Supporting information for “Vertical slice ocean tomography with seismic waves”

Jörn Callies¹, Wenbo Wu¹,², Shirui Peng¹, Zhongwen Zhan¹

¹California Institute of Technology, Pasadena, CA, USA
²Wood Hole Oceanographic Institution, Woods Hole, MA, USA

Contents of this file

1. Texts S1 to S2
2. Figures S1 to S9
3. Table S1

Corresponding author: Jörn Callies, jcallies@caltech.edu
Text S1  Inversion for travel time anomalies

The procedure to infer travel time anomalies relative to an unknown but common reference from measurements of T-wave arrival time changes and origin time corrections (relative to a cataloged event time) is similar to that described in Wu et al. (2020), but here we improve on that inversion in a number of ways. In particular, we invert for the T-wave delays and origin-time corrections separately, which allows us to more naturally use multiple T-wave and land stations as well as a set of T-wave frequencies. Instead of the somewhat ad hoc regularization employed in Wu et al. (2020), we here impose a full set of prior and noise statistics and use a Gauss–Markov estimator. See, for example, Wunsch (2006) for an introduction to the methods employed here.

Our primary goal is to infer time series of the travel time anomalies at the observed frequencies \( \omega_1, \ldots, \omega_l \), stacked into the vector

\[
\tau = \begin{pmatrix} \tau_1 \\ \vdots \\ \tau_l \end{pmatrix}.
\] (1)

These travel time anomalies, ascribed to changes in the ocean's temperature field, are differences between T-wave arrival anomalies and origin time corrections:

\[
\tau = Da \quad \text{with} \quad D = \begin{pmatrix} I_m & -I_m \\ \vdots & \vdots \\ I_m & -I_m \end{pmatrix} \quad \text{and} \quad a = \begin{pmatrix} a_1^{(T)} \\ \vdots \\ a_l^{(T)} \\ a_1^{(L)} \end{pmatrix},
\] (2)

where \( a_i^{(T)} \) contains the T-wave travel time anomalies for each event time at the different frequencies, and \( a_i^{(L)} \) contains the origin time corrections inferred from the land stations. Here, \( I_m \) denotes the identity matrix of size \( m \), the number of unique events. All delays are referenced to predicted arrival times based on the cataloged event times. The predictions are made using the PREM solid earth model for the land stations and a constant T-wave reference velocity of 1.5 km s\(^{-1}\).

The travel time anomalies \( \tau \) are related to the range-averaged temperature anomalies through the sensitivity kernel:

\[
\tau = (K \otimes I_m)T \Delta z,
\]

where \( K \) is the sensitivity matrix.

\[
K = \begin{pmatrix} K_1^T \\ \vdots \\ K_l^T \end{pmatrix} \quad \text{and} \quad T = \begin{pmatrix} T_1 \\ \vdots \\ T_n \end{pmatrix}
\] (3)

are the matrices containing the range-integrated kernels as rows and a stack of the time series of range-averaged temperature anomalies at the \( n \) vertical levels, respectively, and \( \otimes \) denotes a Kronecker product. As described in the main text, we perform an SVD of the sensitivity matrix \( K = U \Lambda V^T \). We normalize such that \( U^T U = I_3 \) and \( h^{-1} V^T V \Delta z = I_l \), where
\( \Delta z = 50 \text{ m} \) is the vertical grid spacing and \( h = 5 \text{ km} \) is a reference depth. With this normalization, the singular vectors are unitless and do not depend on the discretization (as long as it is fine enough). We denote the projections of the temperature anomalies onto the singular vectors by \( \mathbf{c} = h^{-1}(\mathbf{V}^T \otimes \mathbf{I}_m)\Delta z \), so that \( \mathbf{r} = h(\mathbf{U} \Lambda \otimes \mathbf{I}_m)\mathbf{c} \), which can be inverted for the time series \( \mathbf{e} = h^{-1}(\Lambda^{-1}\mathbf{U}^T \otimes \mathbf{I}_m)\mathbf{r} \).

