OBSRANGE: A NEW TOOL FOR THE PRECISE REMOTE LOCATION OF OCEAN BOTTOM SEISMOMETERS

RUSSELL J., EILON Z., AND MOSHER S.

1. Introduction

The last two decades have seen a sea change in the longevity, distribution, and sophistication of temporary ocean bottom seismic installations. The proliferation of ocean bottom seismometer (OBS) deployments has opened up new possibilities for understanding the ocean basins, continental margins, and even inland submerged environments.

However, even straightforward OBS installations present several unique challenges. Foremost among these is the inability to directly measure the location of the sensor at the seafloor. Precise knowledge of station location is essential for almost all seismological analysis. While the location of the ship can be determined with exactitude at the time of deployment, OBS instruments are found to drift by up to hundreds of meters from this point due to water currents and a non-streamlined basal profile.

For broadband OBS deployments, it has long been accepted practice to conduct an acoustic survey in order to triangulate the position of the instrument. To accomplish this, ships send non-directional acoustic pulses into the water column. These are received by the OBS transponder which sends its own acoustic pulse in response. The time elapsed between the ship sending and receiving acoustic pulses is proportional to distance, which (for known ship location) may be used to locate the instrument. It is common for this analysis to be conducted by technicians at OBS instrument centers and provided latterly to PIs and data centers as station metadata. Some codes are proprietary intellectual property of the instrument centers, and others are available for a license fee.

However, standard station location algorithms to date are lacking in certain respects. Water sound speed, "turn-around time" (processing time taken by the OBS transponder between receiving and sending acoustic pulses), and even water depth are often assumed *a priori*. Commonly, no correction is made for the movement of the ship. Robust uncertainty analysis, which would allow practitioners to gauge potential location errors, is either not conducted or communicated.

We present an open-source OBS locator code for use by the marine geophysical community. Our efficient inversion algorithm provides station location in three dimensions, as well as solving for depth-averaged water sound speed and "turn-around time". We use statistical tools to provide robust uncertainties on the station location. We have made the code available in both MATLAB and PYTHON to promote accessibility. In this article we present the theory behind our algorithm, validate the inversion using synthetic testing, demonstrate its utility with real data, and analyse a variety of location survey patterns so as to inform the planning of future OBS experiments.

2. Algorithm

2.1. **The forward problem.** We wish to locate an instrument which rests at unknown position and depth on the ocean floor. Taking the drop coordinates as the center of a Cartesian coordinate system in which x is positive towards East, y is positive towards North, and z is positive upwards from the sea surface, the instrument lies at location (x_O, y_O, z_O). The time taken for an acoustic pulse to travel from the ship to the instrument and back is a function of the sound speed in water (V_P), and the location of the ship, as well as the "turn-around time" (τ) that corresponds to the (fixed) processing time between the OBS transducer receiving a ping and sending its response. In detail, we must account for the possibility that if the ship is under way, its position changes between sending and receiving pings. Thus, the total travel time, T, is:

$$T = \frac{r_s + r_r}{V_P} + \tau \tag{1}$$

where

$$r_s = \sqrt{(x_s - x_O)^2 + (y_s - y_O)^2 + z_O^2} \tag{2}$$

$$r_r = \sqrt{(x_r - x_O)^2 + (y_r - y_O)^2 + z_O^2} \tag{3}$$

where subscript "s" indicates the ship sending a ping and "r" indicates the ship receiving the OBS's response. These positions are related by the velocity ($\mathbf{u} = (u_x, u_y, 0)$) of the ship:

$$\begin{pmatrix} x_s \\ y_s \\ 0 \end{pmatrix} = \begin{pmatrix} x_r \\ y_r \\ 0 \end{pmatrix} - T \begin{pmatrix} u_x \\ u_y \\ 0 \end{pmatrix} \tag{4}$$

