in the modelled map, which is not visible in the observed map, simply because of its low energy. For the *PP* phase, we compare its observed backprojection in Fig. 6(b) with the modelled backprojection in Fig. 6(d). Events I–V are relatively clear in the modelled backprojection, whereas in the observed backprojection, only event II is very clear, and events I, III, IV and V are weak. We find that the backprojected seismic energy of event II is significantly higher than that of other events. In this case, the ocean model appears to underestimate the seismic energy of event II. Event III is relatively clear in modelled map for *PP* phase, implying that it is likely overestimated by the ocean model. Event IV has a relatively high amplitude, with a contribution from event IV, but the larger contribution comes from the artefact of event II. Event VI is not visible in either the observed or modelled maps. There is some incoherent noise in the modelled backprojection map, which corresponds to the imprint of stacking order. Modelled waveforms contain not only coherent *P* or *PP* phases but also *PcP* and *PKP* phases, which act as incoherent noise when aligning on *P* or *PP* phases.

We now study the averaged properties computed on a subset of 60 time-windows. Fig. 7 depicts the average equivalent force field (integration of eq. 1 over frequency and area) for the selected 60 time-windows. There are significant high-force sources in the Southern Ocean, Northwest Pacific and North Atlantic, which allow us to assess the accuracy of these force fields in ocean models. Backprojections of *P* and *PP* of the modelled data were then calculated using two-step and simple stackings, respectively.

To quantify the consistency of observations and simulations systematically, we calculated the correlation coefficients between observed and modelled beam energy for the 60 time-series at all ocean positions of the backprojection maps. Fig. 8 depicts the results. The correlation coefficient value of each point on the ocean in Figs 8(c) and (d) is the correlation coefficient of the observed backprojection amplitude trace and the modelled backprojection amplitude trace on this point. We chose points 1 and 2 to visualize the agreement between the observations and the models. The backprojection of the *P* phase for point 1 has a high correlation, and Fig. 8(a) also shows the consistency of observed and modelled backprojection amplitude changes. Fig. 7 shows a strong microseisms source activity at point 1. However, the correlation for backprojection of the *PP* phase at this point is poor, which is due to the small station order of *PP* at this location (Fig. 3d), and the *PP* phase is submerged by incoherent signals such as surface waves. The *P* phase at point 2 is completely uncorrelated, despite the presence of microseisms sources there. We can invoke here the same reasoning as for the low correlation of *PP* at point 1. *PP* exhibits a good correlation at point 2, and the observed and modelled backprojection amplitude fluctuations agree well (Fig. 8f). Overall, the *P* phase has a high correlation in the western and northern Pacific, the eastern seas

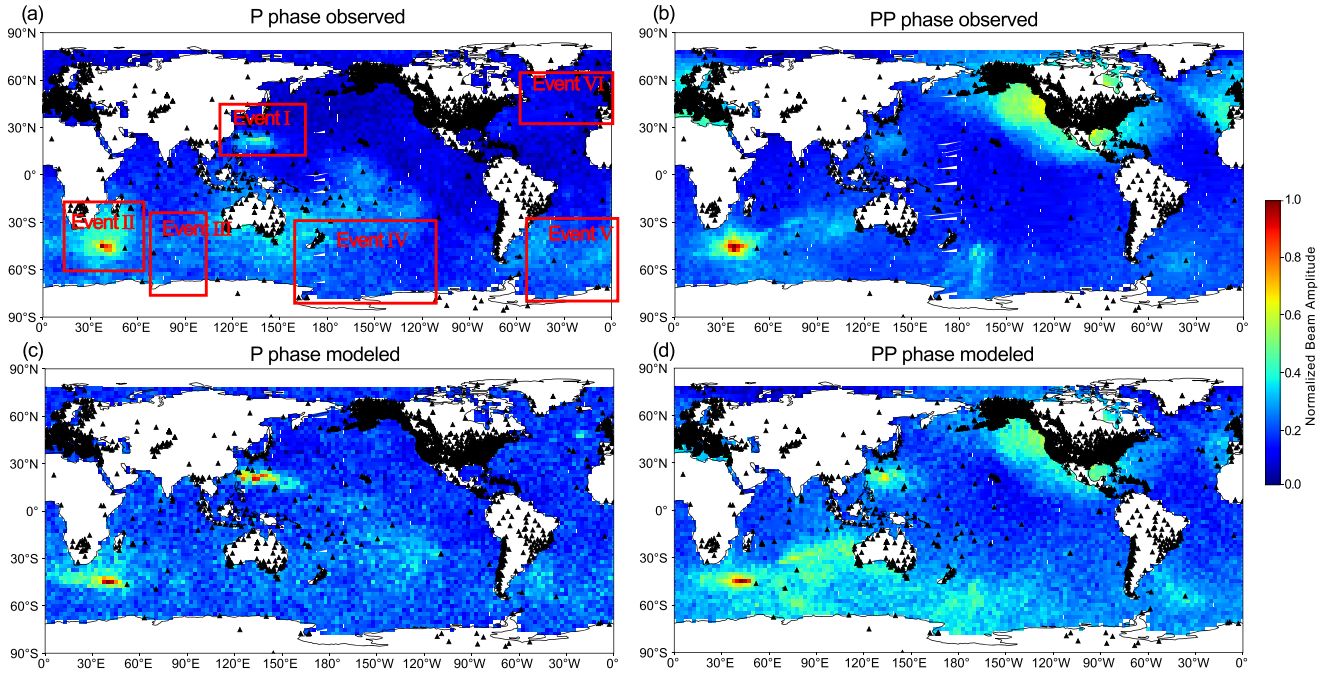

Figure 6. On 2019 August 9, the observed backprojection map (a) and modelled backprojection map (c) for the *P* phase, and the observed backprojection map (b) and modelled backprojection map (d) for the *PP* phase.

