Answer to reviewer 2

Dear Editor, dear authors,

Please consider my comments to the manuscript High Spatial and temporal 3D Sea Ice Thickness Reconstruction Using Seismic Waveform Inversion and Tomography submitted by Hooshmand Zandi et al.

The authors present an imaging study of sea ice. They apply concepts from seismology to two datasets collected in Canada and Svalbard/Norway to constrain space-time patterns of ice thickness. It is certainly one of the first if not the first application of a 2D ice thickness tomography constructed from dispersive vertical component flexural waves reported in the literature, which highlights the significance of the submitted work. The team uses controlled-source or hammer-shot data as well as natural icequake records for the inversion at the two sites. The use of the Spetral Element Method to create complex dispersive signals for the seismic waveform inversion of the water/ice system must be highlighted. The resulting images or distributions are compared against independently obtained thickness data, and the similarities and discrepancies are used to assess the inversion results and to discuss potential sources of bias.

I limit my comments to some general aspects, and I do not provide a line-by-line commentary of style and language that can be improved in places (e.g., lines 220 – 224 still contain place holders for values).

We thank the reviewer for the careful and constructive reading of the manuscript. We have answer all points, some of them required significant developments, which will be included in the revised manuscript and are, for the moment, only copied in this answer to the comments.

I am concerned with the tomography approach. I would appreciate a clarification about the connection between the physics of wave propagation of flexural modes guided in an ice layer with constant and with variable thickness, and the spatial discretization of the inverse problem. The authors write they adopt an approach that is similar to or like tomography:

Line 84: proposing 2D maps of the ice thickness, using a tomography-like approach

Line 85: a regionalization strategy similar to what is used for travel-time tomography

Line 250: a method similar to what is used for travel-time tomography with straight rays

Line 304: using a tomography-like approach

Because it is a pioneering ice thickness tomography study I think the presentation would benefit from spelling out better in what aspects the approach is the same as seismic travel time tomography, which aspects are similar, and how it is different. I understand that the pre-computed database of synthetic dispersive flexural waveforms is based on a model of the water/ice system with a uniform ice thickness along the source/reiver pair (step 1 of the algorithm). That uniform ice thickness is then estimated for each source-receiver pair using waveform inversion (step 4), which is then used in the tomographic inversion to estimate the spatial variability in thickness (step 5).

We thank the reviewer for these remarks/suggestions and agree that the manuscript should state more clearly how our approach relates to travel-time tomography. We have revised the manuscript to this end; the more specific points are addressed in detail in our answers to the following comments, where we also describe the corresponding changes made to the manuscript.

In brief, the approach is the same as travel-time tomography in its inversion step: each source–receiver path provides one path-averaged quantity, here the ice thickness, related linearly to the cell values through straight-ray weights, and the 2D map is obtained by regularized least squares. It is similar in that the path-averaged thickness plays the role of the travel time, although, as we show below, the quantity that averages linearly along a path is the wavenumber. The path value is therefore an equivalent (or apparent) thickness, which equals the path-averaged thickness up to a second-order correction, which we now derive in a new Appendix B.

However, it differs from tavel-time tomography in the nature of the data and in the forward problem: the thickness along each path is not picked but estimated by fitting the dispersive

flexural waveform with a database of waveforms computed for a uniform thickness, and in the passive case the sources must first be located.

The manuscript has been significantly modified, with Appendix, as well as new equations and and discussions directly in the main text. For this reason, we cannot list all changes compared to the previous version, but we will provide a version with highlighted changes.

Considering the observed dispersive waveforms are a result of the integrated propagation I wonder what the assumptions about the thickness-variability and the waveform characterstics are? Does the linearity of Eq 7 that we assume for travel times also hold for ice thickness obtained from fitting dispersive flexural waves?

This is a good question. Unlike in surface-wave tomography, where dispersion is controlled by the velocity structure with depth, the dispersion of flexural waves is controlled by the geometry of the waveguide: the flexural rigidity scales as h^3, so the wavenumber depends non-linearly on the thickness. Therefore, strictly speaking, the linearity of Eq. (11) of the manuscript holds for the wavenumber and not for the thickness itself. It holds for the thickness only up to a second-order term, which we quantify below.

The phase accumulated along a path is the integral of the local wavenumber $k(h,\omega)$, i.e. of the local phase slowness, as in surface-wave tomography (Woodhouse and Dziewonski, 1984; Kennett and Yoshizawa, 2002). The quantity that averages linearly along the path is therefore k, exactly as the slowness in travel-time tomography. The thickness $h_{\mathrm{obs},i}$ retrieved for path i is thus the thickness whose wavenumber equals the path-averaged wavenumber:

$$k\big(h_{\mathrm{obs},i},\omega\big) = \frac{1}{L_i}\int_{L_i} k\big(h(x,y),\omega\big)\,\mathrm{d}\ell. \tag{1}$$

For the frequencies of interest ($f \gtrsim 2$ Hz) the water is deep ($kH \gg 1$, so $\coth(kH) \simeq 1$) and the gravity term $g\,k$ is negligible compared with the flexural term $D\,k^5/\rho_w$. To leading order, Eq. (2) of the manuscript therefore reduces to $k(h,\omega) = \left[12(1-\nu^2)\rho_w\omega^2/E\right]^{1/5} h^{-3/5}$. Note that keeping the inertia of the ice in Eq. (2) changes the exponent only slightly, to between -0.59 and -0.57 from 2 to 40 Hz, and changes the correction below by about 1%.

Only the prefactor $\left[12(1-\nu^2)\rho_w\omega^2/E\right]^{1/5}$ depends on frequency, not the thickness. Since it appears on both sides of Eq. (1), it cancels, and Eq. (1) becomes

$$h_{\mathrm{obs},i} = \left[\frac{1}{L_i}\int_{L_i} h(x,y)^{-3/5}\,\mathrm{d}\ell\right]^{-5/3}, \tag{2}$$

Therfore, the "thickness-related slowness" is $s = h^{-3/5}$, since $s(h_{\mathrm{obs},i})$ is the path average of s. The tomographic inversion is thus of the travel-time type, with s playing the role of the slowness.

We can now directly compare Eq. (2) with Eq. (7) of the original manuscript. For a thickness varying linearly along the path from h_1 to h_2 (a typical case of our dataset), the integral in Eq. (2) can be computed exactly:

$$h_{\mathrm{obs},i} = \left[\frac{5}{2}\,\frac{h_2^{2/5} - h_1^{2/5}}{h_2 - h_1}\right]^{-5/3}. \tag{3}$$

To see how it differs from the path-averaged thickness $\bar{h} = (h_1 + h_2)/2$, we write $h_{1,2} = \bar{h}(1 \mp \varepsilon)$, with $\varepsilon = \Delta h/(2\bar{h})$ and $\Delta h = |h_2 - h_1|$ the total thickness change. A Taylor expansion of Eq. (3) gives:

$$h_{\mathrm{obs},i} = \bar{h}\left[1 - \frac{4}{15}\,\varepsilon^2 - 0.068\,\varepsilon^4 + \dots\right] \simeq \bar{h} - \frac{\Delta h^2}{15\,\bar{h}}. \tag{4}$$

The total relative change ε is not small compared with one ($\varepsilon \le 0.26$ for the thickness ranges of our data, 0.50–0.85 m), so we do not rely on a small-variation argument to justify the truncation. We instead checked it numerically against the exact Eq. (3): for $\varepsilon \le 0.26$ the two differ by about

0.2 mm at most, and the relative bias does not exceed 1.8%. Even for a stronger contrast, 0.80 → 0.35 m ($\varepsilon = 0.39$), the correction is 23.5 mm with Eq. (4) against 24.5 mm with Eq. (3).

Equation (7) in our original manuscript ($h_{\mathrm{obs},i} = \frac{1}{L_i}\int_{L_i} h(x,y)\,\mathrm{d}\ell$), is therefore the leading-order approximation of Eq. (2), and the neglected term is of second order in the thickness variations. For the contrasts of our dataset it does not exceed 1.2 cm (less than 2%, Table 1). We therefore keep Eq.(7) in the revised manuscript, but clarify that $h_{\mathrm{obs},i}$ is an equivalent thickness, which equals the path average up to this second-order correction. This analysis was added to the revised manuscript in Section 2.3.2 and is copied below.

Allowing a different thickness on each source–receiver path refines the level of detail with respect to Moreau et al. (2023), who assumed a single thickness for all paths of an event. However, each path thickness is still obtained by fitting a homogeneous-slab model to a waveform that has propagated through ice of varying thickness. The retrieved value is therefore an equivalent, or apparent, thickness. This is valid under the assumptions that the thickness varies slowly along the path. The phase of the flexural wave is then the integral of the local wavenumber, provided that the local wavenumber changes little over one wavelength λ and that reflections and mode conversions are negligible:

$$\frac{\lambda}{k}\left|\frac{\mathrm{d}k}{\mathrm{d}\ell}\right| \ll 1, \tag{7}$$

where ℓ is the curvilinear coordinate along the path. For the frequencies in this study, the water is deep enough for the gravity term of Eq. (2) to be negligible, so that:

$$k(h,\omega) = \left[12(1-\nu^2)\rho_w\omega^2/E\right]^{1/5} h^{-3/5} \tag{8}$$

Since the prefactor does not depend on the thickness, $\mathrm{d}k/k = -(3/5)\,\mathrm{d}h/h$, and Eq. (7) is equivalent to a condition on the relative thickness gradient:

$$\frac{\lambda}{h}\left|\frac{\mathrm{d}h}{\mathrm{d}\ell}\right| \ll 1. \tag{9}$$

For our data, with wavelengths of 8 to 18 m and thickness changes of 0.35 m or less over 200 to 600 m, this thickness-gradient parameter is about 0.05 or less.

In Appendix B, we demonstrate that, under this condition, the equivalent thickness is:

$$h_{\mathrm{eq}} \simeq \bar{h} + \frac{\Delta h^2}{24}\left.\frac{\partial^2 k/\partial h^2}{\partial k/\partial h}\right|_{\bar{h}}, \tag{10}$$

where $\bar{h}$ is the average thickness along the path. The numerical verification and the limits of this approximation are discussed in Appendix B, and our synthetic waveform tests with linear profiles confirm that the thickness fitted with a homogeneous slab agrees with the equivalent thickness within a fraction of a millimeter.

Because the wavenumber is a convex function of the thickness, the equivalent thickness is always smaller than the path-averaged thickness. To second order, the difference is $4\sigma_h^2/(5\bar{h})$, where σ_h^2 is the variance of the thickness along the path. For the thickness ranges of our datasets (0.50 to 0.85 m over 200 to 600 m), this term does not exceed 1.2 cm (less than 2%).