We cannot, however, observe the travel time anomalies directly. We only observe changes in the travel time between repeating earthquakes. The design matrix thus takes differences of travel time anomalies between event pairs:

\[
\mathbf{E} = \begin{pmatrix}
\mathbf{I} \otimes \mathbf{X}^{(T)} & 0 \\
0 & \mathbf{X}^{(L)}
\end{pmatrix}.
\]

The pair matrices \( \mathbf{X}^{(T)} \) and \( \mathbf{X}^{(L)} \) take these differences between events at all T-wave and land stations involved. For example, if there were four pairs connecting five events observed at one T-wave section, we might have

\[
\mathbf{X}^{(T)} = \begin{pmatrix}
-1 & 1 \\
-1 & 1 \\
-1 & 1 \\
-1 & 1
\end{pmatrix}.
\]

If pairs 1, 2, and 4 are observed at land station 1, and pairs 1 and 3 are observed on land station 2, the land station pair matrix would be

\[
\mathbf{X}^{(L)} = \begin{pmatrix}
-1 & 1 \\
-1 & 1 \\
-1 & 1 \\
-1 & 1
\end{pmatrix}.
\]

The observed arrival time offsets, again measured relative to the cataloged event times, are then

\[
\delta = \mathbf{Ea} + \mathbf{n} \quad \text{with} \quad \delta = \begin{pmatrix}
\delta_{1}^{(T)} \\
\vdots \\
\delta_{l}^{(T)} \\
\delta_{l}^{(L)}
\end{pmatrix}
\]

containing a stack of the T-wave offsets measured at different frequencies and the land offsets. The observational noise is denoted by \( \mathbf{n} \).

We specify prior statistics for the projections onto the singular vectors (i.e., for \( \mathbf{c} \)), and we assign these covariances to the T-wave anomalies \( \mathbf{a}_{1}^{(T)}, \ldots, \mathbf{a}_{l}^{(T)} \):

\[
\mathbf{R} = \begin{pmatrix}
\mathbf{h}^2(\mathbf{U} \Lambda \otimes \mathbf{I}_m)(\mathbf{\Sigma} \otimes \mathbf{C})(\mathbf{U} \Lambda \otimes \mathbf{I}_m)^T & 0 \\
0 & 0
\end{pmatrix} + \mathbf{\sigma}^2 \mathbf{I}_{l+1} \otimes \mathbf{I}_m,
\]

\( 3 \)
where $\Sigma$ is a diagonal matrix that specified the prior variances in the time series of the projections onto the singular vectors (Table S1). The diagonal nature of $\Sigma$ means that we assign no prior correlations between the projections onto the singular vectors. The matrix $C$ specifies the time correlation $C_{ij} = e^{-|t_i - t_j|/\lambda}$, which we assume to be identical for all singular vectors. The expected standard deviation of origin time corrections is specified as $\sigma = 5$ s, and $I_{l+1}$ denotes a square matrix of size $l + 1$ filled with ones. This prescribes perfect correlation between the origin time correction and the part of the $T$-wave anomalies that is due to origin time errors.

In addition to the random process with exponential correlation function, we also include deterministic annual and semi-annual cycles as well as a linear trend. We do so by appending their coefficients to the vector $a$ and extending the matrices $E$, $R$, and $D$. For the design matrix,

$$
E \rightarrow \left( \begin{array}{c|c} E & A \\ \hline 0 & \end{array} \right)
$$

with

$$
A = \left( I_l \otimes X^{(T)}t \quad I_l \otimes X^{(T)} \cos \omega t \quad I_l \otimes X^{(T)} \sin \omega t \quad I_l \otimes X^{(T)} \cos 2\omega t \quad I_l \otimes X^{(T)} \sin 2\omega t \right),
$$

where $t$ contains the event times referenced to the midpoint between the first event and last event. For the prior covariance matrix,

$$
R \rightarrow \left( \begin{array}{cc} R & 0 \\ \hline 0 & h^2 (I_5 \otimes U^\Lambda) \Xi (I_5 \otimes U^\Lambda)^T \\ \end{array} \right),
$$

where $\Xi$ is a diagonal matrix containing the prior variances for the trend or seasonal amplitudes for the projections onto the singular vectors. Like for the random components, we assign no prior correlation between the trends and seasonal amplitudes of these projections. For the difference matrix,

$$
D \rightarrow \left( \begin{array}{c} D \\ B \end{array} \right),
$$

such that the trends and seasonal variations are included in the travel time anomalies $\tau$.