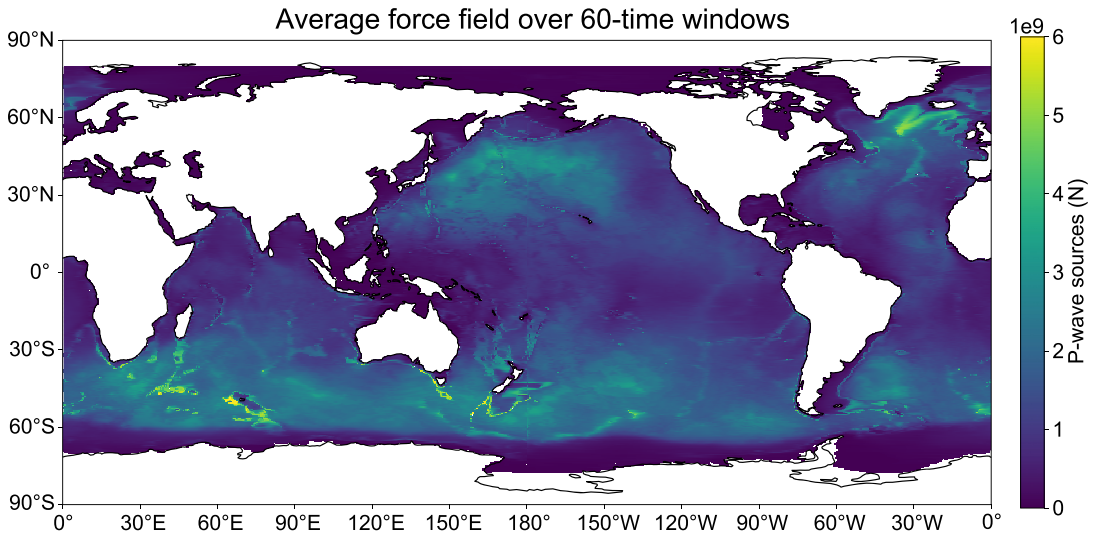

Figure 7. 60 time-windows average force field as modelled from ocean hindcast data.

of the United States, and the North Atlantic (southern Greenland), which are also the regions where the sources are concentrated in the Northern Hemisphere (Fig. 7). The *PP* phase exhibits a very high correlation over the entire Southern Ocean region from 30° S to 60° S, which is also the region where the sources are concentrated in the Southern Hemisphere (Fig. 7).

3.2 Events classification

In areas where ocean storm activity is concentrated, the high correlation coefficient computed over these 60 time-windows between the model and observations (*P* or *PP*) means that we can rely on the model to define seismically effective microseisms events. To quantify more precisely what should be the equivalent force threshold above which, when exploring the model, we can expect to isolate a source in the data side, we further explore continuous seismic data. We calculated the backprojections of *P* and *PP* at 00:00–03:00 of UTC each day in 2015 and 2020 from observed seismic data and compared them directly with ocean models (i.e. Fig. 1b, without synthetic backprojection) for the corresponding period to derive the observable thresholds for microseisms events in ocean models. We used the standard level set-based activate contour method (Osher & Sethian 1988) to isolate microseisms events from the ocean model and calculate their equivalent forces. The microseisms events obtained from the backprojection calculated from the global array

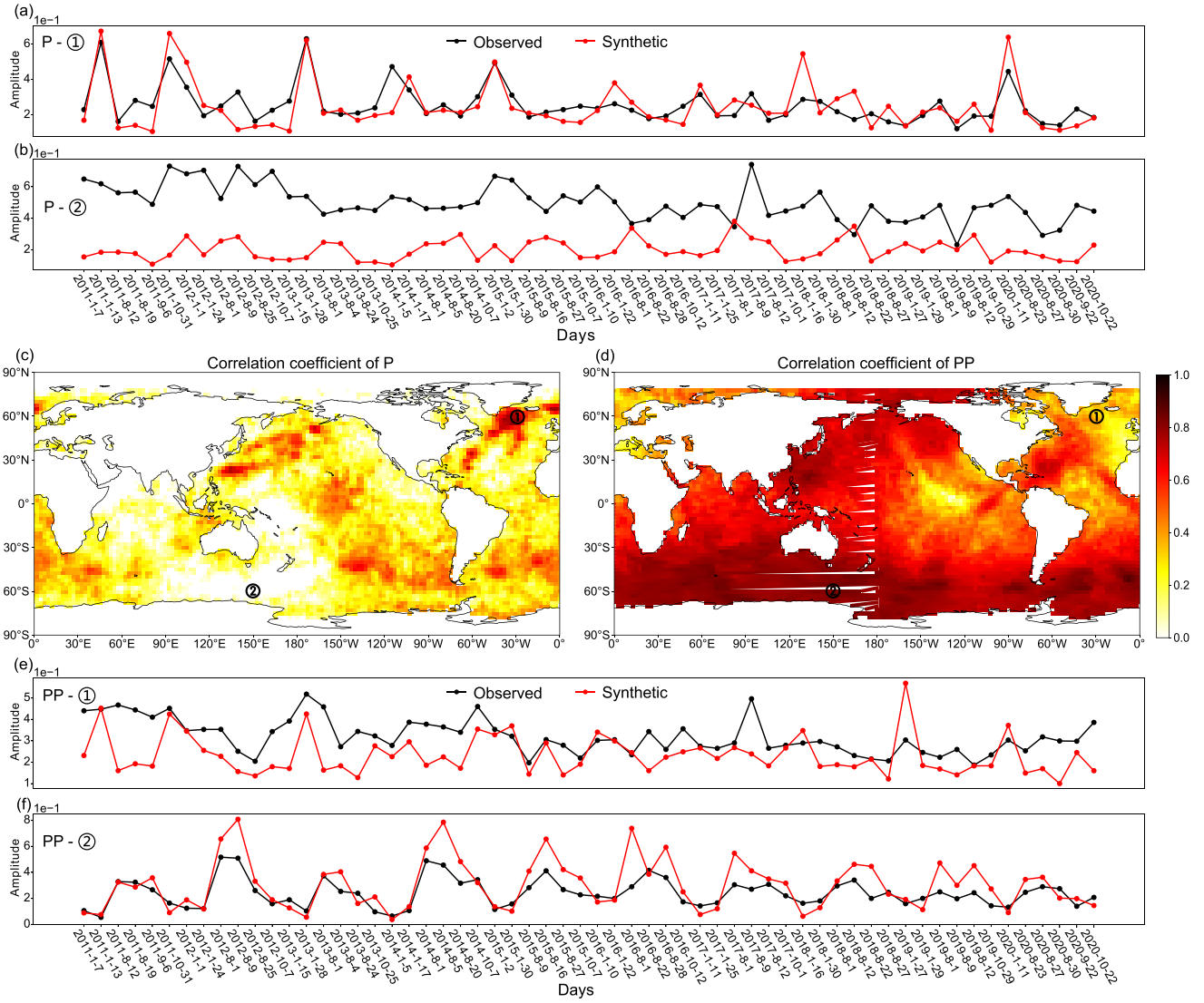

Figure 8. Correlation coefficient map of *P* backprojection (c) and *PP* backprojection (d). (a) 60 time-windows *P* phase backprojection normalized amplitude values at point 1. (b) 60 time-windows *P* phase backprojection normalized amplitude values at point 2. (e) 60 time-windows *PP* phase backprojection normalized amplitude values at point 1. (f) 60 time-windows *PP* phase backprojection normalized amplitude values at point 2.

can be regarded as globally valid, as discussed in the previous section. To minimize the impact of the remaining inherent noise in our global-stacking approach, we calculate the *P* and *PP* backprojection of some regional arrays to add additional constraints.