Say the inversion yields a spatially variable thickness profile for a source-receiver pair, is the associated waveform propagating along that inferred profile the same as the signal propagating along a homogeneous slab with an 'equivalent' average or effective thickness? What is this equivalence? This aspect should at least be demonstrated with a synthetic waveform modelling example or otherwise confirmed, or the limits of this assumption should be discussed.

This question relates closely to the previous one, and was also raised by the other reviewer, as well as by one reviewer of Moreau et al. (2023). We have therefore added the follwoing analysis to the revised manuscript so that the ambiguity is definitely removed.

We assume here that by inversion, the reviewer refers to the final result of the tomography, since thickness retrieval for one source-receiver pair is always constant in our approach (although we could try in the future to fit directly a varying thickness-related waveform).

From a general point of view, the reviewer is correct: the thickness retrieved by fitting a constant-thickness model to a waveform that has propagated through ice of varying thickness is not, in general, the path-averaged thickness. It is an *equivalent* thickness, defined by the condition that the accumulated phase is the same as in a homogeneous slab. As explained in the answer to the previous question, for the slowly varying thickness profiles of our dataset, it differs from the path-averaged thickness by at most 1.2 cm.

For a given angular frequency ω, the wavenumber $k(h,\omega)$ of the flexural-gravity wave in ice of thickness h is the root of the dispersion relation, Eq. (2) of the manuscript. When the thickness varies slowly along a path of length L, $h = h(x)$ with $0 \le x \le L$, the local wavenumber is $k(h(x),\omega)$, and the phase accumulated over the path is the integral of the local wavenumber, as in the path-average treatment of surface waves (Woodhouse and Dziewonski, 1984; Kennett and Yoshizawa, 2002) and for guided modes in waveguides of linearly varying thickness (Moreau et al., 2014):

$$\varphi(\omega) = \int_0^L k\big(h(x),\omega\big)\,\mathrm{d}x. \tag{5}$$

A constant-thickness model with thickness h_c gives $\varphi_c(\omega) = k(h_c,\omega)\,L$. The two have the same dispersion (hence the same waveform shape) if

$$k\big(h_{\mathrm{eq}},\omega\big) = \frac{1}{L}\int_0^L k\big(h(x),\omega\big)\,\mathrm{d}x = \langle k\rangle, \tag{6}$$

where $\langle\cdot\rangle$ is the average along the path. This is Eq. (1) of our previous answer, with $h_{\mathrm{eq}} = h_{\mathrm{obs},i}$ for the path considered. The equivalent thickness h_{eq} is therefore the thickness whose wavenumber equals the *path average of the wavenumber* (the analogue of the average slowness of Kennett and Yoshizawa, 2002; Boiero and Socco, 2011), not the thickness whose value equals the path average of the thickness. Because k is a nonlinear function of h, these two averages differ in general. The difference is small when the thickness varies slowly along the path, so that perturbation theory can be applied (Kennett and Yoshizawa, 2002). The same distinction between an "equivalent" and a "mean" property arises in near-surface surface-wave analysis: Boiero and Socco (2011) showed that, when the phase slowness varies linearly along the path, the dispersion curve obtained from an array corresponds to the average slowness beneath the array, whereas for non-linear variations it may not correspond to any specific subsurface velocity model.

Assume a linearly varying thickness profile, as in Moreau et al. (2014), with total change $\Delta h = |h(L) - h(0)|$ and path-averaged thickness $\bar{h} = \langle h\rangle$. The local deviation $h(x) - \bar{h}$ takes all values between $-\Delta h/2$ and $\Delta h/2$ with equal weight along the path, so that its average is zero and its variance is $\sigma_h^2 = \langle (h-\bar{h})^2\rangle = \Delta h^2/12$. A second-order Taylor expansion of k about $\bar{h}$ gives, at each position x,

$$k\big(h(x),\omega\big) \simeq k(\bar{h}) + \left.\frac{\partial k}{\partial h}\right|_{\bar{h}} \big[h(x)-\bar{h}\big] + \frac{1}{2}\left.\frac{\partial^2 k}{\partial h^2}\right|_{\bar{h}} \big[h(x)-\bar{h}\big]^2, \tag{7}$$

Averaging Eq. (7) along the path, the first-order term vanishes because $\langle h - \bar{h}\rangle = 0$, and

$$\langle k\rangle \simeq k(\bar{h}) + \frac{1}{2}\left.\frac{\partial^2 k}{\partial h^2}\right|_{\bar{h}} \sigma_h^2 = k(\bar{h}) + \frac{\Delta h^2}{24}\left.\frac{\partial^2 k}{\partial h^2}\right|_{\bar{h}}. \tag{8}$$

Expanding the left-hand side of Eq. (6) to first order, $k(h_{\mathrm{eq}}) \simeq k(\bar{h}) + \left.\frac{\partial k}{\partial h}\right|_{\bar{h}} (h_{\mathrm{eq}} - \bar{h})$, and equating both expressions gives

$$\boxed{h_{\mathrm{eq}} \simeq \bar{h} + \frac{\Delta h^2}{24}\left.\frac{\partial^2 k/\partial h^2}{\partial k/\partial h}\right|_{\bar{h}}} \tag{9}$$

where the derivatives, evaluated at $\bar{h}$, follow from implicit differentiation of Eq. (2) of the manuscript.

The correction is even in Δh, so it is identical for thickness increasing or decreasing along the path. Since $k \propto h^{-3/5}$ in the flexural regime (see our answer to the previous comment), k decreases with h ($\partial k/\partial h < 0$) and is convex ($\partial^2 k/\partial h^2 > 0$). It follows that $h_{\mathrm{eq}} \leq \bar{h}$ for any thickness profile, not only for a linear one: since $k(h)$ is convex, it lies above its tangent at $\bar{h}$, $k(h) \geq k(\bar{h}) + \frac{\partial k}{\partial h}\big|_{\bar{h}} (h - \bar{h})$. Averaging along the path, where $\langle h - \bar{h} \rangle = 0$, gives $\langle k \rangle \geq k(\bar{h})$, and, because k decreases with h, $h_{\mathrm{eq}} \leq \bar{h}$.

In practice, the fit is performed on a band-limited pulse and not at a single frequency. Equation (6) holds at each frequency, so the equivalent thickness is in principle frequency dependent, $h_{\mathrm{eq}}(\omega)$, and the thickness retrieved by the fit is a weighted average of $h_{\mathrm{eq}}(\omega)$ over the source spectrum. This weighting has a negligible effect here because the correction of Eq. (9) depends on frequency only through the exponent of $k(h)$, which varies very little over the band. In the flexural regime $k \propto h^{-3/5}$, for which

$$\frac{\partial^2 k/\partial h^2}{\partial k/\partial h} \simeq -\frac{8}{5\,\bar{h}} \quad \Rightarrow \quad h_{\mathrm{eq}} \simeq \bar{h} - \frac{\Delta h^2}{15\,\bar{h}}, \tag{10}$$

which does not depend on ω and is the same as Eq. (4) with $\sigma_h^2 = \Delta h^2/12$. As explained previously, the inertia of the ice, neglected in the pure flexural regime, makes the local exponent vary slightly with frequency, because the ratio of the inertia of the ice to that of the water loading increases with the wavenumber. Numerically, evaluating Eq. (9) with the full dispersion relation, Eq. (2), from 5 to 40 Hz, the correction varies by less than 0.1 mm (for $\bar{h} = 0.675$ m and $\Delta h = 0.35$ m it is -12.0 mm over the whole band, against -12.1 mm for the pure power law). Evaluating Eq. (9) at the central frequency of the pulse therefore reproduces the fitted value closely, as confirmed by the numerical tests below.

We generated waveforms for a linear thickness profile with the perturbation of Eq. (2) of the manuscript and searched, with a constant-thickness model, for the thickness that maximizes the normalized cross-correlation . We used the following settings. Source: Gaussian-windowed pulse centered at $f_0 = 5$, 10 and 15 Hz; propagation distances $L = 200$ to 600 m; $H = 5$ m, $E = 3.8$ GPa, $\nu = 0.321$, $\rho_i = 910$, $\rho_w = 1010$ kg m^{-3}. The fitted thickness changed by less than 0.2 mm between increasing and decreasing profiles, and by less than 0.05 mm between distances and frequencies (Table 1).

Case	$\bar{h}$ (m)	Δh (m)	$\bar{h} - h_{\mathrm{eq}}$ fitted (mm)	$\bar{h} - h_{\mathrm{eq}}$ Eq. (9) (mm)
0.725 → 0.625 m	0.675	0.10	1.0	1.0
0.750 → 0.600 m	0.675	0.15	2.2	2.2
0.775 → 0.575 m	0.675	0.20	3.9	3.9
0.800 → 0.550 m	0.675	0.25	6.2	6.1
0.825 → 0.525 m	0.675	0.30	8.9	8.8
0.850 → 0.500 m	0.675	0.35	12.2	12.0

Table 1: Bias of the path-averaged thickness with respect to the fitted equivalent thickness, for a thickness decreasing along the path. For each contrast Δh, 15 simulations were run ($L = 200$, 300, 400, 500, 600 m and $f_0 = 5$, 10, 15 Hz). The fitted bias is the mean over these 15 cases; it varies by less than 0.05 mm between cases. The second-order estimate of Eq. (9), evaluated with the full dispersion relation, does not depend on L or f_0, and differs from the fitted value by less than 0.25 mm in all cases.

For the configurations typical of our dataset, and under the slowly varying thickness assumption, the equivalent thickness is therefore smaller than the mean thickness by at most 1.2 cm

(less than 2%) for the sharpest contrast $0.50 \rightarrow 0.85$ m. Identifying the retrieved thickness with the path-averaged thickness thus introduces a systematic underestimation of the mean thickness of at most this size. Its sign is known, and it can be corrected with Eq. (9). The corresponding correction therefore does not affect our conclusions.

The bias grows as Δh^2 and, at fixed Δh, as the mean thickness decreases (it depends on the relative contrast $\Delta h/\bar{h}$): for thinner ice and stronger contrasts, such as $0.30 \rightarrow 0.80$ m and $0.20 \rightarrow 0.70$ m over 200 m, we find a bias of about 3 cm and 4 cm (6% and 9%), and the second-order formula then deviates from the fitted value by 1.6 and 3 mm. Given that, in our tomography results, the transition between thick ice near the shore and thinner ice towards the center seems sharper near the shore, we cannot exclude that such gradients are present in our data. This could explain the underestimation of our inverted thickness when compared to Serriperri et al. (2022). Our numerical tests assume a linear thickness profile, a slowly varying thickness (no mode conversion or reflection) and noise-free synthetic waveforms. For non-linear thickness profiles, Eq. (2) or Eq. (6) can be used directly instead of Eq. (9) for the inversions, but this is out of the scope of our manuscript.