It follows that, to a close approximation,

$$\begin{aligned} r_s &\approx r_r - T\,(\mathbf{u} \cdot \hat{\mathbf{r}}_r) \\ &= r_r - \delta r \end{aligned}$$

where $\hat{\mathbf{r}}_r$ is the unit-vector pointing from the instrument to the ship at the time of receiving. If we know the distance δr we can account for the send-receive timing offset related to a change in ship's position, by computing a correction time, $\delta T = \delta r / V_P$. Substituting this into equation (1), we have

$$T + \delta T = \frac{2 r_r}{V_P} + \tau \tag{5}$$

2.2. **The inverse problem.** If travel times are known between the OBS and certain locations, but the position of the OBS is not, equation (5) can be thought of as a non-linear inverse problem, of the form $\mathbf{d} = g(\mathbf{m})$, where $g(\mathbf{m})$ represents the forward-model. The model contains five parameters: $\mathbf{m} = \{x_O, y_O, z_O, V_P, \tau\}$. The data, $\mathbf{d}$, are a vector of corrected travel times, $T + \delta T$ (note that δT is itself a function of $\mathbf{m}$; this will be adjusted iteratively). Uncorrected travel-time residuals predicted from the starting model with magnitude >500 ms are considered anomalous and are removed before beginning the inversion. This type of problem can be solved iteratively using Newton's method (**??**):

$$\mathbf{m}_{k+1} = \mathbf{m}_k + \left[\mathbf{G}^{\mathrm{T}}\mathbf{G}\right]^{-1}\mathbf{G}^{\mathrm{T}}\left(\mathbf{d} - g(\mathbf{m}_k)\right) \tag{6}$$

where $\mathbf{G}$ is a matrix of partial derivatives: $G_{ij} = \partial d_i / \partial m_j$, as follows:

$$\begin{aligned}
\frac{\partial d_i}{\partial x_O} &= -\frac{2(x_i - x_O)}{V_P\, r_i} \\
\frac{\partial d_i}{\partial y_O} &= -\frac{2(y_i - y_O)}{V_P\, r_i} \\
\frac{\partial d_i}{\partial z_O} &= \frac{2 z_O}{V_P\, r_i} \\
\frac{\partial d_i}{\partial V_P} &= -\frac{2\, r_i}{V_P^2} \\
\frac{\partial d_i}{\partial \tau} &= 1
\end{aligned}$$

We use the drop coordinates and water depth (if available from multibeam) as a starting model, along with $V_P = 1500$ m/s and τ =13 ms. If we consider the setup of the problem, there is some degree of trade off between the water depth and the water velocity. Simplistically, if all survey measurements are made at a constant distance from the station (*e.g.*, if the survey is a circle centered on the station) then these parameters co-vary perfectly. As a result, the inverse problem is ill-posed and, like all mixed-determined problems, requires regularization. We use constraint equations to damp perturbations in V_P, which is not likely to vary substantially from 1500 m/s, and τ, which should not vary substantially from ∼13 ms (Ernest Aaron, *pers. comm.*):

$$\mathbf{F} = \begin{bmatrix}\mathbf{G}\\ \mathbf{H}\end{bmatrix} \qquad \mathbf{f} = \begin{bmatrix}\mathbf{d} - g(\mathbf{m})\\ \mathbf{0}\end{bmatrix} \tag{7}$$

where

$$\mathbf{H} = \begin{pmatrix} 0 & & & & \\ & 0 & & & \\ & & 0 & & \\ & & & \gamma_{V_P} & \\ & & & & \gamma_\tau \end{pmatrix} \tag{8}$$

We have had success using $\gamma_{V_P} = 5 \times 10^{-8}$ and $\gamma_\tau = 0.2$. Finally, we implement global norm damping to stabilize the inversion, through parameter $\epsilon = 10^{-10}$, such that the equation to be solved becomes:

$$\mathbf{m}_{k+1} = \mathbf{m}_k + \left[\mathbf{F}^{\mathrm{T}}\mathbf{F} + \epsilon\mathbf{I}\right]^{-1}\mathbf{F}^{\mathrm{T}}\mathbf{f} \tag{9}$$

This equation is solved iteratively, until the root-mean-squared (RMS) of the misfit $(T + \delta T - g(\mathbf{m})|)$ decreases by less than 0.1 ms compared to the previous iteration. This criterion is usually reached after ∼4 iterations.