Once we specify the noise statistics $N = \sigma_n^2 I$ with $\sigma_n = 0.01$ s, we now have all information required for a Gauss–Markov estimate:

$$
\tilde{a} = P E^T N^{-1} \delta \quad \text{with} \quad P = (R^{-1} + E^T N^{-1} E)^{-1}.
$$

The desired estimate of the travel time anomalies is then $\tilde{\tau} = D \tilde{a}$, and the posterior covariance $P$ can be propagated to this estimate as $D P D^T$. Similarly, estimates of the projections onto singular vectors $\tilde{c}$, of the trends, and seasonal amplitudes of the signal can be obtained from $\tilde{a}$, and their uncertainties can be calculated from $P$.

The prescription $\sigma_n = 0.01$ s is meant to capture errors due to the changes in the source location between repeating events (Wu et al., 2020), due to noise affecting the cross-correlation.
<table>
<thead>
<tr>
<th></th>
<th>corr. time (days)</th>
<th>random (mK)</th>
<th>trend (mKyr$^{-1}$)</th>
<th>12 mo. (mK)</th>
<th>6 mo. (mK)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diego Garcia</td>
<td>15</td>
<td>15 10 5</td>
<td>4 2 1</td>
<td>15 10 5</td>
<td>15 10 5</td>
</tr>
<tr>
<td>Cape Leeuwin</td>
<td>30</td>
<td>20 5 5</td>
<td>4 2 1</td>
<td>20 5 5</td>
<td>5 2 2</td>
</tr>
</tbody>
</table>

Table S1: Parameters of the prior covariances of the projections onto the singular vectors. Where three parameters are given, they are for the three singular vectors. For the seasonal signals, the cos and sin terms are each assigned half the indicated prior variance.

function, and due to hydrophone motion. The moored CTBTO hydrophones are displaced by local currents, which contributes measurement error because the moorings are not navigated. Nichols and Bradley (2017) estimated that the CTBTO hydrophones at Wake Island, similar in design to the ones used here, can be displaced by 1/20 of the length of their mooring lines. The mooring lines for H08S2 and H01W2 are 562 m and 570 m long, respectively, producing maximum horizontal displacements of less than 30 m. A horizontal displacement of 30 m corresponds to a travel time anomaly of 0.02 s, which is equal to 2$\sigma_n$. A more refined prescription of $N$ might split the errors into their contributing factors and take the respective error correlations into account. We leave this refinement to future work.

Text S2  Cycle-skipping correction

Acoustic modes are dispersive. If the sounds speed profile is perturbed, a dispersive mode’s phase and group will shift by different amounts. If the differential shift between a mode’s phase and group is large enough, the correlation function will peak not at the correct phase shift but at the next peak over. The modal dispersion causes a cycle skip.

We here illustrate this process using a simple example of a Gaussian wave packet propagating through Munk’s canonical sound speed profile. The observed $T$-wave delays show a tell-tale sign of this process, suggesting that the same process occurs in the real system. The simple physics causing the cycle skipping allows for a robust identification and correction procedure, described below.

To illustrate the process causing cycle skipping, we consider modal propagation through a range-independent ocean with slowness profile $S(z)$, following Munk et al. (1995). A signal propagating a distance $L$ has a phase travel time $\tau = Ls$, where $s = k/\omega$ is the modal phase slowness that equals the slowness $S(z)$ at the turning depths. The modal structure $P$ is defined by

$$\frac{d^2 P}{dz^2} + (\omega^2 S^2 - k^2) P = 0$$

with boundary conditions $P = 0$ at $z = 0$ (the surface) and $dP/dz = 0$ at $z = -h$ (the bottom). We apply the normalization $\int dz P^2 = h$. The group slowness is

$$s_g = \frac{dk}{d\omega} = \frac{1}{sh} \int dz S^2 P^2,$$
where the integration is over the full depth. The mode’s group travel time is $\tau_g = L s_g$.