In complement to the changes made to Section 2.3.2 of the manuscript, we have added the following appendix B to include this analysis.

Appendix B: Equivalent thickness and path-averaged thickness

This appendix details the relation between the thickness retrieved by fitting a constant-thickness model to a waveform that has propagated through ice of varying thickness (the equivalent thickness $h_{\mathrm{eq}} = h_{\mathrm{obs},i}$) and the path-averaged thickness $\bar{h} = \langle h \rangle$ used in Eq. (11).

B1: Definition of the equivalent thickness

For a given angular frequency ω, the wavenumber $k(h, \omega)$ of the flexural–gravity wave in ice of thickness h is the root of the dispersion relation, Eq. (2). When the thickness varies slowly along a path of length L, $h = h(x)$ with $0 \leq x \leq L$, the phase accumulated over the path is the integral of the local wavenumber as in the path-average treatment of surface waves (Woodhouse and Dziewonski, 1984; Kennett and Yoshizawa, 2002), and for guided modes in waveguides of linearly varying thickness (Moreau et al., 2014):

$$\varphi(\omega) = \int_0^L k\big(h(x), \omega\big)\, \mathrm{d}x. \tag{23}$$

A constant-thickness model with thickness h_c gives $\varphi_c(\omega) = k(h_c, \omega)\, L$. Both have the same dispersion, and hence the same waveform shape, if

$$k\big(h_{\mathrm{eq}}, \omega\big) = \frac{1}{L} \int_0^L k\big(h(x), \omega\big)\, \mathrm{d}x = \langle k \rangle, \tag{24}$$

where $\langle \cdot \rangle$ is the average along the path. This is Eq. (12) with $h_{\mathrm{eq}} = h_{\mathrm{obs},i}$. The equivalent thickness is therefore the thickness whose wavenumber equals the path average of the wavenumber, and not the thickness whose value equals the path average of the thickness. Because k is a nonlinear function of h, these two averages differ in general. The same distinction arises in near-surface surface-wave analysis (Boiero and Socco, 2011).

B2: Exact result for a linear profile and second-order expansion

In the flexural regime, Eq. (2) reduces to $k \propto h^{-3/5}$, and Eq. (24) leads to the power mean of Eq. (13). For a thickness varying linearly from h_1 to h_2 along the path, the integral can be

computed exactly:

$$h_{\mathrm{eq}} = \left[\frac{5}{2} \frac{h_2^{2/5} - h_1^{2/5}}{h_2 - h_1} \right]^{-5/3} . \tag{25}$$

We write $h_{1,2} = \bar{h}(1 \mp \varepsilon)$, with $\varepsilon = \Delta h/(2\bar{h})$ and $\Delta h = |h_2 - h_1|$. Exchanging h_1 and h_2 leaves the result unchanged, so only even powers of ε appear:

$$h_{\mathrm{eq}} = \bar{h} \left[1 - \frac{4}{15} \varepsilon^2 - 0.068\, \varepsilon^4 + \dots \right] \simeq \bar{h} - \frac{\Delta h^2}{15\, \bar{h}}. \tag{26}$$

The total relative change ε is not small compared with one ($\varepsilon \leq 0.26$ for the thickness ranges of our data), so the truncation is not justified by a small-variation argument but was checked numerically against the exact Eq. (25): for $\varepsilon \leq 0.26$ the two differ by about 0.2 mm at most (the term in ε^4 is less than 2% of the one retained), and the relative bias does not exceed 1.8%. Even for a stronger contrast, 0.80 to 0.35 m ($\varepsilon = 0.39$), the correction is 23.5 mm with Eq. (26) against 24.5 mm with Eq. (25).

The same result follows from a perturbation expansion of k about $\bar{h}$ for the linear profile, whose local deviation $h(x) - \bar{h}$ takes all values between $-\Delta h/2$ and $\Delta h/2$ with equal weight, so that its path average is zero and its variance is $\sigma_h^2 = \Delta h^2/12$:

$$k\big(h(x), \omega\big) \simeq k(\bar{h}) + \left.\frac{\partial k}{\partial h}\right|_{\bar{h}} \big[h(x) - \bar{h}\big] + \frac{1}{2} \left.\frac{\partial^2 k}{\partial h^2}\right|_{\bar{h}} \big[h(x) - \bar{h}\big]^2. \tag{27}$$

Averaging along the path, the first-order term vanishes and $\langle k \rangle \simeq k(\bar{h}) + \frac{\Delta h^2}{24}\, \partial^2 k/\partial h^2$. Expanding the left-hand side of Eq. (24) to first order and equating both expressions gives

$$h_{\mathrm{eq}} \simeq \bar{h} + \frac{\Delta h^2}{24} \left.\frac{\partial^2 k/\partial h^2}{\partial k/\partial h}\right|_{\bar{h}}, \tag{28}$$

where the derivatives follow from implicit differentiation of Eq. (2). For $k \propto h^{-3/5}$ the ratio of derivatives is $-8/(5\bar{h})$, which gives back Eq. (26). The correction is even in Δh, so it is identical for a thickness increasing or decreasing along the path. Moreover, k decreases with h and is convex, so that $k(h) \geq k(\bar{h}) + \partial k/\partial h|_{\bar{h}}\, (h - \bar{h})$. Averaging along the path gives $\langle k \rangle \geq k(\bar{h})$ and hence $h_{\mathrm{eq}} \leq \bar{h}$ for any thickness profile, not only for a linear one.

B3: Effect of the source spectrum

The fit is performed on a band-limited pulse and not at a single frequency. Equation (24) holds at each frequency, so the equivalent thickness is in principle frequency dependent, and the fitted thickness is a weighted average of $h_{\mathrm{eq}}(\omega)$ over the source spectrum. This weighting has a negligible effect because the correction of Eq. (28) depends on frequency only through the local exponent of $k(h)$. The inertia of the ice makes this exponent vary slightly with frequency, from -0.58 to -0.57 between 5 and 40 Hz, because the ratio of the inertia of the ice to that of the water loading increases with the wavenumber. Evaluating Eq. (28) with the full dispersion relation, Eq. (2), from 5 to 40 Hz, the correction varies by less than 0.1 mm (for $\bar{h} = 0.675$ m and $\Delta h = 0.35$ m it is -12.0 mm over the whole band, against -12.1 mm for the pure power law).

B4: Numerical verification with synthetic waveforms

We generated waveforms for a linear thickness profile by integrating the local wavenumber obtained from Eq. (2) along x, and searched, with a constant-thickness model, for the thickness that maximizes the normalized cross-correlation. The source was a Gaussian-windowed pulse centered at $f_0 = 5$, 10 and 15 Hz, with propagation distances $L = 200$ to 600 m, and $H = 5$ m,

$E = 3.8$ GPa, $\nu = 0.321$, $\rho_i = 910$ and $\rho_w = 1010$ kg m^{-3}. The fitted thickness changed by less than 0.2 mm between increasing and decreasing profiles, and by less than 0.05 mm between distances and frequencies (Table B-1).

Profile	$\bar{h}$ (m)	Δh (m)	$\bar{h} - h_{\mathrm{eq}}$ fitted (mm)	$\bar{h} - h_{\mathrm{eq}}$ Eq. (28) (mm)
0.725 → 0.625 m	0.675	0.10	1.0	1.0
0.750 → 0.600 m	0.675	0.15	2.2	2.2
0.775 → 0.575 m	0.675	0.20	3.9	3.9
0.800 → 0.550 m	0.675	0.25	6.2	6.1
0.825 → 0.525 m	0.675	0.30	8.9	8.8
0.850 → 0.500 m	0.675	0.35	12.2	12.0

Table B-1: Bias of the path-averaged thickness with respect to the fitted equivalent thickness, for a thickness decreasing along the path. For each contrast, 15 simulations were run ($L = 200$, 300, 400, 500, 600 m and $f_0 = 5$, 10, 15 Hz). The fitted bias is the mean over these 15 cases; it varies by less than 0.05 mm between cases. The second-order estimate of Eq. (28), evaluated with the full dispersion relation, does not depend on L or f_0, and differs from the fitted value by less than 0.25 mm in all cases.

For the sharpest contrast of our data (0.50 to 0.85 m), the equivalent thickness is therefore smaller than the mean thickness by at most 1.2 cm (less than 2%). Identifying the retrieved thickness with the path-averaged thickness thus introduces a systematic underestimation of at most this size. Its sign is known, and it can be corrected with Eq. (28).

B5: Limits

The bias grows as Δh^2 and, at fixed Δh, as the mean thickness decreases (it depends on the relative contrast $\Delta h/\bar{h}$). For thinner ice and stronger contrasts, such as 0.30 to 0.80 m and 0.20 to 0.70 m over 200 m, we find a bias of about 3 and 4 cm (6% and 9%), and the second-order formula then deviates from the fitted value by 1.6 and 3 mm. These cases are outside the range of our data. The tests also assume a linear thickness profile, a slowly varying thickness (no mode conversion or reflection) and noise-free synthetic waveforms. For non-linear profiles, Eq. (13) or Eq. (24) can be used directly instead of Eq. (28); the sign of the bias remains the same by convexity of k.

This then affects the resolution of the spatial variability, which is another term or concept that should be discussed in more depth. What is the minium length scale over which a thickness variation can be resolved? As in seismic travel time tomography, a connection to the involved wavelength(s) should exist, so I assume reporting the wavelength range of the dispersive signals is also relevant for the discussion.

We thank the reviewer for this remark. The icequake waveforms considered in this study are typically dominated by frequencies between 5 and 15 Hz, corresponding to wavelengths of approximately 8–18 m.

The spatial resolution of the tomographic images cannot, however, be inferred from the wavelength alone. It depends on the spatial sensitivity of the measurements, the density and azimuthal coverage of the source–receiver paths, measurement uncertainties, and the regularization of the inversion. In finite-frequency travel-time tomography, a measurement is sensitive to a region around the geometrical ray rather than exclusively to the ray itself. An indicative transverse scale is provided by the first Fresnel zone, whose diameter at the midpoint of a path of length L is approximately $D_F \simeq \sqrt{\lambda L}$. For wavelengths of 10–15 m and path lengths of 200–600 m, this gives diameters of approximately 45–95 m. This estimate illustrates the poten-

tial importance of finite-frequency effects; it should not be interpreted as a strict lower bound on tomographic resolution.