2.3. **Errors and uncertainty.** In order to estimate the uncertainty in our model we bootstrap survey timing data with a balanced resampling approach. In each iteration the algorithm inverts a random sub-sample of the true data set, with the constraint that all data points are eventually sampled an equal number of times. This approach provides empirical probability distributions of possible model parameters, but does not straightforwardly offer quantitative estimates of model uncertainty because the goodness of data fit for each run in the bootstrap iteration is ignored (that is, within each iteration, a model is found that best fits the randomly sub-sampled dataset, but in the context of the full dataset, the fit and uncertainty of that particular model may be very poor). For more statistically robust uncertainty estimates, we perform a grid search over (x_O, y_O, z_O) within a region centred on the bootstrapped mean location, $(x_{O_{\mathrm{best}}}, y_{O_{\mathrm{best}}}, z_{O_{\mathrm{best}}})$. For each perturbed location, (x'_O, y'_O, z'_O), we use an F-test to compare the norm of the data prediction error to the minimum error, assuming they each have a χ^2 distribution. The effective number of degrees of freedom, ν is calculated as

$$\nu = N_f - \mathrm{tr}(\mathbf{F}\mathbf{F}_{\mathrm{inv}}) \tag{10}$$

where $\mathbf{F}_{\mathrm{inv}} = \left[\mathbf{F}^{\mathrm{T}}\mathbf{F} + \epsilon\mathbf{I}\right]^{-1}\mathbf{F}^{\mathrm{T}}$, N_f is the length of vector $\mathbf{f}$, and tr() denotes the trace. Using the F-test, we can evaluate the statistical probability of the true OBS location departing from our best-fitting location by a given value.

Some care is required in implementing this grid search. Since z_O covaries with V_P, varying z_O quickly leads to large errors in data prediction as $|z'_O - z_{O_{\mathrm{best}}}|$ increases if one holds V_P fixed. As a result, it appears as if the gradient in the error surface is very sharp in the z direction, implying this parameter is very well resolved; in fact, the opposite is true. We find the empirical covariance of z_O, V_P, and τ by performing principal component analysis on the bootstrap model solutions. We then use the largest eigenvector to project perturbations in z_O within the grid search onto the other two parameters, adjusting them appropriately as we progress through the grid search.

3. Results

3.1. **Demonstration on synthetic data.** We validated our algorithm by checking that it correctly recovers the (known) location of synthetic test stations. Synthetic two-way travel times were computed by interpolating the ship's position within a fixed survey pattern at one-minute intervals, sending straight-line rays to the instrument and back, and adding the turn-around time. This travel time includes the change in ship's position between sending and receiving; since the position of the ship at the time it receives the acoustic pulse is itself dependent on the travel time, we iterated on this value until the time and position converged to give an error of $< 10^{-6}$ s. Only the location and absolute time at the time the ship receives the acoustic pulse was recorded for the inversion, mimicking the data obtained from the EdgeTech deck box. We then added Gaussian random noise to the resultant travel times using a standard deviation of 4 ms, to account for measurement noise, errors in ship GPS location, and local changes in water velocity. Lastly, we randomly dropped out ∼20% of the travel time data points, simulating the occasional null return from the acoustic survey. This testing procedure was designed to mimic the idiosyncrasies of real acoustic surveys as closely as possible.

Figure 1 shows the result of an inversion at a single station. For this inversion, we included a correction for a Doppler shift introduced by the ship's motion, estimating ship velocity from the timing and location of survey points. The inversion was successful in locating the OBS station: the estimated location is 3.02 m from the the true location (Figure 1). This misfit is extremely small in the context of ∼320 m of drift, a survey radius of ∼3700 m, and a water depth of ∼5300 m. Moreover, the true location falls well within the uncertainty bounds estimated from the F-test and the bootstrap analysis.