We now perturb the slowness profile $S(z)$ to $S(z) + \Delta S(z)$. The difference in the modal phase slowness between the reference and perturbed profiles is

$$\Delta s = \frac{1}{s h} \int dz P^2 S \Delta S.$$  \hspace{1cm} (17)

The change in group slowness, in contrast, is

$$\Delta s_g = \frac{1}{s h} \int dz \left[ \left(2 - \frac{s_g}{s} \right) P^2 + \omega (P^2)_\omega \right] S \Delta S,$$  \hspace{1cm} (18)

where $(P^2)_\omega = 2 PP_\omega$, and $P_\omega$ is defined by

$$\frac{d^2P_\omega}{dz^2} + \left(\omega^2 S^2 - k^2\right)P_\omega = -2 \left(\omega S^2 - ks_g\right) P$$  \hspace{1cm} (19)

with the same boundary conditions as for $P$ and the additional constraint that $\int dz P_\omega = 0$. Because variations in the slowness are much smaller than their mean values, the sensitivity profiles are roughly $P^2/h$ for the phase and $[P^2 + \omega (P^2)_\omega]/h$ for the group.

These phase and group sensitivity profiles thus have different structures (Fig. S2). We illustrate their behavior using the lowest mode at $\omega/2\pi = 2.5$ Hz for the canonical temperate profile of Munk (1974) and Munk et al. (1995; eq. 2.2.13). Above 1.6 km depth, the group sensitivity to sound speed perturbations is larger than the phase sensitivity. If the temperature anomalies that cause these sound speed perturbations are dominated by this depth range, as is typical, the group delay $\Delta \tau_g = L \Delta s_g$ will exceed the phase delay $\Delta \tau = L \Delta s$. If the group delay is more than a half period larger than the phase delay, cycle skipping is likely. In this scenario, cycle skipping is always forward: the skip goes in the same direction as the delay. It is important to correct this cycle skipping because otherwise large anomalies would be overestimated systematically.

As an example, consider a slowness perturbation that is caused by a single low dynamical mode (Fig. S2). For the buoyancy frequency profile $N = N_0 e^{z/d}$ underlying the canonical sound speed profile, the sound speed perturbation is proportional to $N^2 G$. The dynamical mode $G$ for vertical displacement is defined by

$$\frac{d^2G}{dz^2} + \frac{N^2}{c^2} G = 0$$  \hspace{1cm} (20)

with $G = 0$ at $z = 0$ and $z = -h$, and a normalization $\int dz N^2 G^2 = h$ is applied. We use $h = 5$ km, but this has little impact on the results because neither the first acoustic mode nor the first dynamical mode has much amplitude in the bottom 1 km or so. For the first dynamical, $\Delta \tau_g/\Delta \tau = 1.61$, so cycle skipping is likely for $|\Delta \tau| > 0.33$ s. This is about where we observe cycle skipping to occur in the real $T$ waveforms received at Cape Leeuwin.

To further illustrate the cycle skipping, we consider a signal with the source function $e^{-\sigma z^2} \cos \omega t$. We again use the frequency $\omega/2\pi = 2.5$ Hz, and we impose a bandwidth $\sigma/2\pi = 0.5$ Hz corresponding to the filtering we use for the real data. We then propagate this signal
Figure S1: Sensitivity kernels calculated using two-dimensional numerical calculations as well as assuming the propagation of the first acoustic mode through a range-independent ocean. These kernels are shown for all considered frequencies and for both paths.

assuming a dispersion relation linearized around $\omega$. We impose phase lags $\Delta \tau = 0.0$ s, 0.3 s, and 0.6 s and group lags $\Delta \tau_g = 1.61 \Delta \tau$, corresponding to propagation through an anomaly caused by the first dynamical mode (Fig. S3). Correlating the delayed signals with the signal without lag, the signal with $\Delta \tau = 0.3$ s produces a maximum of the correlation function at the correct phase lag. But the signal with $\Delta \tau = 0.6$ s experiences cycle skipping: the maximum of the correlation function shifts by one period to a lag of 1.0 s. The secondary maximum to the left is at the correct lag. If cycle skipping can be identified, it can easily be corrected by picking the adjacent peak in the correlation function.

This type of cycle skipping caused by dispersive modal propagation can be identified in the real data. If the first acoustic mode dominates and temperature anomalies are surface-intensified, the cycle skipping causes an increase in the magnitude of the low-frequency (say, 2.5 Hz) phase delay. At the same time, it causes a decrease in the differential delay between a higher frequency (say, 4 Hz) and this low frequency. Plotting the measured low-frequency and differential delays against one another reveals three distinct clusters: a center cluster with small low-frequency delays that is not affected by cycle skipping and adjacent clusters with larger-magnitude low-frequency delays that suffer from forward or backward cycle skipping (Fig. S4, S5).
Figure S2: Illustration of dispersive propagation of low-frequency acoustic modes. Shown are profiles of (a) the buoyancy frequency $N$, (b) Munk’s canonical slowness profile $S$ (blue) and the phase slownesses $s$ of the first acoustic mode (solid orange) and two higher modes (dotted orange), (c) the first acoustic mode (solid) and two higher modes (dotted), (d) the phase and group sensitivities $P^2S/sh$ (green) and $[(2-s_g/s)P^2 + \omega(P^2)^2]S/sh$ (red) for the first acoustic mode, and (e) the first dynamical mode (solid) and two higher modes (dotted).