In this study, the inversion is based on straight-ray sensitivity weights and does not explicitly account for finite-frequency effects. Consequently, the recovered thickness variations should be interpreted in light of both the path geometry and the spatial resolution assessed through synthetic tests. Multiple paths with different azimuths and lengths help constrain the spatial distribution of thickness, although they do not by themselves guarantee resolution below the characteristic finite-frequency sensitivity scale. Our checkerboard tests, presented in the response to the next comment, provide an empirical assessment of the length scales that can be recovered for the acquisition geometry and inversion procedure considered here. The corresponding figures show the scale of the thickness variations that can be recovered in our datasets.

To cover all these points, we have completely re-written the discussion section to include these comments to the manuscript.

A new Section 4.2, "Array geometry and up-scaling", has been added to the manuscript to address both comments, and is copied thereafter. It discusses the influence of station geometry on ray coverage and tomographic resolution, the limitations of the original Svalbard array, and the potential benefits of more spatially distributed and clustered deployments for large-scale ice monitoring. It also presents synthetic checkerboard tests comparing the original configuration with a more spatially distributed array. The full modified Discussion section is given below.

4 Discussion

The method introduced in this paper appears very promising for 3D mapping of sea ice thickness. However, some questions remain regarding its limitations and potential improvements.

4.1 Comparison with independent estimates and origin of the Svalbard offset

For the dataset recorded in Canada with controlled sources, the thickness estimates obtained from tomographic inversions are consistent with those derived from frequency–wavenumber spectrum inversions, suggesting that both methods yield comparable results under controlled acquisition conditions. In contrast, for the dataset recorded in Svalbard with natural sources, the two methods produce slightly different thickness estimates.

Our tomographic thicknesses in the dense-array area are systematically lower than those of Serripierri et al. (2022), by about 5–10 cm, and both are lower than the drillings (62 cm on 1 March and 70–73 cm on 26 March). The drilling offset is consistent with the porous top 10–15 cm of the ice (compacted snow), which softens the flexural wave and lowers the apparent thickness. This layer affects both methods, but does not explain the difference between them. Three effects may contribute to this difference:

- The source positions are constrained to lie on the ice, whereas the initial catalog contains events located outside the fjord. Moving them inside shortens the path. Since the accumulated phase $k(h)\,L$ is fixed by the waveform and $k \propto h^{-3/5}$ in the flexural regime, the retrieved thickness varies as $L^{5/3}$, and a shorter path gives a thinner ice. The constraint also removes location errors towards the outside but retains those towards the inside, so that the remaining errors are not centred on zero. For random errors of 10–25 m on paths of 200–300 m, this yields a mean underestimation of about 1.5–5.6 cm. This systematic bias should therefore be considered when interpreting small-scale thickness variations, particularly near the shoreline, where icequake sources are concentrated and ray coverage may be less uniform.

- The thickness retrieved along a path is an equivalent thickness, not the path-averaged thickness of Eq. (11) (Appendix B). This introduces an underestimation of the mean

thickness of at most 1.2 cm for the contrasts observed in our dataset. This effect does not apply in the same way to Serripierri et al. (2022), who fit the wavenumbers directly in the frequency–wavenumber spectrum over an aperture of about 50 m. The corresponding bias scales approximately with the square of the thickness change over this aperture: a change of 0.35 m over 300 m corresponds to about 6 cm over 50 m, giving a bias of the order of 0.4 mm, which is negligible.

- The two methods also differ in their spatial and temporal averaging. The estimates of Serripierri et al. (2022) are obtained from noise correlations accumulated over several hours and from sensor pairs separated by less than 150 m. The spatial Fourier transform further averages the wavefield over an aperture of about 50 m. In contrast, our estimates are obtained from individual icequake paths extending over several hundred metres, mostly from the shoreline towards the array. The two methods therefore sample the ice at different spatial scales and with different temporal averaging.

These differences in source geometry, the definition of the retrieved thickness, and the spatial and temporal averaging provide plausible explanations for part of the offset, but we have not demonstrated that they account for all of it. The temporal and spatial smoothing inherent in the noise-correlation estimates may also contribute to the differences in the observed trends: the estimates from the two methods may reflect different degrees of smoothing of the underlying thickness variations.

The Canadian experiment provides a useful comparison. There, the sources surround the array and are close to it, and no comparable offset is observed between the tomographic and frequency–wavenumber estimates along the fibre. This suggests that the source geometry and the associated path coverage may contribute to the discrepancy observed in Svalbard, rather than the difference arising from the principle of the tomographic method itself.

The accuracy of the tomographic inversion depends on the density and spatial distribution of ray paths crossing each cell, which in this case are controlled by the number and locations of icequakes. These vary substantially from one day to the next. Although incorporating prior information from the previous day mitigates this effect, variations in ray coverage may still affect the estimated thickness. This dependence on acquisition geometry is particularly important when considering how to improve the spatial resolution and extend the method to larger areas, as discussed in the following section. A possible extension would be to perform tomography using noise correlations between all pairs of geophones, thereby increasing the number of inter-sensor paths within the dense array. However, the main objective of this paper is to demonstrate the potential for thickness estimation over larger distances, of the order of 1 km, rather than to optimize small-scale imaging over apertures of approximately 50 m.

4.2 Array geometry and up-scaling

The dense Svalbard array was designed as a proof of concept to show that passive seismology can provide information on sea ice comparable to that obtained from active experiments. At the time, we anticipated that heterogeneities, cracks, attenuation, and porosity would prevent waveforms from propagating more than a few tens of metres from their sources, as is the case in the jumbled ice of the Canada experiment. The array was therefore designed for small-scale monitoring using conventional array methods: noise interferometry to recover the Green's function between sensor pairs, tomography based on inter-sensor paths, and beamforming to characterize the azimuthal distribution of the wavefield and the source mechanisms (Moreau et al., 2020a; Serripierri et al., 2022). These methods require a high density of sensors. Surprisingly, the data contained thousands of icequakes recorded at distances of a few hundred metres, opening the possibility of source localization and tomography at a larger scale. The array geometry was, however, not optimal for this purpose.

For event localization and path-thickness estimation, such a high station density is not required: five stations are sufficient for triangulation (Moreau et al., 2020b, 2023), and we used five stations selected for their large aperture and geometrical coverage. The other stations provide additional thickness estimates along nearly parallel paths, since they are close to each other. These additional measurements mainly help reduce the noise on the estimated thickness, while improving azimuthal coverage only marginally. Together with the sources, which are located mostly along the shoreline, a compact array produces rays that converge towards the array and images that are elongated along the ray directions.

The interpretation of small-scale thickness variations near the shoreline requires particular caution. Although many icequake sources are located along the shoreline, the resulting ray paths do not necessarily provide uniform spatial coverage or a broad range of azimuths near the sources. Moreover, rays originating from nearby sources may share common path segments, so that errors associated with these sources may be mapped into the same region rather than averaging out. Consequently, thickness estimates near the shoreline are less well constrained than in the central part of the array, where rays cross at different azimuths. Source-location errors may contribute to small-scale artefacts, although their estimated magnitude is insufficient to explain the largest observed thickness contrasts.

The ability to constrain a region depends not only on the number of paths but also on how they traverse the regions to be distinguished. Paths that do not cross a thickness contrast constrain the thickness along their trajectories but provide little information on the contrast between regions. The contrast between two regions, and the thickness of a region without receivers, can be constrained by paths that cross both regions. A region without any receiver can therefore be constrained if sources are located within it and if the fraction γ of each path lying in this region varies between paths. In this case, an uncertainty δh_{obs} on the observed path thickness translates approximately into an uncertainty of $\delta h_{\mathrm{obs}}/\gamma$ on the thickness of the region, which increases as the fraction decreases.

The spatial resolution also depends on the wavelength of the flexural waves. For the ice properties listed in the table of mechanical parameters for Svalbard, the wavelength varies from approximately 20 m at 2 Hz to 5 m at 50 Hz. The icequake waveforms considered here are typically dominated by frequencies between 5 and 15 Hz, corresponding to wavelengths of approximately 8–18 m. These wavelengths provide a relevant physical scale for discussing spatial sensitivity, but the resolution of a tomographic image cannot be inferred from wavelength alone.

In finite-frequency travel-time tomography, a measurement is sensitive to a region around the geometrical ray rather than exclusively to the ray itself. An indicative transverse scale is provided by the first Fresnel zone, whose diameter at the midpoint of a source–receiver path of length L is approximately $D_F \simeq \sqrt{\lambda L}$. For wavelengths of 10–15 m and path lengths of 200–600 m, this gives diameters of approximately 45–95 m. This estimate illustrates the potential importance of finite-frequency effects, but it should not be interpreted as a strict lower bound on the tomographic resolution. The effective resolution also depends on the distribution and azimuthal diversity of the paths, measurement uncertainties, and the regularization and parameterization of the inversion.

Our inversion uses straight-ray sensitivity weights and does not explicitly account for finite-frequency effects. Consequently, narrow thickness anomalies located close to, but not intersected by, a geometrical ray may have a weaker influence on the measured waveform than predicted by the straight-ray approximation. Conversely, a set of paths with different azimuths and lengths can provide complementary constraints on the spatial distribution of thickness. The length scale that can actually be recovered must therefore be assessed for the acquisition geometry and inversion procedure considered here, rather than inferred from wavelength or ray density alone.

To illustrate the influence of station geometry on the recovery of spatial variations, we performed synthetic checkerboard tests (Figures 1 and 2). Figure 1 shows the checkerboard model used to generate the synthetic thickness observations. In Figure 2a, the stations are those of the

original configuration used in the Svalbard experiment, whereas in Figure 2b they are relocated to provide broader and more uniform coverage of the study area. The corresponding tomographic results are shown in Figures 2c and d. The broader geometry leads to a substantially better recovery of the checkerboard pattern, demonstrating the strong influence of station placement on ray coverage and inversion quality. These tests provide an empirical assessment of the spatial scales recoverable with the tested geometries and inversion settings.

The optimal design depends on the size of the monitored area, from a few hundred metres to several kilometres, on the propagation distance and source distribution, and on the desired spatial resolution, but also on the ability of an icequake to propagate over long distances. Finally, long paths cross more thickness variations and cracks, making the assumption of slowly varying thickness (Section **??**) more difficult to satisfy. Shorter paths are therefore more likely to meet this assumption, but require more stations or clusters to cover a given area. Determining the minimum number of stations and their optimal distribution for a given area and target resolution requires a dedicated study.

4.3 Limitations and perspectives

Potential improvements to our approach could involve exploring other types of cost functions. The use of cross-correlations of signals is known to be non-convex, owing to the oscillatory behavior of the waveforms. Alternative misfit functions, for example those based on spectrogram correlations or on optimal transport distances (Métivier et al., 2022), could improve the convergence of the inversion in steps 2–4 of the procedure.