In order to obtain statistics on the general quality of the synthetic recovery, we performed this test for 10,000 synthetic OBS stations, as follows: For each iteration, a synthetic station location was determined relative to a fixed drop point by drawing x- and y- drifts from zero-centered Gaussian distributions with standard deviations of 100 m (only in rare cases are stations thought to drift further than ∼200 m). The depth perturbation, turn-around time, and water velocity perturbation were similarly randomly selected, with mean values of 5000 m, 13 m/s, and 1500 m/s and standard deviations of 50 m, 3 ms, and 10 m/s, respectively. For tests of the basic location algorithm, we held the survey geometry constant, using the *PACMAN* configuration with a radius of 1 Nm (Section 3.4).

The results of these tests show that on average our inversion is highly successful in correctly locating the OBS stations. The mean location errors in the x-, y-, and z- directions were 0.038 m, 0.152 m, and -0.599 m respectively, demonstrating there was no systematic bias in the locations. The mean errors in water velocity and turn-around time were indistinguishable from zero, showing that estimation of these parameters was also not biased. The mean absolute horizontal location error was 2.31 m, with a standard deviation of 1.22 m. 95% of the absolute horizontal station location errors were less than 4.58 m. There was no relationship observed between station drift (*i.e.*, the distance between the synthetic OBS station and the drop point) and the location error, indicating that as long as stations settle within the survey bounds they will be well located. A corollary to this observation is that location estimates should not be biased by incorrectly recorded drop locations.

We observed a strong trade-off between water velocity and depth, which was responsible for the somewhat larger standard error in station depth estimates, which was 9.6 m. This uncertainty is likely of negligible concern for most OBS practitioners, but if precise depths are important then a survey geometry that includes more tracks towards and away from the station would be preferable (in addition to verification using acoustic echo-sounders that implement precise water-velocity profiles from XBT data).

3.2. **Application to PacificArray deployment.** We applied the location algorithm to acoustic surveys carried out during the *Young Pacific ORCA (OBS Research into Convecting Asthenosphere)* deployment in the central Pacific ocean during April and May of 2018. The OBS array comprised 30 SIO broadband instruments deployed from the R/V Kilo Moana in water depths of $\sim$4400-4800 m. Acoustic surveys were carried out using an EdgeTech 8011M Acoustic Transceiver command and ranging unit, attached to a hull-mounted 12 kHz transducer. The relatively calm seas allowed for ideal survey geometry at almost all sites, with a ship speed of $\leq$8 knots at a maximum radius of $\sim$1.3 Nm.

An example station inversion, as well as the graphical outputs of the location software, is shown in Figures 2-4. Ship velocity is estimated from the survey data by differencing survey points. In theory, this could be used to correct doppler shifts (Figure 2c) in travel time (as in the synthetic tests), but we found that this correction did not substantially improve data fit for real stations and so did not apply it to this data set, although it is included as an option in the location codes. The small RMS data misfit of $\sim$1.6 ms attests to the quality of the survey measurements and the appropriateness of our relatively simple location algorithm. The southwestwards drift of $\sim$ 340 m demonstrates that ocean currents can substantially displace the final OBS location from their surface drop point.

The 30 stations in this array drifted an average distance of 170 m. The mean data RMS misfit was 1.96 ms and the estimated 95% percentile location error based on the bootstrap analysis was 1.13 m. The water depth estimated by the inversion was systematically shallower than that measured using the shipboard multibeam instrument, differing by an average value of 18.6 m. Assuming the multibeam depths, which are computed using a water sound speed profile that is validated daily by XBT measurements, are correct, this discrepancy indicates that the inversion systematically overestimates sound speed slightly.