According to the observability in the backprojection map, we divide the microseisms events isolated from the ocean model into five categories of 0 to 4, as shown in Table 1. Backprojection map examples are presented as supporting information (Figs S4–S7, Supporting Information). The distributions of the equivalent forces over time for these five types of events are depicted in Figs 9(a) and (b). There are more class-4 (strong) events on days 150–250. This is because it is winter in the Southern Hemisphere, where the ocean activity is high. The wider ocean in the Southern Hemisphere makes the strong events more numerous (Figs 9c and d). The corresponding Northern Hemisphere winters are 0–50 and 300–365 days, with fewer class-4 events. Although strong ocean activity occurs in the Northern Hemisphere winter, it is limited to the Northern Atlantic and Pacific (Figs 9c and d). The remaining days correspond to spring and fall in the Southern or Northern Hemispheres when overall ocean activities are lower. Each event level spreads significantly in terms of actual force value, with extensive overlap between the levels. However, as expected, class-4 events correspond to stronger modelled sources. The equivalent forces for the events can be roughly divided into three layers, with separation limits of 2×10^{12} and 6×10^{11} N. Forces greater than 2×10^{12} N correspond to events with clear teleseismic *P* wave and are thus of direct interest for deep seismic imaging. Between 2×10^{12} and 6×10^{11} N, the oceanic events are possibly usable, while below they can be discarded for use as single events.

We further explore the relationship between detectability with SNR and site noise level. We take the median absolute deviation (MAD) of the ocean force field as the background force value, and plot the ratio of the maximum equivalent force to MAD (MAX/MAD) as the SNR (e.g. Nishida & Takagi 2022) on Figs 9(a) and (b). The MAX/MAD curves (from ocean models) are consistent with the class 3 and 4 events (from observed seismic data, see also Nishida & Takagi 2022). In order to compare our detection with a proxy of this local noise, we plot the mean squared amplitude (MSA)

of all the stations as the site noise. As expected in the secondary microseism frequency range, the MSA curves are lower in the Northern Hemisphere summer (days 150–200) and higher in winter. This incoherent noise, mostly dominated by surface waves in the target frequency range, decrease the coherence of global stacking.

Figs 9(c) and (d) show the spatial distribution of all events sorted as a function of their class for both 2015 and 2020. The events are concentrated in the Southern Ocean, the western and northern Pacific Ocean, and the North Atlantic. The distribution of events corresponds to a higher bathymetric coefficient C_P, such as in southern New Zealand and the South Pacific, where events can be seen along grey traces (high *P*-wave amplification). Globally, they are roughly concentrated in the mid-latitudes range of 30°–60°.

Based on this seismic validation of the ocean-based model, we can now propose a catalogue of microseisms events derived directly from ocean hindcast data. Fig. 10 depicts every 3-hr event in 2015 and 2020 with a force greater than 6×10^{11} N. Events in the Northern Hemisphere peak from January to March and October to December, while events in the Southern Hemisphere peak from April to September. Spatially, as in Figs 9(c) and (d), it is concentrated in the mid-latitude 30°–60° range. The colour of the dots represents the evolution of time, highlighting some storm tracks related to ocean dynamics. For instance, warm-toned dots in the Pacific Ocean off eastern Asia and the Atlantic Ocean offshore of the east coast of the United States correspond to typhoons or hurricanes in the Northern Hemisphere summer.