An important limitation of the present approach is that the tomographic inversion assumes straight-ray propagation and represents each source–receiver path by a single equivalent thickness. Although the checkerboard tests assess the effects of ray coverage and regularization within this framework, they do not independently validate the approximation for flexural waves propagating through ice of spatially variable thickness. As discussed in Appendix B, the equivalent thickness differs from the path-averaged thickness by a second-order correction, which is less than 1.2 cm for the thickness contrasts considered here. However, this analysis assumes slowly varying thickness and does not account for possible wave scattering, reflection, or mode conversion caused by lateral heterogeneity. A complete assessment of these effects would require full-wave simulations in a laterally heterogeneous ice layer, which are beyond the scope of this study. The recovered thickness variations should therefore be interpreted with these limitations in mind.

A further limitation is that the waveform inversion uses 2D SEM simulations to fit observations recorded in a three-dimensional environment. The simulations represent wave propagation in the vertical plane containing the source and receiver and therefore neglect out-of-plane propagation and three-dimensional heterogeneity. These effects may introduce differences between observed and synthetic waveforms, potentially affecting waveform correlations and inferred thicknesses. Because the misfit is based on normalized cross-correlation, it is less sensitive to overall amplitude differences but remains sensitive to discrepancies in waveform phase and shape. Future work using three-dimensional simulations could help quantify the impact of these effects on the thickness estimates.

For larger-scale tomographies, using a regular mesh could become prohibitive in terms of computational cost. To address this issue, we suggest that an adaptive mesh, in which the cell size is driven by ray density, could substantially reduce the computational load by using larger cells where ray density is low and smaller cells where it is high. For this approach to be successful, further developments are required to compute the spatial gradient in the operator $\mathbf{L}$ of Eq. (**??**).

Finally, extending the database with a varying Young's modulus could open the possibility to produce simultaneous maps of the ice thickness and rigidity, by exploiting waveforms from the non-dispersive modes on the horizontal channels. Young's modulus is a key parameter for sea

ice models, which currently assume a constant value, while field observations show important variations between 1 and 9 GPa (Moreau et al., 2020a; Kuchly et al., 2025), with significant impact on the response of sea ice to geophysical forcing in the marginal ice zone.

Beyond these inversion and parameterization improvements, a natural extension of this work is to include more realistic ice physics in the forward-modeling and inversion framework. The current study assumes a homogeneous, isotropic, and snow-free ice layer, which is a useful first approximation but does not fully represent natural sea ice. In reality, ice can be heterogeneous and anisotropic because of temperature variations, brine, air bubbles, cracks, and crystal orientation. Snow cover can also affect the dispersion of flexural waves.

Extending the spectral-element database and inversion framework to account for these effects, for example through depth-dependent or anisotropic elastic properties and the inclusion of a snow layer, would improve thickness estimates in more complex environments and represents a promising direction for future research.

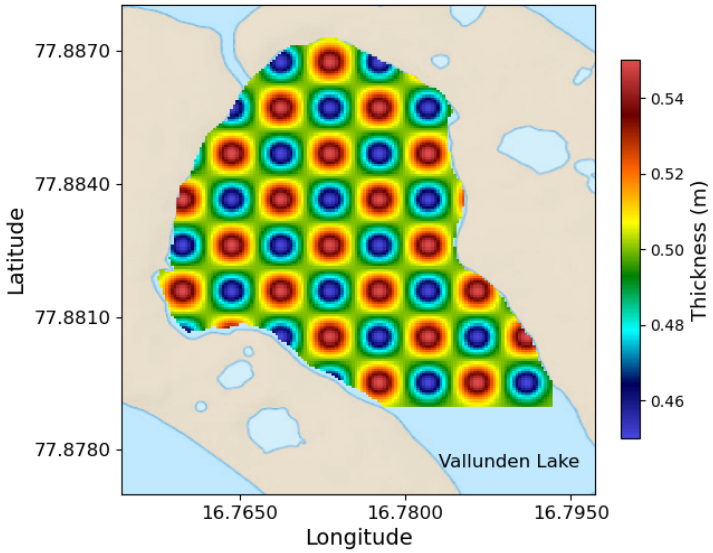


Figure 1: Synthetic checkerboard model used to generate the ice thickness observations for evaluating the influence of station geometry on the tomographic inversion. (Figure 15 of the revised manuscript.)

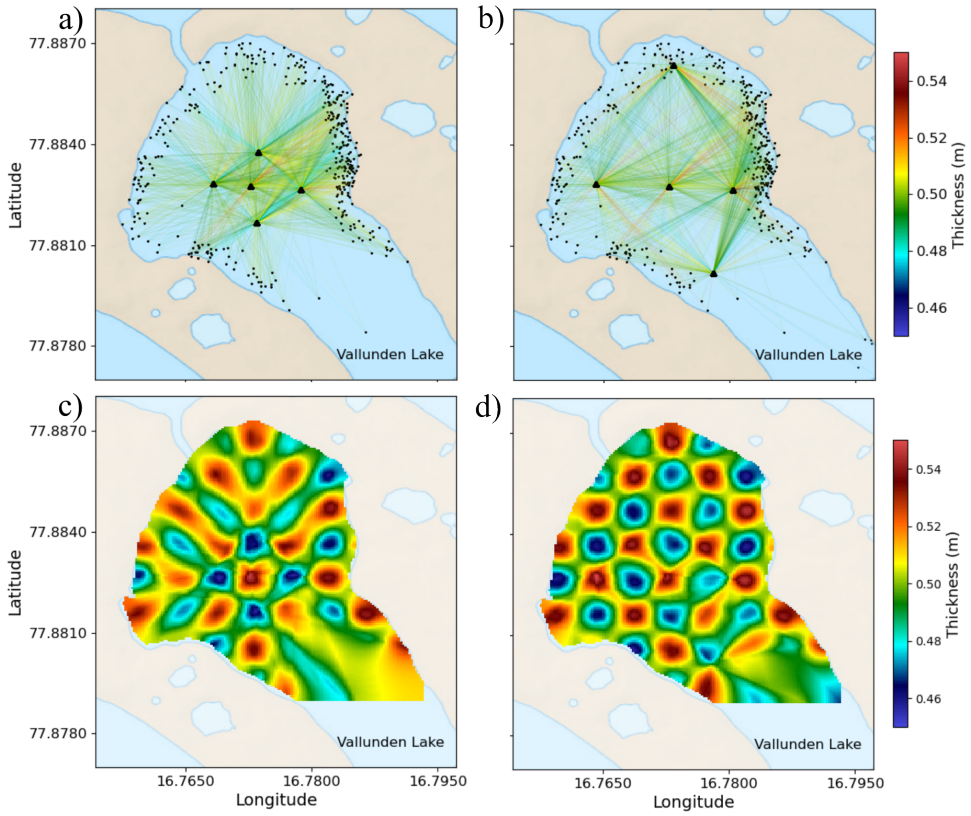


Figure 2: Influence of station geometry on tomographic reconstruction. (a) Original station configuration used for the first-step inversion to estimate ice thickness and source locations. (b) Optimized station configuration providing broader and more uniform ray coverage. (c) Tomographic inversion result using the original configuration. (d) Tomographic inversion result using the optimized configuration, showing an improved recovery of the checkerboard anomalies. (Figure 16 of the revised manuscript.)

Regrettably, the authors mention checkerboard tests (lines 294 – 300), but I can not see checkerboard test results in the submitted material, and this is also not referred to in the assessment of the inversion results shown in Figure A-1.

We thank the reviewer for pointing this out. We have added the checkerboard test figures to Appendix C (Figs. C-1–C-6) and revised Sections 3.1 and 3.2 to include a quantitative discussion of the results for the controlled-source and natural-source datasets, respectively. For the natural-source data, we assess resolution on two days with good and poor ray coverage and discuss the implications for the recovered spatial variations.

Appendix C: Checkerboard Resolution Tests

This appendix presents synthetic checkerboard tests designed to evaluate the spatial resolution of the ice-thickness tomographic inversion for the controlled-source and natural-source datasets. Synthetic thickness models with alternating anomalies of different spatial scales are used to assess how well the inversion recovers the imposed patterns under the ray coverage of each experiment.

C1: Controlled-source data

Three checkerboard tests with different anomaly sizes are performed to assess the spatial resolution of the controlled-source tomographic inversion. The tests use the 296 selected paths out

of 528 possible paths.

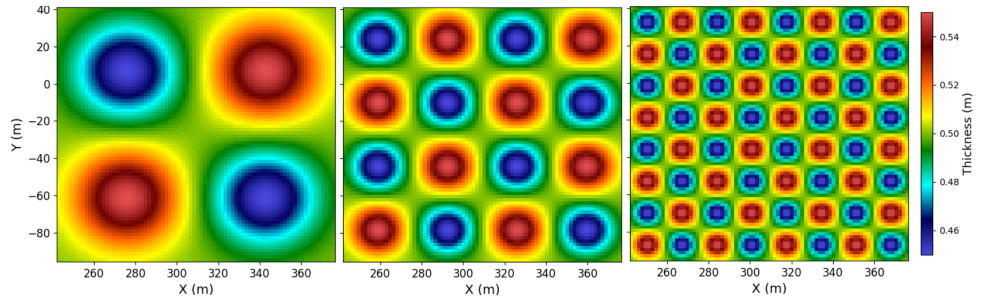


Figure C-1: Synthetic checkerboard models used for the controlled-source resolution tests.

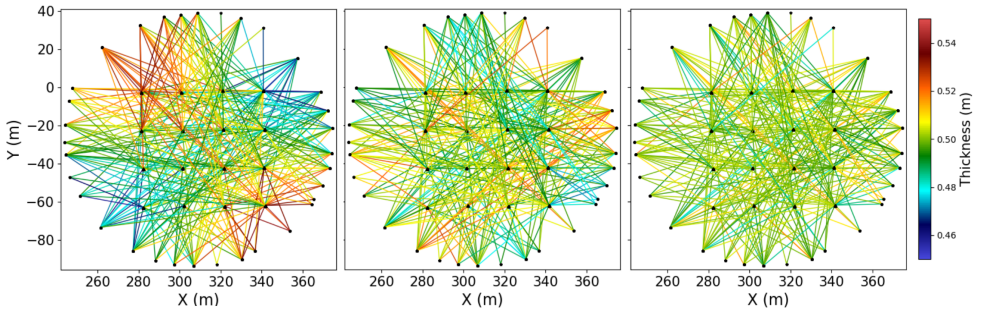


Figure C-2: Source–receiver paths for the 296 selected paths, colored according to their thickness values, as indicated by the colorbar.