We identify pairs in these clusters by employing a Gaussian mixture model with three components that have shared covariances. We correct pairs in the left cluster by shifting to the next maximum of the correlation function to the right of the original maximum, and we correct pairs in the right cluster by shifting to the next maximum to the left. We then probe for additional cycle skips (or reversals of corrections applied after the cluster analysis) by looking for reductions in the cost of the inversion for travel time anomalies (cf. Wu et al., 2020). We allow such additional corrections if the Gaussian mixture model indicates a probability greater than $0.1\%$ to belong to a cluster different from that identified as most likely. For the path to Diego Garcia, this procedure corrects 40 pairs to the left and 56 pairs to the right (out of a total of 3831 used pairs). For the path to Cape Leeuwin, cycle skipping is more common, both because the path is longer and because travel time anomalies are somewhat larger. Here, 230 pairs are corrected to the left, and 393 pairs are corrected to the right (out of a total of 3032 used pairs).

References


Figure S3: Illustration of cycle skipping using synthetic signals. The top panel shows three signals that have experienced different amounts of phase lags $\Delta \tau$ and group lags $\Delta \tau_g = 1.61\Delta \tau$. The triangles trace the peak that is at the center of the signal without lag, illustrating the phase shift. The bottom panel shows the cross-correlation functions between the lagged signals and the reference signal without lag. The maximum of the cross-correlation function occurs at the correct phase lag (marked by triangle) for the signal with $\Delta \tau = 0.3$ s. Cycle skipping occurs for the signal with $\Delta \tau = 0.6$ s: the maximum (marked by circle) is offset by one period from the correct lag, which coincides with a secondary maximum (marked by triangle).


Figure S4: Cycle-skipping corrections for the path to Diego Garcia. Shown are the low-frequency delays (2.0 Hz) vs. the differential delays (4.0 Hz minus 2.0 Hz). The green dots indicate identified cycle skipping that is corrected to the right, the blue dots indicate cycle skipping corrected to the left, and orange dots indicate pairs not identified as being affected by cycle skipping. The identifications are (a) after the cluster analysis and (b) after addition corrections (or reversals) based on the inversion cost.
Figure S5: Cycle-skipping corrections for the path to Cape Leeuwin. Shown are the low-frequency delays (2.5 Hz) vs. the differential delays (4.0 Hz minus 2.5 Hz). The green dots indicate identified cycle skipping that is corrected to the right, the blue dots indicate cycle skipping corrected to the left, and orange dots indicate pairs not identified as being affected by cycle skipping. The identifications are (a) after the cluster analysis and (b) after additional corrections (or reversals) based on the inversion cost.
Figure S6: Temperature anomalies for the path to Diego Garcia for the early part of the observed period. Shown are (b), (d) the anomalies from the Argo and ECCO products and (a), (c), (e) the projections onto the first two singular vectors inferred from $T$ waves and applied to the Argo and ECCO products. Note the arsinh-scaled color map.
Figure S7: Temperature anomalies for the path to Cape Leeuwin for the early part of the observed period. Shown are (b), (d) the anomalies from the Argo and ECCO products and (a), (c), (e) the projections onto the first two singular vectors inferred from $T$ waves and applied to the Argo and ECCO products. Note the arsinh-scaled color map.
Figure S8: Temperature anomalies for the path to Diego Garcia projected onto the first two singular vectors. Shown are (a) the time series inferred from T-wave data, (b) the projections of anomalies from the Argo product, and (c) projections of anomalies from the ECCO product.
Figure S9: Temperature anomalies for the path to Cape Leeuwin projected onto the first two singular vectors. Shown are (a) the time series inferred from T-wave data, (b) the projections of anomalies from the Argo product, and (c) projections of anomalies from the ECCO product.