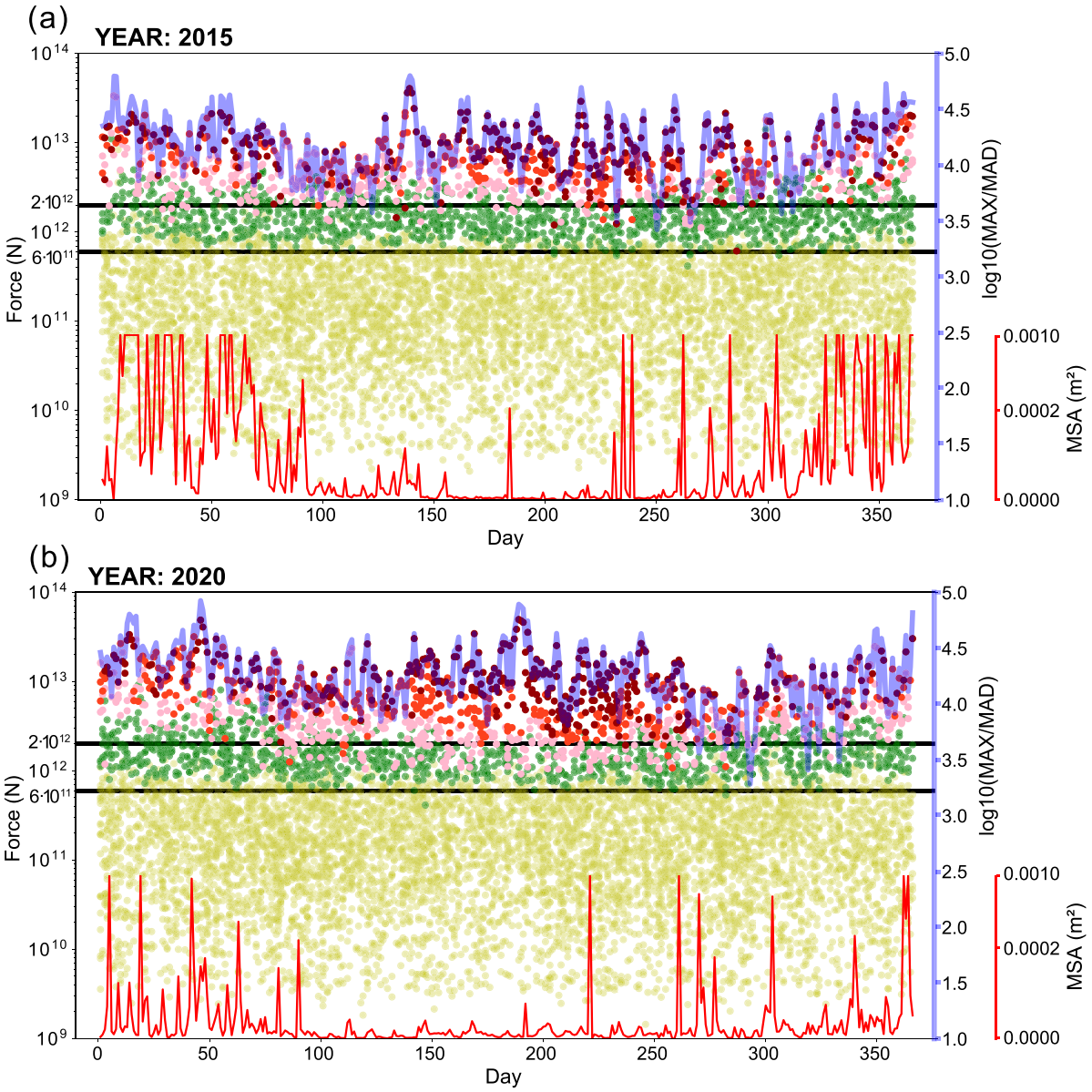

Figure 9. (a) and (b) Equivalent forces of five categories of microseisms events for 2015 and 2020. The light blue curves are SNR ($\log_{10}$(MAX/MAD)) of the ocean force fields, and the red curves are the MSA of the station (because the value of MSA varies widely, we limit it to 0–0.001). (c) and (d) Spatial distribution of five different classes of microseisms events in the ocean for 2015 and 2020. The background ocean map is the average of the *P*-wave bathymetric coefficients over 3–10 s.

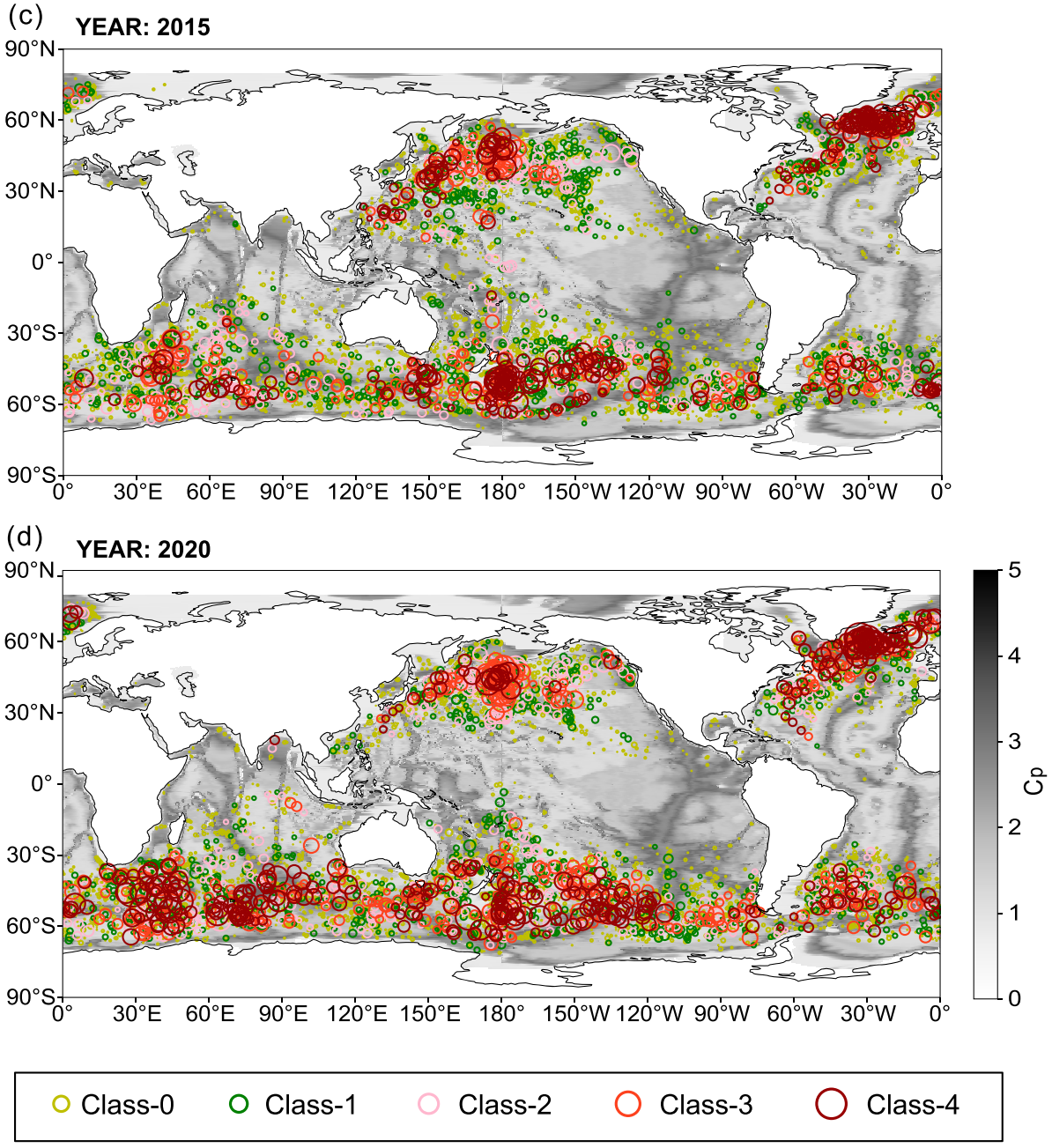

Figure 9. Continued.

Table 1. Five categories of microseisms events based on the observation.

Events	0	1	2	3	4
Descriptions of the back projection	Not detectable	Weak from regional arrays	Not detectable on the global map, but clear for at least a regional array	Weak in the global map, but confirmed with the ocean wave model	Obvious in the backprojection map of global stations

4 DISCUSSION AND CONCLUSIONS

We build a catalogue of microseisms events, which are defined by the production of teleseismic *P*-wave energy. We anticipate that such a catalogue will be of interest for interferometric applications, particularly deep Earth imaging. Modelled ocean seismic forcing is directly used to infer the catalogue following quantitative validation with seismic data. When compared using our backprojection approach on a set of selected dates spread over several years, the overall dynamics of the ocean model fit very well with seismic data. Furthermore, we defined a daily-based comparison for two