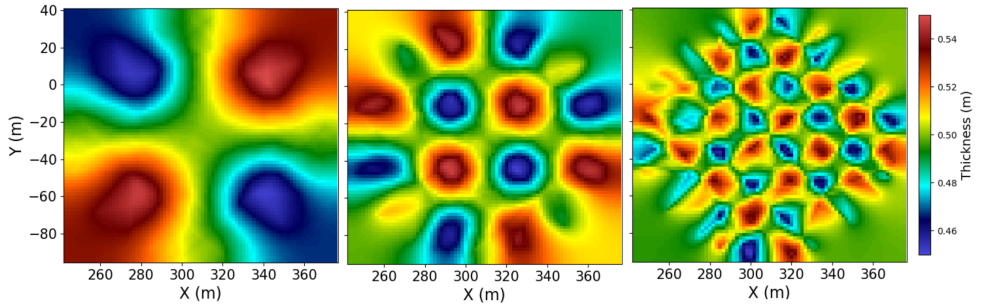


Figure C-3: Tomographic reconstructions of the three synthetic checkerboard models for the controlled-source dataset.

C2: Natural-source data

Checkerboard tests are performed using paths retained after quality control. Results are presented for 7 March 2019, one of the days with the best path coverage, and 18 March 2019, which has the fewest paths and the poorest coverage.

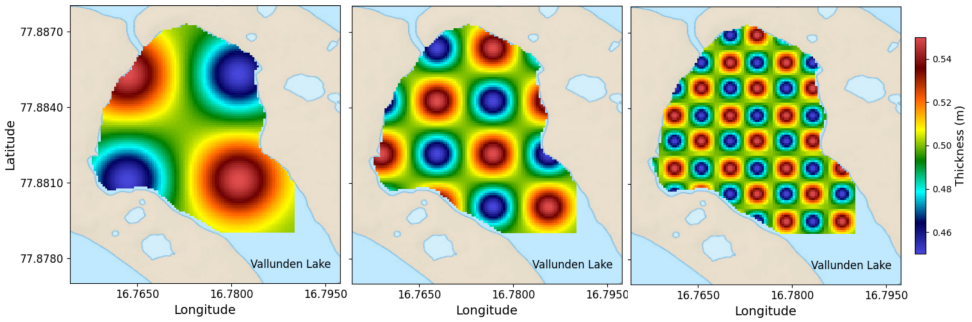


Figure C-4: Synthetic checkerboard models used for the natural-source resolution tests.

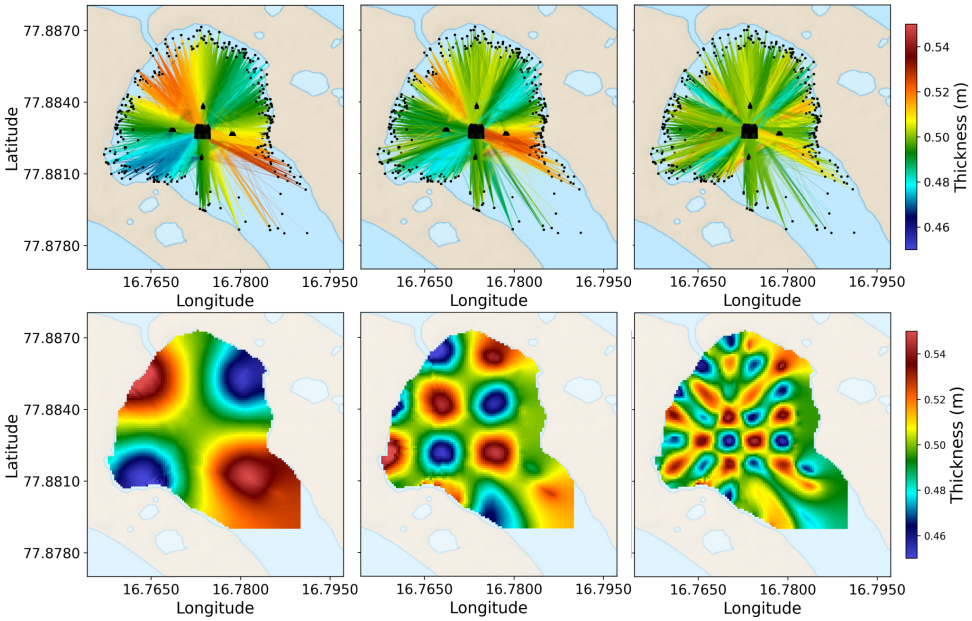


Figure C-5: Source–receiver paths for the selected paths, colored according to their thickness values (top), and the corresponding tomographic reconstructions for 7 March 2019 (bottom).

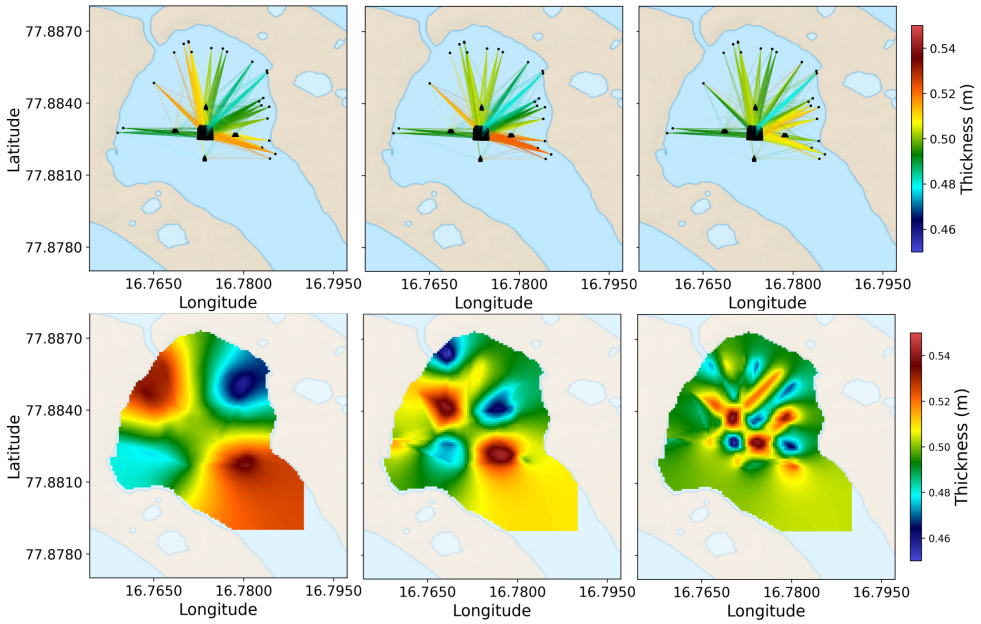


Figure C-6: Source–receiver paths for the selected paths, colored according to their thickness values (bottom), and the corresponding tomographic reconstructions for 18 March 2019 (bottom).

I caution that a checkerboard test can be useful to analyze data coverage and resolution in the straight ray framework, but it does not allow an assessment of the fundamental resolution issue reported above, i.e., if the physics of wave propagation of flexural modes guided in an ice layer with constant and with variable thickness is compatible with the spatial discretization of the inverse problem.

We agree with the reviewer that a checkerboard test assesses the ray coverage and the effect of the regularization, but not the validity of the physics. The principle of the method is the one of classical tomography. As shown in our answer on linearity, the flexural wavenumber is, up to a factor that does not depend on the thickness, a function of $h^{-3/5}$, which plays the role of the slowness. Retrieving a thickness map from path-averaged values is therefore the same problem as retrieving a velocity map from path-averaged travel times or phase velocities, as in surface-wave tomography (Woodhouse and Dziewonski, 1984; Kennett and Yoshizawa, 2002): the quantity inverted is a thickness-related slowness instead of a velocity. The fundamentals of the method (path integral, discretization, regularization, resolution analysis) are those of this well-established framework, and we do not think that they need to be re-established here for guided waves. The validity of the specific approximations that come with guided waves in ice (path average of the wavenumber, dispersion) is what we address in Appendix B. Full-wave simulations in a laterally heterogeneous plate, which would be needed for a complete sensitivity analysis, are beyond the scope of this paper because of their computational cost. This limit is now clearly stated in the revised manuscript.

In addition, the complete chain (waveform inversion on each path, path-average assumption and tomography) is tested against independent observations, in conditions where the thickness is not uniform. In the Canada dataset, the tomography reproduces the sharp thickness discontinuity across the ridge and agrees closely with the thicknesses obtained independently from frequency–wavenumber inversions of DAS data along the fibre line (Kuchly et al., 2026), and with the drillings, up to a systematic offset of about 5 cm that we attribute to the consolidated slush layer at the base of the ice. In the Svalbard dataset, the thicknesses of 75–80 cm measured in the northern channel in mid-March are reproduced, the average increase of 10–15 cm in the dense-array area follows the independent estimates and the drillings, and the thicker ice

along the shoreline is consistent with the physical expectation of Marchenko et al. (2021). A fundamental inconsistency between the physics of wave propagation in ice of variable thickness and the tomographic model would appear as discrepancies larger than the few centimetres observed, especially across sharp contrasts such as the ridge. These comparisons constrain the mean thickness and the large contrasts, and not the small structures discussed in the other comments.

We have modified the "Limitations and perspectives" to cover those points (together with those raised by the other reviewer):

Limitations and perspectives

Potential improvements to our approach could involve exploring other types of cost functions. The use of cross-correlations of signals is known to be non-convex, owing to the oscillatory behavior of the waveforms. Alternative misfit functions, for example those based on spectrogram correlations or on optimal transport distances (Métivier et al., 2022), could improve the convergence of the inversion in steps 2–4 of the procedure.

An important limitation of the present approach is that the tomographic inversion assumes straight-ray propagation and represents each source–receiver path by a single equivalent thickness. Although the checkerboard tests assess the effects of ray coverage and regularization within this framework, they do not independently validate the approximation for flexural waves propagating through ice of spatially variable thickness. As discussed in Appendix B, the equivalent thickness differs from the path-averaged thickness by a second-order correction, which is less than 1.2 cm for the thickness contrasts considered here. However, this analysis assumes slowly varying thickness and does not account for possible wave scattering, reflection, or mode conversion caused by lateral heterogeneity. A complete assessment of these effects would require full-wave simulations in a laterally heterogeneous ice layer, which are beyond the scope of this study. The recovered thickness variations should therefore be interpreted with these limitations in mind.

A further limitation is that the waveform inversion uses 2D SEM simulations to fit observations recorded in a three-dimensional environment. The simulations represent wave propagation in the vertical plane containing the source and receiver and therefore neglect out-of-plane propagation and three-dimensional heterogeneity. These effects may introduce differences between observed and synthetic waveforms, potentially affecting waveform correlations and inferred thicknesses. Because the misfit is based on normalized cross-correlation, it is less sensitive to overall amplitude differences but remains sensitive to discrepancies in waveform phase and shape. Future work using three-dimensional simulations could help quantify the impact of these effects on the thickness estimates.