Without accurate seafloor corroboration from an ROV, it is not possible to directly verify the locations of stations within the Pacific ORCA array. However, we obtain indirect support for the success of the location algorithm by considering the drift of all stations within this array (Figure **??**). Taken together, the direction and magnitude of drift depicts a pattern of clockwise rotation with a minimum diameter of $\sim$500 km. This system appears to correlate with a meso-scale ocean gyre, with a direction, location, and approximate size that is consonant with large-scale patterns of geostrophic flow modeled in this location within the time frame of our deployment. We speculate that a large fraction of the OBSs' lateral drift is achieved in the upper $\boldsymbol{X}$ m of the ocean, where horizontal flow amplitudes are highest [#REF]. The fact that we are able to discern this pattern from our estimated locations is a testament to the accuracy of the OBSrange algorithm. This observation also raises the intriguing possibility of using OBS instruments as ad hoc depth-integrated flow meters for the oceans.

3.3. Comparison to previous tools.

- Re-run synthetics with damped Vw, etc.
- Talk about null points
- Show difference in error between our method and theirs
- latlon2xy conversion

We compared our location algorithm with a tool developed by engineers at Scripps Institution of Oceanography (SIO) that has previously been used to locate OBS on the seafloor. This unpublished tool, hereby referred to as *SIOgs*, performs a grid search in x–y holding z fixed at the reported drop-point depth and assumes a water velocity of 1500 m/s and turn-around time of 13 ms. The grid search begins with grid cells of 100×100 m and iteratively reduces their size to 0.1×0.1 m. In contrast to our algorithm, *SIOgs* does not account for: 1) the δT (Doppler) correction due to the changing ship position between sending and receiving, 2) the ellipsoidal shape of the earth when converting between latitude–longitude and x–y, 3) variations in z, τ, and V_p, and 4) automated identification and removal of low-quality travel-time data. Furthermore, *SIOgs* provides no information about uncertainty or resolution of model parameters.

To quantitatively compare our algorithm with *SIOgs*, as well as the importance of the 4 additional features that our algorithm includes, we performed 9 separate inversions of a synthetic dataset for a *PACMAN* survey geometry with 1 Nm radius (Figure 5). The inversions include the complete *OBSrange* algorithm as well as several variants where features have been damped or removed to assess their importance; details of the inversions are given in Table 1. Our algorithm locates the station to within ∼1.5 m of the true location, while *SIOgs* locates it ∼42 m from the true position, far beyond the 95% F-test contour (Figure 5a).

The grid search method is very susceptible to mislocations due to anomalous data. Inversion *SIOgs no QC* in Figure 5 includes a single anomalous travel-time measurement 2000 ms off from its true value, causing the station to be mislocated by ∼130 m. Although such outliers can be manually removed, they may easily be overlooked. *OBSrange* includes an automatic quality control step based on travel-time residuals of the starting location that removes anomalous residuals with magnitudes >500 ms.

In addition to showing the full potential of *OBSrange*, we demonstrate the importance of including the "Doppler" correction (δT in equation (5)), and of accounting for Earth's ellipsoidal shape when converting latitude and longitude to x–y.