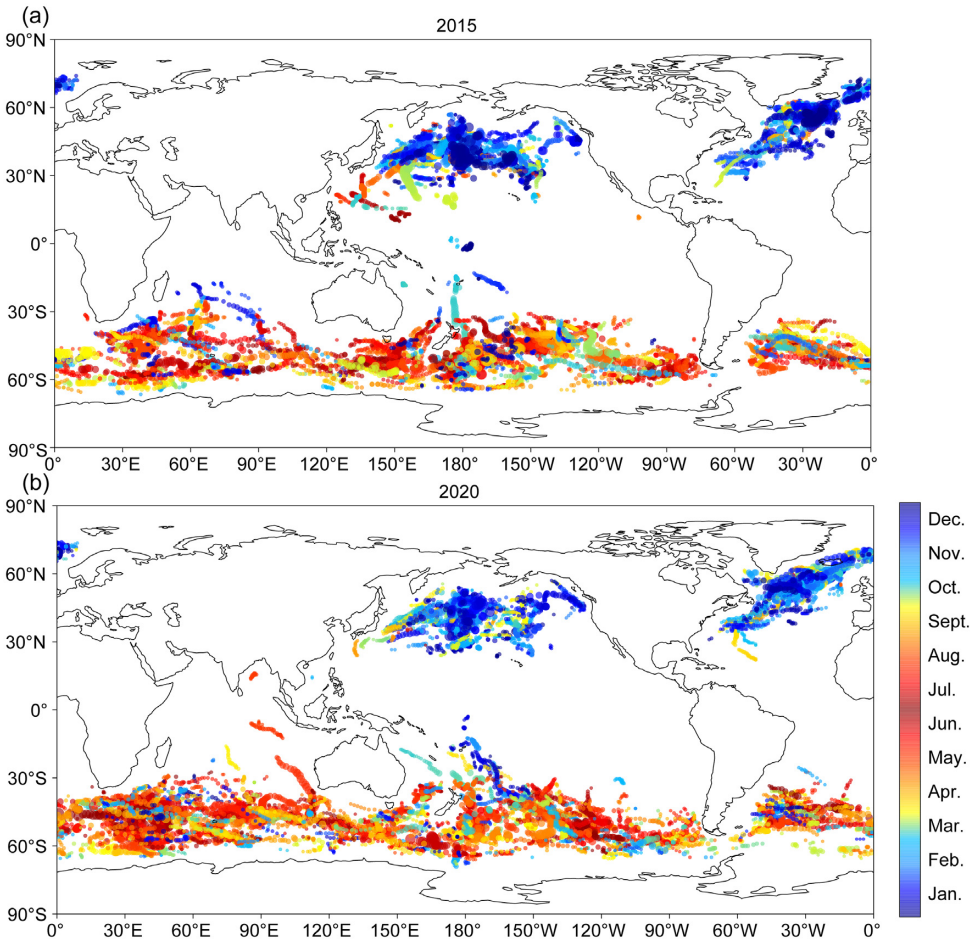

Figure 10. All microseisms events with force greater than 6×10^{11} N in 2015 and 2020.

different years, allowing us to define a threshold value for the modelled microseisms force for which we can expect to detect significant *P* waves at long distances. With this threshold force, we explore the available microseisms model from 1994 to 2020 with a 3-hr temporal resolution. As expected, we verified that microseisms events are always more active in winter and are roughly concentrated in the mid-latitude range of 30°–60° in the ocean.

Regarding the seismic data-based validation of the model, a global station distribution gives better results than a selection of small aperture arrays (Fig. S8 (Supporting information) shows results for small aperture arrays, which are limited by detection range and resolution). The spatial resolution of the global station's backprojection is very high in most regions of the world when looking for teleseismic *P*-wave ray parameters in the secondary microseisms frequency band. In this study, we did not introduce traveltime correction factors as it was done for earthquake backprojections (Ishii *et al.* 2005; Kiser & Ishii 2017). As a result, there are inevitably travel time errors, especially for long-distance *PP* waves (Nishida & Takagi 2022). Globally distributed stations have an averaging effect, and the participation of multiple propagation paths will weaken the traveltime anomaly. To avoid shadow zones caused by uneven station distribution around the world (or at least regions with low SNR backprojections), we use the *PP* phase to extend the backprojection range to 180°. High-precision global coverage comparison of the model and data is achieved through the combination of the *P* phase (to constrain sources in the Northern Hemisphere) and the *PP* phase (to constrain sources in the Southern Hemisphere). However, it is also important to note that some events are still detected in some regions despite inherent low stacking orders. The auto-focusing method developed by Nishida & Takagi (2022) uses the curvature of the wavefronts to account for and distinguish *P* and *PP*. An interesting perspective could be to adapt such an auto-focusing beamformer to a global stacking approach to both tackle the *P* and *PP* ambiguity and the inherent lack of coverage and resolution limits of a regional array. The unbalanced distribution stations and their effect on the stacking order for the *P* wave are mitigated by a two-step stacking approach. This is crucial when accounting for large differences in the noise field spatially, and especially when considering the possibility of energetic surface waves originating from the vicinity of some stations. We could further extend these to the horizontal components and possibly study the generation of *S* waves from ocean forcing on a global scale (Liu *et al.* 2016; Nishida & Takagi 2016; Retailleau & Gualtieri 2021).

The modelled microseisms data preserves the initial ocean model's relative frequency information. On average, we observe a good fit between the modelled source spectra and the spectra we obtain at any source location after backprojection over the ocean's surface (Fig. S9a, Supporting Information). For a single station observation (no backprojection, Figs S9b and c, Supporting Information) some discrepancies appear that are related to our decision to only simulate for *P* wavefield when real seismic records are mostly dominated by Rayleigh waves; plus all complexity resulting from

the 3-D structure seen as deviations in the real data from the 1-D velocity model that we use as an input in our workflow (Fig. 2).

The Southern Ocean, without continental barriers, experiences intense winter storm activity, which triggers widespread microseisms events. However, this spread tended to decrease our ability to seismically detect them. Not only are *P* waves excited in this range, but also surface waves (Rayleigh and Love waves; e.g. Stutzmann *et al.* 2012; Reading *et al.* 2014; Gal *et al.* 2015; Gualtieri *et al.* 2021). This mid-latitude range with prominent secondary microseisms corresponds to the Ferrel cell in the atmospheric circulation, where the prevailing westerlies are the main driving force for ocean current movement (Stommel 1957; Harman 1987; Nishida 2017). Tropical cyclones (typhoons and hurricanes) move from low to high latitudes along the west coast of the Pacific Ocean and the Atlantic Ocean in a southwest-northeast direction. The distribution areas of the corresponding seismic sources are relatively concentrated. The tropical cyclone enters the mid-latitudes and transforms into extratropical, and its intensity may not decrease under the action of the westerlies (Klein *et al.* 2000; Retailleau & Gualtieri 2021). The overall number of events has not changed significantly over the years (Fig. S10a, Supporting Information). We also investigated the relationship between the force and the cumulative number of events (Fig. S10b, Supporting Information). We believe that interpreting such statistics as a Gutenberg–Richter power law for earthquakes is difficult in this context. The main problem arises from the spatiotemporal definition of the ocean event. In our analysis, for instance, a large event can easily splinter into several smaller ones, making it difficult to determine an equivalent *b*-value; however, this is not the case for earthquake detection. Further analysis, such as using a refined temporal mesh for the ocean hindcast data, may aid in understanding the scaling laws for microseisms sources.