For larger-scale tomographies, using a regular mesh could become prohibitive in terms of computational cost. To address this issue, we suggest that an adaptive mesh, in which the cell size is driven by ray density, could substantially reduce the computational load by using larger cells where ray density is low and smaller cells where it is high. For this approach to be successful, further developments are required to compute the spatial gradient in the operator $\mathbf{L}$ of Eq. (17).

Finally, extending the database with a varying Young's modulus could open the possibility to produce simultaneous maps of the ice thickness and rigidity, by exploiting waveforms from the non-dispersive modes on the horizontal channels. Young's modulus is a key parameter for sea ice models, which currently assume a constant value, while field observations show important variations between 1 and 9 GPa (Moreau et al., 2020a; Kuchly et al., 2025), with significant impact on the response of sea ice to geophysical forcing in the marginal ice zone.

Beyond these inversion and parameterization improvements, a natural extension of this work is to include more realistic ice physics in the forward-modeling and inversion framework. The current study assumes a homogeneous, isotropic, and snow-free ice layer, which is a useful first

approximation but does not fully represent natural sea ice. In reality, ice can be heterogeneous and anisotropic because of temperature variations, brine, air bubbles, cracks, and crystal orientation. Snow cover can also affect the dispersion of flexural waves.

Extending the spectral-element database and inversion framework to account for these effects, for example through depth-dependent or anisotropic elastic properties and the inclusion of a snow layer, would improve thickness estimates in more complex environments and represents a promising direction for future research.

Considering the phenomenon of cycle-skipping, the reader can be interested to learn about preprocessing steps or choices such as windowing and tapering before evaluating the correlation misfit between observations and synthetics (Eq 5).

We thank the reviewer for this suggestion. We have added the following text to Section 2.3, immediately after the definition of the correlation misfit (Eq. (5)).

For natural-source data, most of the observed signal energy lies below 30 Hz, although some waveforms contain frequencies reaching approximately 40 Hz. We therefore use a 15 Hz Ricker source wavelet to generate the synthetic waveforms and apply a low-pass filter with a cutoff frequency of approximately 2.5 times the central frequency (about 38 Hz) to the observed data. For controlled-source data, the observed signal energy extends to approximately 100 Hz; we consequently use a 40 Hz Ricker wavelet and a low-pass cutoff of approximately 100 Hz. In both cases, a Hann window is applied to the waveform segment containing the flexural wave, tapering its beginning and end, and the misfit is calculated over the entire windowed segment.

The correlation-based misfit function may contain multiple local minima associated with cycle-skipping. We mitigate this risk through extensive exploration of the parameter space using simulated annealing, followed by a grid search to refine individual path-thickness estimates. Although this strategy reduces the risk of becoming trapped in local minima, it cannot guarantee the complete avoidance of cycle-skipping.

Related to the quality of the waveform inversion, the manuscript can perhaps benefit from additional comments about the suitability of 2D synthetics applied to waveforms observed in a 3D medium.

To cover this point, in the modified *Discussion* section (see above), we now acknowledge the limitations of using 2D synthetics to fit waveforms recorded in a 3D environment. Our 2D SEM simulations capture wave propagation in the vertical plane containing the source and receiver, but do not account for out-of-plane propagation or three-dimensional heterogeneity. These effects may contribute to waveform differences between the synthetics and observations and may affect the inferred thickness. We also clarify that the normalized correlation misfit reduces sensitivity to overall amplitude differences, but does not eliminate discrepancies in waveform shape or phase caused by three-dimensional effects. Full 3D simulations are beyond the scope of this study and remain a direction for future work.

I would also welcome a better discussion of the obtained tomography results shown in Figures 12, 13, and A-1. The character of the images is understandably streaky. What should a reader consider to avoid this artefact in future deployments? Were the stations deployed with doing a tomography in mind?

We thank the reviewer for this comment. We have added a short paragraph to the "Inversion of icequakes in Svalbard" subsection to clarify the origin of the streak-like features, their relationship with uneven ray coverage and limited azimuthal diversity, and the implications for future deployments:

The streak-like features in the tomographic images are consistent with the uneven spatial distribution and limited azimuthal diversity of the source–receiver paths, particularly near the edges of the study area. The checkerboard tests confirm that spatial resolution varies with ray

coverage, with poorer recovery in less well-sampled areas. As discussed in Section 4.2, a broader and more spatially distributed station geometry could improve tomographic reconstruction.

We also clarify that the array was originally designed around a large central 2D grid to capture guided waves propagating in different directions, supplemented by four smaller, distant linear arrays serving as virtual sources for noise-correlation analysis (see our response above: The dense Svalbard array was designed as a proof of concept...). The existing discussion in Section 4.2 further explains the limitations of this geometry for tomography and illustrates how broader station coverage can improve reconstruction.

To assess the influence of spatial discretization, we repeated the inversion for 10 March using cell sizes of 6, 12 and 18 m, using the same smoothing parameter. The streaks remain visible at all three resolutions, although their appearance and the level of spatial detail change with cell size. The 6 m discretization provides greater flexibility to represent small-scale variations in the densely sampled central area, whereas increasing the cell size reduces the number of unknowns and limits the spatial detail that can be represented. However, the streak-like artefacts persist even at coarser resolutions, indicating that they cannot be attributed to fine discretization alone. Increasing the regularization can suppress some of these structures, but at the cost of poorer agreement between the predicted and observed path-averaged thicknesses. No single choice of cell size or smoothing parameter fully compensates for the uneven path coverage. The results of this comparison are shown in Fig. C-7.

For future deployments intended for tomography, we recommend optimizing station locations to obtain more spatially uniform path coverage and, in particular, a wider range of path azimuths across the entire study area, including its margins. Distributing stations more broadly, rather than concentrating most of them in the centre, and combining this geometry with sources distributed around the area would improve the constraints on lateral variations:

To illustrate the influence of station geometry on the recovery of spatial variations, we performed synthetic checkerboard tests (Figures 15 and 16). Figure 15 shows the checkerboard model used to generate the synthetic thickness observations. In Figure 16a, the stations are those of the original configuration used in the Svalbard experiment, whereas in Figure 16b they are relocated to provide broader and more uniform coverage of the study area. The corresponding tomographic results are shown in Figures 16c and d. The broader geometry leads to a substantially better recovery of the checkerboard pattern, demonstrating the strong influence of station placement on ray coverage and inversion quality. These tests provide an empirical assessment of the spatial scales recoverable with the tested geometries and inversion settings.

The optimal design depends on the size of the monitored area, from a few hundred metres to several kilometres, on the propagation distance and source distribution, and on the desired spatial resolution, but also on the ability of an icequake to propagate over long distances. Finally, long paths cross more thickness variations and cracks, making the assumption of slowly varying thickness (Section 2.3.3) more difficult to satisfy. Shorter paths are therefore more likely to meet this assumption, but require more stations or clusters to cover a given area. Determining the minimum number of stations and their optimal distribution for a given area and target resolution requires a dedicated study.

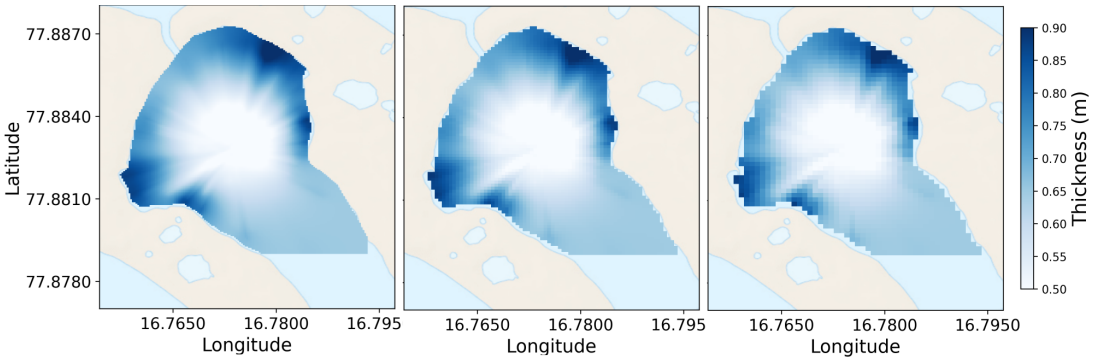


Figure C-7: Comparison of tomographic reconstructions of sea-ice thickness on 10 March 2019 using cell sizes of (a) 6 m, (b) 12 m, and (c) 18 m. The same smoothing parameter was used for all three inversions. The streak-like features persist across the different discretizations, indicating that they cannot be eliminated simply by increasing the cell size.

Is the icequake distribution exclusively along the shore a surprise, or was this expected? It appears that the blue-ish thick ice zone correlates very well with the epicenter distribution; can this be an artifact? Is this a limit of the resolution?

The distribution of the icequakes along the shore was expected: these are tidal icequakes (the correlation with the tides is demonstrated in Moreau et al. (2023)), which occur where the ice is under maximum stress, that is, along the shore. The thicker ice along the shore was not a surprise either, since it is also reported in Marchenko et al. (2021) and is consistent with stronger thermal exchanges with the solid ground. Both observations have the shoreline as a common cause, which can explain their spatial correlation without invoking an artifact.

We cannot, however, totally rule out a bias associated with the distribution of the sources, because the whole thickness map of the fjord is not known independently. Two arguments make us believe that the thick ice along the shore is not an artifact. First, the known biases of the method act in the opposite direction: since the accumulated phase is fixed by the data, a source position error δL on the path length changes the retrieved thickness by $\delta h/h \simeq (5/3)\,\delta L/L$, and the constraint that the sources lie on the ice forbids the errors that would lengthen the paths beyond the shoreline. The remaining errors shorten the paths and therefore lead to a thinner, and not a thicker, apparent ice. These errors are in addition limited to a few centimetres, whereas the contrasts are of 20 cm or more. Second, the thickness measured in the field in the northern channel (75–80 cm in mid-March) is reproduced by the inversion, and the tendency of thicker ice near the sources is not observed in the Canada dataset, where the sources surround a denser array and where the tomography reproduces the sharp thickness change between the two types of ice. We therefore do not think that this correlation reveals a resolution or methodological limit, although the resolution is lower close to the shoreline, where the rays are less numerous and less diverse in azimuth, as discussed in our answers on resolution and on the robustness of the features.