3.4. Exploration of survey pattern geometries.

4. Discussion

5. Conclusion

6. Figures and Tables

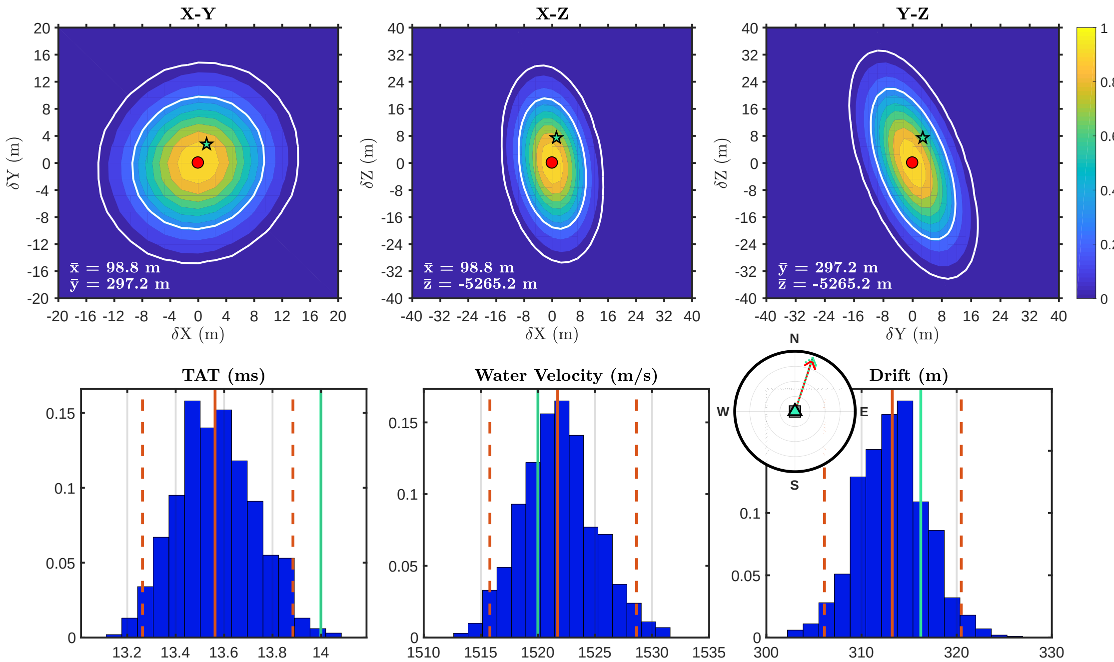

Figure 1. Test of location algorithm using synthetic data. A comparison of the true input values (green star and lines) with the inverted model parameters (red circle and red solid lines) demonstrates that the location, depth, and water velocity are extremely well recovered, and the estimated uncertainties on these parameters are consonant with the actual misfit. Top three plots show slices through the F-test surface, contoured by probability. Bottom three plots show histograms from a bootstrap analysis with 95th percentile values indicated by dashed red lines. Inset shows the direction of true (green dashed) and estimated (red) drift with respect to the starting location.