Furthermore, a detailed exploration of source geometry (localized, spread and all types) should be conducted for interferometric applications; thus, we decide not to restrict the catalogue to a particular geometry and instead rely on spatially integrated forces. Larger forces tend to imply a larger area (Fig. S10c, Supporting Information), which may be damaging to interferometric imaging when focusing on single events. For this purpose, we recommend focusing on events that generate larger forces over moderate or smaller areas (events in the grey dashed box in Fig. S10d, Supporting Information). Bathymetric effects contribute significantly to source excitation, especially between Greenland and Iceland, southeastern New Zealand, and around the Kerguelen Islands. Intense sources often appear in these regions. In the first approximation, our knowledge of the bathymetry appears sufficient compared to the seismic resolution and inherent diffraction limit in this frequency range for the *P* wave.

Quantification of event detectability is influenced by several parameters. As previously discussed, the two main parameters are stacking order (controlled by the number and the distribution of stations) and equivalent forces of events. Influencing parameters also include noise levels near seismic stations (everything that do not correspond to ocean generated *P* waves), the size of events and global noise level. For these reasons, a constant detectability threshold that would be consistent anywhere at any time cannot be defined. For simplicity, we rather focus on an ad hoc detectability parameter based on the ocean model and the equivalent forces; We argue that forces greater than 6×10^{11} N can be observed by local arrays, and forces greater than 2×10^{12} N can be observed by global stations. This question of a full, data-based, spatiotemporal, definition of detectability, would require further investigations.

Our catalogue includes significant events that occurred simultaneously in different regions. We can already advance the hypothesis that it could be an issue when a single source is targeted for interferometric applications. A synchronous source that is not properly unaccounted for would result in significant bias in the correlograms (e.g. Liu *et al.* 2019 for surface wave case). Furthermore, synchronous sources, such as spread ones, may not be a particular issue if properly accounted for when interpreting correlograms (e.g. Sager *et al.* 2022 at a smaller scale). A simple investigation of the catalogue reveals that it is rare to have only one dominant source at any given time. Generally, we believe that a more detailed investigation of sources between Greenland and Iceland during the Northern Hemisphere winter would be beneficial for imaging purposes. As discussed in numerous other seismological studies, this area can be generally considered a major source on Earth. For the Southern Ocean, dominant sources are often spatially broad and synchronous with each other, making event-based seismic interferometry more difficult.

ACKNOWLEDGMENTS

We acknowledge the support from the National Natural Science Foundation of China under grant 422308064, and the European Research Council under the European Union's Horizon 2020 research and innovation program (grant agreement no. 742335, F-IMAGE), and from the French National Research Agency (ANR) under the project TERRACORR (ANR-20-CE49-0003), and China Scholarship Council. Most of the computations presented in this paper were performed using the GRICAD infrastructure (https://gricad.univ-grenoble-alpes.fr), which is supported by Grenoble research communities. The authors would like to thank the editor Christoph Sens-Schönfelder, and two anonymous reviewers for their time and excellent suggestions that helped to improve the manuscript. The authors would also like to thank Fabrice Ardhuin and Lucia Gualtieri for their help in using the ocean model used in this study. We used data from numerous seismic networks, and we thank all of the people involved in the data acquisition, stations maintenance, QC and distribution. All networks citations are in supplementary documents.

DATA AVAILABILITY

The ocean hindcast data can be found at ftp://ftp.ifremer.fr/ifremer/ww3/HINDCAST/SISMO/. The catalogue in this study is available at the Zenodo https://doi.org/10.5281/zenodo.7269139

REFERENCES

Ardhuin, F., Gualtieri, L. & Stutzmann, E., 2015. How ocean waves rock the Earth: two mechanisms explain microseisms with periods 3 to 300 s, *Geophys. Res. Lett.,* **42**(3), 765–772.

Ardhuin, F., Stutzmann, E., Schimmel, M. & Mangeney, A., 2011. Ocean wave sources of seismic noise, *J. geophys. Res.,* **116**(C9), doi:10.1029/2011jc006952.

Backus, M., Burg, J., Baldwin, D. & Bryan, E., 1964. Wide-band extraction of mantle P waves from ambient noise, *Geophysics,* **29**(5), 672–692.

Berger, J., Davis, P. & Ekström, G., 2004. Ambient Earth noise: a survey of the Global Seismographic Network, *J. geophys. Res.: Solid Earth,* **109**(B11), doi:10.1029/2004jb003408.

Boué, P., Poli, P., Campillo, M., Pedersen, H., Briand, X. & Roux, P., 2013. Teleseismic correlations of ambient seismic noise for deep global imaging of the Earth, *Geophys. J. Int.,* **194**(2), 844–848.

Euler, G. G., Wiens, D. A. & Nyblade, A. A., 2014. Evidence for bathymetric control on the distribution of body wave microseism sources from temporary seismic arrays in Africa, *Geophys. J. Int.,* **197**(3), 1869–1883.

Farra, V., Stutzmann, E., Gualtieri, L., Schimmel, M. & Ardhuin, F., 2016. Ray-theoretical modeling of secondary microseism P waves, *Geophys. J. Int.,* **206**(3), 1730–1739.

Gal, M., Reading, A. M., Ellingsen, S. P., Gualtieri, L., Koper, K. D., Burlacu, R. & Hemer, M. A., 2015. The frequency dependence and locations of short-period microseisms generated in the Southern Ocean and West Pacific, *J. geophys. Res.: Solid Earth,* **120**(8), 5764–5781.

Gerstoft, P., Shearer, P. M., Harmon, N. & Zhang, J., 2008. Global P, PP, and PKP wave microseisms observed from distant storms, *Geophys. Res. Lett.,* **35**(23), doi:10.1029/2008gl036111.

Gualtieri, L., Bachmann, E., Simons, F. J. & Tromp, J., 2021. Generation of secondary microseism love waves: effects of bathymetry, 3-D structure and source seasonality, *Geophys. J. Int.,* **226**(1), 192–219.