This is now explained in the revised manuscript:

The interpretation of small-scale thickness variations near the shoreline requires particular caution. Although many icequake sources are located along the shoreline, the resulting ray paths do not necessarily provide uniform spatial coverage or a broad range of azimuths near the sources. Moreover, rays originating from nearby sources may share common path segments, so that errors associated with these sources may be mapped into the same region rather than averaging out. Consequently, thickness estimates near the shoreline are less well constrained than in the central part of the array, where rays cross at different azimuths. Source-location errors may contribute to small-scale artefacts, although their estimated magnitude is insufficient to explain the largest observed thickness contrasts.

Related, for example, we see a small area of intermittently increased ice thickness at about 16.78 lat / 77.8840 lon on 11 and 12 March (Figs 13, A-1). Figure 12 shows that in this time period icequakes occur at that location, but not before and not after. The authors should better qualify which of the features shown in the distributions are robust and which can be likely or possibly artifacts.

We thank the reviewer for this comment, which points to a real limitation of passive tomography, and we have revised the text accordingly.

The small area of thicker ice near 77.885°N, 16.78°E is crossed by rays on all days, but it is sampled in two different ways. On most days the rays only transit through it: the area is a small fraction of each path, which is several hundred metres long, so each ray is weakly sensitive to it, and the rays come from different sources, with different azimuths and independent errors. On 11 and 12 March, the events located in the area are not natural icequakes: they were produced by collisions of very large (5m x 10m) ice blocks carried out by another research group during our deployment, whose waveforms are very similar to those of icequakes and which were therefore included in the catalog by the automatic detection (Moreau et al., 2023). This explains why events occur there on these two days and not before or after. All the rays emitted by these events start in the same place, so that the area is a segment common to all of these paths, and any error common to the rays of these events, for instance on the source position, is attributed to this area and does not average out. In addition, the ice there was being deliberately disturbed (blocks, broken ice), so that the local thickness and the waves generated may differ from the surrounding ice.

The daily prior does not compensate for this. Cells whose ray coverage decreases from one day to the next receive a weight of 0.01 (the weight is 1 where the coverage is equal or higher), so that the structure obtained on days with local sources is not carried over to the following days, and the area is then determined by the few transiting rays and by the smoothing. We therefore cannot say, from the maps alone, whether the feature on 11 and 12 March reflects thicker or disturbed ice that is visible only when sources are present, or an artefact of the errors common to these events. We consider it a possible artefact. In the revised manuscript we now flag them in Figure 12, and discuss their impact on the resulting tomography in Figure 13. We agree that the tomography is more reliable where rays cross with different azimuths (the centre of the array) than near the sources, which lie mostly on the shoreline. This is a classical limitation of straight-ray tomography with sources confined to the boundary of the domain. Errors on the source position are one contribution: for a flexural wave the accumulated phase $k(h)\,L$ is fixed by the data and $k \propto h^{-3/5}$, so that $\delta h/h \simeq (5/3)\,\delta L/L$. A location error of 10 m on a 300 m path changes the thickness by a few centimetres. This cannot explain thickness contrasts of 20 cm, which would require location errors of about 50–70 m, larger than the ±25 m search window. It can, however, blur or create small-scale features of a few centimetres.

What we consider robust.

- The large-scale pattern (thinner ice in the centre, thicker ice along the shore) is present on all days, with rays coming from many azimuths. Thicker ice along the shoreline is expected physically, since shallower water inhibits convection and reduces the heat flux to the ice base (Marchenko et al., 2021), and field measurements of 75–80 cm in the northern channel in mid-March are reproduced by the inversion.

- The same method applied to the controlled-source experiment in Canada, where the sources are distributed all around and close to the array, does not show such features: the tomography reproduces the sharp thickness discontinuity across the ridge, agrees closely with the thicknesses obtained from frequency–wavenumber inversions of DAS data (of the type used by Serripierri et al. (2022)) along the fibre line, and with the drillings, up to a

systematic offset of about 5 cm that we attribute to the consolidated slush layer at the base of the ice.

- The average thickness in the dense-array area follows the same increase (10–15 cm) as independent estimates and as the drillings.

What we consider fragile. Isolated patches that appear only on days when icequakes occur at the same location, and streaks aligned with the ray directions, particularly on days with few rays, should not be interpreted individually.

In the revised manuscript, we have clarified these points in Sections 3.2, 4.1, and 4.2. In Section 3.2, we identify the main thickness patterns that we consider robust and distinguish them from small-scale features that may be artefacts. In Section 4.1, we quantify the potential bias introduced by constraining the source locations to the ice. Finally, in Section 4.2, we discuss how limited ray coverage and azimuthal diversity near the shoreline affect the interpretation of small-scale thickness variations.

We also specifically discuss the anomaly observed on 11 and 12 March and why it may be related to the artificial ice-block collisions and the associated ray geometry:

The main thickness patterns observed in the daily reconstructions are considered robust: in particular, the thicker ice near the shoreline, the thinner ice in the central region, and the overall increase in thickness during the measurement period. These features are supported by their persistence across multiple days and, where available, by independent thickness measurements. In contrast, isolated small-scale anomalies and streak-like features, particularly in areas with sparse ray coverage or limited azimuthal diversity, may reflect inversion artefacts rather than genuine variations in ice thickness. Their interpretation therefore requires caution. One example is the small patch of increased thickness observed on 11 and 12 March near 77.885°N, 16.78°E. The seismic events recorded at this location on these two days were generated by another research group through deliberate collisions of large ice blocks, rather than by natural icequakes. The corresponding thickness anomaly should therefore be treated as a possible artefact. Because these events originated from approximately the same location, their ray paths share a common segment, meaning that errors common to these rays may be mapped into this area rather than averaging out. In addition, the deliberate disturbance of the ice may have affected its local thickness and the generated waveforms.

I understand the motivation to provide time-lapse images, and the associated average temporal evolution shown in Figure 14 is valuable and interesting. However, if the quality of the spatial variability is too difficult to constrain (which again relates to the resolution), it may be considered to leave these results out of the manuscript and provide instead an average tomography result or 2D change images associated with longer time intervals and hence better data coverage.

We thank the reviewer for this helpful suggestion. We agree that the spatial variability in the time-lapse reconstructions must be interpreted in light of the limitations in spatial resolution. The reliability of the reconstructed patterns depends on ray coverage and azimuthal diversity, as well as on regularization. In particular, the prior-data term constrains the solution towards the prescribed prior thickness model, meaning that some local features may be insufficiently constrained by the observations and partly influenced by the prior.

We have therefore clarified the limitations of the spatial interpretation, distinguishing large-scale patterns supported by independent observations from isolated small-scale features that may be affected by limited ray coverage or regularization. We have retained the daily reconstructions in the Appendix to document the temporal variability, while Figure 14 presents the evolution of the average thickness over the dense-array area. This average shows an increase of approximately 10–15 cm during the deployment, consistent with independent estimates and drilling measurements. We consider this temporal evolution valuable, while acknowledging that

the interpretation of small-scale spatial features remains limited by data coverage and regularization.

List of minor points

Lines 220 – 224 still contain place holders for values (mentioned in the introduction of the review).

We thank the reviewer for pointing this out. This issue has now been addressed in the revised manuscript, where the placeholders have been replaced with the corresponding numerical values.

References

Boiero, D. and Socco, L. V.: The meaning of surface wave dispersion curves in weakly laterally varying structures, Near Surface Geophysics, 9, 561–570, 2011.

Kennett, B. and Yoshizawa, K.: A reappraisal of regional surface wave tomography, Geophysical Journal International, 150, 37–44, 2002.

Kuchly, S., Auvity, B., Mokus, N., Bureau, M., Nicot, P., Fourgeaud, A., Dansereau, V., Eddi, A., Perrard, S., Dumont, D., and Moreau, L.: An integrated multi-instrument methodology for studying marginal ice zone dynamics and wave-ice interactions, EGUsphere, 2025, 1–23, https://doi.org/10.5194/egusphere-2025-3304, 2025.

Kuchly, S., Moreau, L., Zanchi, V., Mokus, N., Dansereau, V., Smith, M. M., Dumont, D., Perrard, S., and Eddi, A.: High-Resolution Measurement of Sea Ice Mechanical Characteristics Using Distributed Acoustic Sensing, Seismological Research Letters, (under review), 2026.

Marchenko, A., Morozov, E., Ivanov, A., Elizarova, T., and Frey, D.: Ice thickening caused by freezing of tidal jet, Russian Journal of Earth Sciences, 21, 1–8, https://doi.org/10.2205/2021ES000761, 2021.

Métivier, L., Brossier, R., Kpadonou, F., Messud, J., and Pladys, A.: A review of the use of optimal transport distances for high resolution seismic imaging based on the full waveform, MathematicS In Action, 11, 3–42, https://doi.org/10.5802/msia.15, 2022.

Moreau, L., Minonzio, J.-G., Talmant, M., and Laugier, P.: Measuring the wavenumber of guided modes in waveguides with linearly varying thickness, The Journal of the Acoustical Society of America, 135, 2614–2624, https://doi.org/10.1121/1.4869691, 2014.

Moreau, L., Boué, P., Serripierri, A., Weiss, J., Hollis, D., Pondaven, I., Vial, B., Garambois, S., Larose, É., Helmstetter, A., Stehly, L., Hillers, G., and Gilbert, O.: Sea Ice Thickness and Elastic Properties From the Analysis of Multimodal Guided Wave Propagation Measured With a Passive Seismic Array, Journal of Geophysical Research: Oceans, 125, e2019JC015 709, https://doi.org/10.1029/2019JC015709, 2020a.

Moreau, L., Weiss, J., and Marsan, D.: Accurate Estimations of Sea-Ice Thickness and Elastic Properties From Seismic Noise Recorded With a Minimal Number of Geophones: From Thin Landfast Ice to Thick Pack Ice, Journal of Geophysical Research: Oceans, 125, e2020JC016 492, https://doi.org/10.1029/2020JC016492, 2020b.

Moreau, L., Seydoux, L., Weiss, J., and Campillo, M.: Analysis of microseismicity in sea ice with deep learning and Bayesian inference: application to high-resolution thickness monitoring, The Cryosphere, 17, 1327–1341, https://doi.org/10.5194/tc-17-1327-2023, 2023.

Serripierri, A., Moreau, L., Boue, P., Weiss, J., and Roux, P.: Recovering and monitoring the thickness, density, and elastic properties of sea ice from seismic noise recorded in Svalbard, The Cryosphere, 16, 2527–2543, https://doi.org/10.5194/tc-16-2527-2022, 2022.

Woodhouse, J. and Dziewonski, A.: Mapping the upper mantle: Three dimensional modelling of earth structure by inversion of seismic waveforms, Journal of Geophysical Research, 89, 5953–5986, 1984.