Gualtieri, L., Stutzmann, E., Capdeville, Y., Ardhuin, F., Schimmel, M., Mangeney, A. & Morelli, A., 2013. Modelling secondary microseismic noise by normal mode summation, *Geophys. J. Int.,* **193**(3), 1732–1745.

Gualtieri, L., Stutzmann, E., Farra, V., Capdeville, Y., Schimmel, M., Ardhuin, F. & Morelli, A., 2014. Modelling the ocean site effect on seismic noise body waves, *Geophys. J. Int.,* **197**(2), 1096–1106.

Gutenberg, B., 1958. Microseisms, in *Advances in Geophysics*(Vol., **5,** pp. 53–92). Elsevier.

Harman, J.R., 1987. Westerlies, middle-latitude west winds, in: *Climatology. Encyclopedia of Earth Science*. Springer, Boston, MA.

Hasselmann, K., 1963. A statistical analysis of the generation of microseisms, *Rev. Geophys.,* **1**(2), 177, doi:10.1029/rg001i002p00177.

Hillers, G., Graham, N., Campillo, M., Kedar, S., Landès, M. & Shapiro, N., 2012. Global oceanic microseism sources as seen by seismic arrays and predicted by wave action models, *Geochem. Geophys. Geosyst.,* **13**(1), doi:10.1029/2011gc003875.

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

Kennett, B. L., Engdahl, E. R. & Buland, R., 1995. Constraints on seismic velocities in the Earth from traveltimes, *Geophys. J. Int.,* **122**(1), 108–124.

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

Klein, P. M., Harr, P. A. & Elsberry, R. L., 2000. Extratropical transition of western North Pacific tropical cyclones: an overview and conceptual model of the transformation stage, *Weather Forecast.,* **15**(4), 373–395.

Koper, K. D., Hutko, A. R., Lay, T. & Sufri, O., 2012. Imaging short-period seismic radiation from the 27 February 2010 Chile (Mw 8.8) earthquake by back-projection of P, PP, and PKIKP waves, *J. geophys. Res.: Solid Earth,* **117**(B2), doi:10.1029/2011JB008576.

Landès, M., Hubans, F., Shapiro, N. M., Paul, A. & Campillo, M., 2010. Origin of deep ocean microseisms by using teleseismic body waves, *J. geophys. Res.,* **115**(B5), doi:10.1029/2009jb006918.

Li, L., Boué, P. & Campillo, M., 2020a. Observation and explanation of spurious seismic signals emerging in teleseismic noise correlations, *Solid Earth,* **11**(1), 173–184.

Li, L., Boué, P., Retailleau, L. & Campillo, M., 2020b. Spatiotemporal correlation analysis of noise-derived seismic body waves with ocean wave climate and microseism sources, *Geochem. Geophys. Geosyst.,* **21,** doi:10.1029/2020gc009112.

Liu, Q., Koper, K. D., Burlacu, R., Ni, S., Wang, F., Zou, C. & Reading, A. M., 2016. Source locations of teleseismic P, SV, and SH waves observed in microseisms recorded by a large aperture seismic array in China, *Earth planet. Sci. Lett.,* **449,** 39–47.

Liu, X., Beroza, G. C. & Nakata, N., 2019. Isolating and suppressing the spurious non-diffuse contributions to ambient seismic field correlations, *J. geophys. Res.: Solid Earth,* **124**(9), 9653–9663.

Longuet-Higgins, M. S., 1950. A theory of the origin of microseisms, *Philos. Trans. R. Soc. Lond. Ser. A, Math. Phys. Sci.,* **243**(857), 1–35.

Meschede, M., Stutzmann, É., Farra, V., Schimmel, M. & Ardhuin, F., 2017. The effect of water column resonance on the spectra of secondary microseism P waves, *J. geophys. Res.: Solid Earth,* **122**(10), 8121–8142.

Montagner, J.-P. & Kennett, B. L. N., 1996. How to reconcile body-wave and normal-mode reference earth models, *Geophys. J. Int.,* **125**(1), 229–248.

Moresi, L. & Mather, B.R.,, 2019. Stripy: a Python module for (constrained) triangulation in cartesian coordinates and on a sphere, *J. Open Source Softw.,* **4**(38), 1410, doi:10.21105/joss.01410.

Nishida, K. & Takagi, R., 2016. Teleseismic S wave microseisms, *Science,* **353**(6302), 919–921.

Nishida, K., 2017. Ambient seismic wave field, *Proc. Japan Acad. Ser. B,* **93**(7), 423–448.

Nishida, K. & Takagi, R., 2022. A global centroid single force catalog of P-wave microseisms, *J. geophys. Res.: Solid Earth,* **127**(4), e2021JB023484, doi:10.1029/2021JB023484.

Nissen-Meyer, T., van Driel, M., Stähler, S. C., Hosseini, K., Hempel, S., Auer, L. & Fournier, A., 2014. AxiSEM: broadband 3-D seismic wavefields in axisymmetric media, *Solid Earth,* **5**(1), 425–445.

Obrebski, M., Ardhuin, F., Stutzmann, E. & Schimmel, M., 2013. Detection of microseismic compressional (P) body waves aided by numerical modeling of oceanic noise sources, *J. geophys. Res.: Solid Earth,* **118**(8), 4312–4324.

Osher, S. & Sethian, J. A., 1988. Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations, *J. Comput. Phys.,* **79**(1), 12–49.

Pedersen, H. & Colombi, A., 2018. Body waves from a single source area observed in noise correlations at arrival times of reflections from the 410 discontinuity, *Geophys. J. Int.,* **214**(2), 1125–1135.

Peterson, J. R., 1993. Observations and modeling of seismic background noise (No. 93-322), *US Geological Survey*.

Poli, P., Campillo, M. & Pedersen, H., 2012. Body-wave imaging of earth's mantle discontinuities from ambient seismic noise, *Science,* **338**(6110), 1063–1065.

Pyle, M. L., Koper, K. D., Euler, G. G. & Burlacu, R., 2015. Location of high-frequency P wave microseismic noise in the Pacific Ocean using multiple small aperture arrays, *Geophys. Res. Lett.,* **42**(8), 2700–2708.

Rascle, N. & Ardhuin, F., 2013. A global wave parameter database for geophysical applications. Part 2: model validation with improved source term parameterization, *Ocean Modell.,* **70,** 174–188.

Reading, A. M., Koper, K. D., Gal, M., Graham, L. S., Tkalčić, H. & Hemer, M. A., 2014. Dominant seismic noise sources in the Southern Ocean and West Pacific, 2000-2012, recorded at the Warramunga Seismic Array, Australia, *Geophys. Res. Lett.,* **41**(10), 3455–3463.

Retailleau, L. & Gualtieri, L., 2021. Multi-phase seismic source imprint of tropical cyclones, *Nat. Commun.,* **12**(1), 1–8.

Retailleau, L., Landès, M., Gualtieri, L., Shapiro, N. M., Campillo, M., Roux, P. & Guilbert, J., 2018. Detection and analysis of a transient energy burst with beamforming of multiple teleseismic phases, *Geophys. J. Int.,* **212**(1), 14–24.

Rost, S. & Thomas, C., 2002. Array seismology: methods and applications, *Rev. Geophys.,* **40**(3), 2–1.

Ruigrok, E., Draganov, D. & Wapenaar, K., 2008. Global-scale seismic interferometry: theory and numerical examples, *Geophys. Prospect.,* **56**(3), 395–417.

Sager, K., Tsai, V. C., Sheng, Y., Brenguier, F., Boué, P., Mordret, A. & Igel, H., 2022. Modelling P waves in seismic noise correlations: advancing fault monitoring using train traffic sources, *Geophys. J. Int.,* **228**(3), 1556–1567.

Stommel, H., 1957. A survey of ocean current theory, *Deep Sea Res. (1953),* **4,** 149–184.

Stutzmann, E., Ardhuin, F., Schimmel, M., Mangeney, A. & Patau, G., 2012. Modelling long-term seismic noise in various environments, *Geophys. J. Int.,* **191**(2), 707–722.

Toksöz, M. N. & Lacoss, R. T., 1968. Microseisms: mode structure and sources, *Science,* **159**(3817), 872–873.

Van Driel, M., Krischer, L., Stähler, S. C., Hosseini, K. & Nissen-Meyer, T., 2015. Instaseis: instant global seismograms based on a broadband waveform database, *Solid Earth,* **6**(2), 701–717.

Vinnik, L. P., 1973. Sources of microseismicP waves, *Pageoph,* **103,** 282–289.

Walker, K. T. & Shearer, P. M., 2009. Illuminating the near-sonic rupture velocities of the intracontinental KokoxiliMw7.8 and Denali faultMw7.9 strike-slip earthquakes with global P wave back projection imaging, *J. geophys. Res.,* **114**(B2), doi:10.1029/2008jb005738.

Wang, W., Gerstoft, P. & Wang, B., 2018. Seasonality of P wave microseisms from NCF-based beamforming using ChinArray, *Geophys. J. Int.,* **213**(3), 1832–1848.

Ward Neale, J., Harmon, N. & Srokosz, M., 2018. Improving microseismic P wave source location with multiple seismic arrays, *J. geophys. Res.: Solid Earth,* **123**(1), 476–492.

Xia, H. H., Song, X. & Wang, T., 2016. Extraction of triplicated PKP phases from noise correlations, *Geophys. J. Int.,* **205**(1), 499–508.

Xie, J., Luo, Y. & Yang, Y., 2021. Retrieving PmP travel times from a persistent localized microseismic source, *Geophys. Res. Lett.,* **48,** e2021GL094827, doi:10.1029/2021GL094827.

Xu, Y., Koper, K. D., Sufri, O., Zhu, L. & Hutko, A. R., 2009. Rupture imaging of the Mw 7.9 12 May 2008 Wenchuan earthquake from back projection of teleseismic P waves, *Geochem. Geophys. Geosyst.,* **10**(4), doi:10.1029/2008GC002335.

SUPPORTING INFORMATION

Supplementary data are available at *GJI* online.

Figure S1. (a) Slowness of *P* and *PP* phases. (b) *PP* backprojection of single pixel source modelled data. Modelled data are synthetized as depicted in Fig. 2, but with only one pixel source at (45° S, 40° E). Yellow curve is the great circle from the source to North America. In the middle of the curve (red box), the higher amplitude is an artefact caused by *P* phase when performing *PP* phase backprojection for North American stations.

Figure S2. Backprojection and stacking order maps for different distance interval and triangle area configurations.

Figure S3. (a) Events that cannot be detected in the *PP* backprojection shadow zones of the eastern United States and western Europe, but which can be detected by the *P* backprojection. (b) The strong events in the south of Greenland and Iceland are very clear in the *P* backprojection maps, but indistinguishable on the *PP* backprojection map. (c) The observability of event in south of Africa for *P* backprojection is even stronger than for *PP* backprojection. (d) Strong events in the southern Pacific can be detected by both *P* and *PP* backprojections. (e) Event in the south of Australia and New Zealand is clearly visible in *PP* backprojection map. But it cannot be distinguished in the *P* backprojection map.

Figure S4. Class-4 events on the global backprojection map. (a) and (b) The sources in the red boxes are obvious in the backprojection map of global stations. Their corresponding microseism events in (c) and (d) red boxes are defined as class-4.

Figure S5. Class-3 events on the global backprojection map, local arrays map. (a) *P* backprojection and (b) *PP* backprojection for global stations. *P* backprojection for arrays (c) CI, (d) YS and (e) YY. The microseism events in the red boxes of force field (f) are defined as class-3. These events have some imprints visible in the global *PP* backprojection but are not obvious. They can be confirmed with ocean wave model (f). In addition, they are clearly visible for the individual arrays.

Figure S6. Class-2 event on the local arrays backprojection map. (a) *P* backprojection and (b) *PP* backprojection for global stations. *P* backprojection for arrays (c) CI, (d) YS and (e) YY. The event is not detectable in the global backprojection map, but is clearly detectable for both the YS and YY arrays.

Figure S7. Class-1 events on the local arrays backprojection map. (a) *P* backprojection and (b) *PP* backprojection for global stations. *P* backprojection for arrays (c) CI, (d) YS and (e) YY. This event is weak for a single array YY.

Figure S8. Synthetic data backprojection maps of *P* phase for (a) network CI, (c) network IV and (e) network CI + IV. Synthetic data backprojection maps of *PP* phase for (b) network CI, (d) network IV and (f) network CI + IV.

Figure S9. Comparison of spectra of synthetic and observed waveform on 2019 August 9. (a) Spectra of beam trace. Delay and sum of all traces for point (45° S, 37.5° E). Spectra of single station (b) GR-GRA1 and (c) IC-LSA.

Figure S10. (a) Number of events with force $> 2 \times 10^{12}$ N. (b) Relation between the forces and cumulative number of events from 1994 to 2020. (c) Forces and surface areas of five classes microseism events of 2015 and 2020. The colour scale is the same as in Fig. 9. (d) Forces and surface areas of all the events from 1994 to 2020.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the paper.