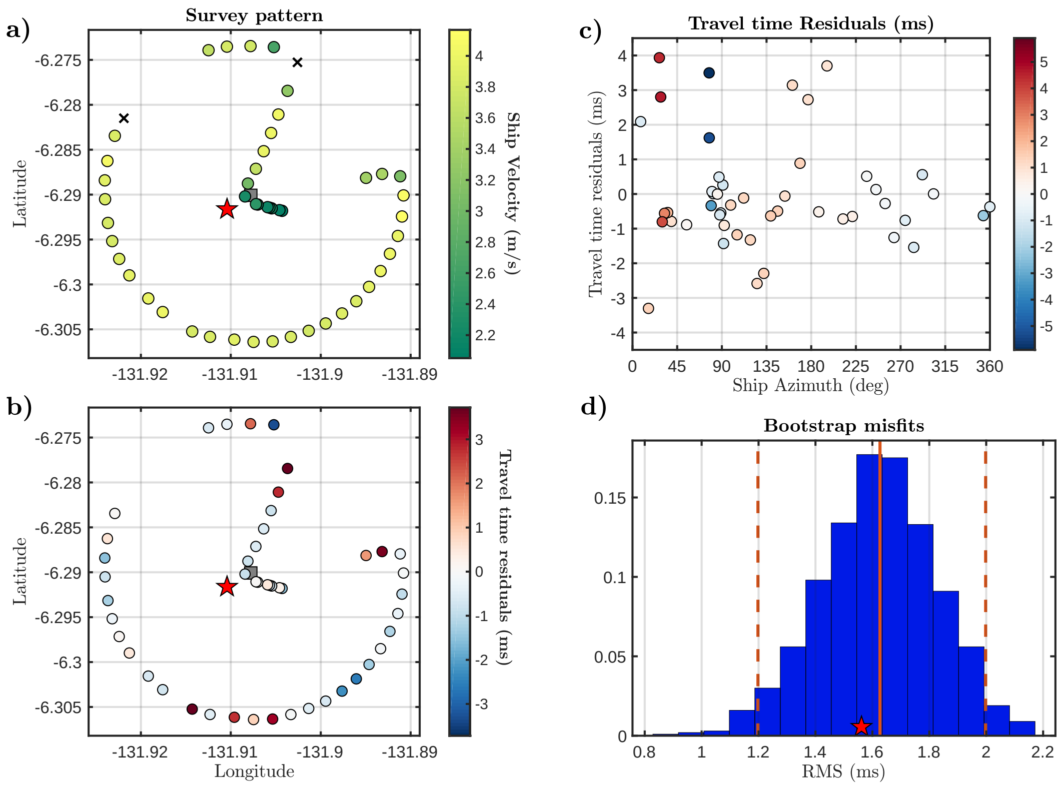

FIGURE 2. Example inversion at station EC03 in the 2018 Young Pacific ORCA deployment. a) Map view of acoustic survey; colored circles are successful acoustic range measurements, black crosses are bad measurements rejected by automatic quality control, grey square is drop location, red star is final location. b) Map view of data residuals based on travel times computed using bootstrap mean station location. c) Data residuals plotted as a function of azimuth, colored by the computed doppler correction (not used in this inversion). d) Histogram of data RMS from the bootstrap; the RMS of the final model is shown as a red star.

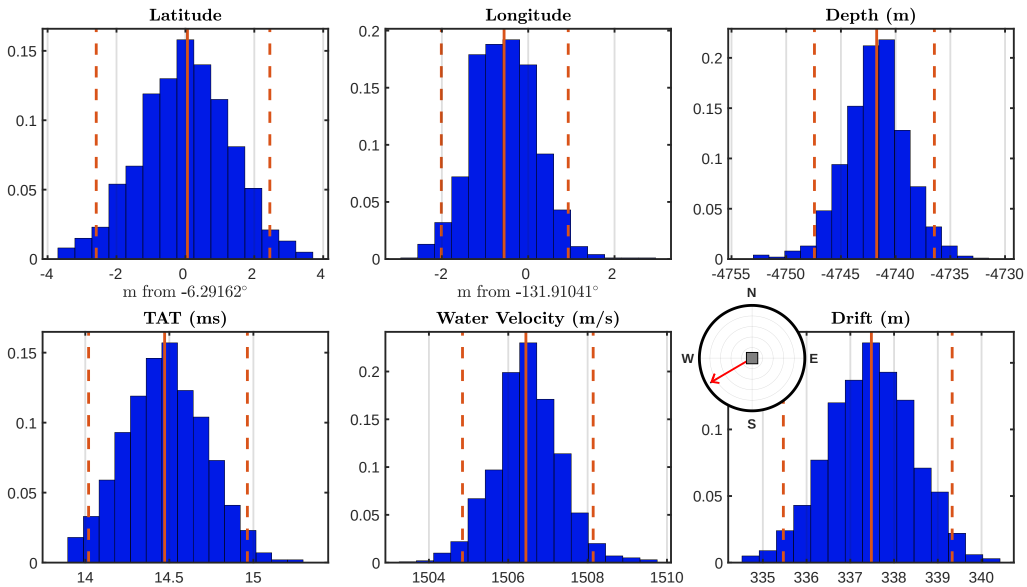

FIGURE 3. Histograms of model parameters from the bootstrap inversion of station EC03 in the 2018 Young Pacific ORCA deployment. Red solid line shows mean value, while dashed lines indicate 95th percentiles. Latitude and longitude are plotted in meters from the mean point, for ease of interpretation. The inset plot shows the mean drift azimuth from the drop location (grey square).

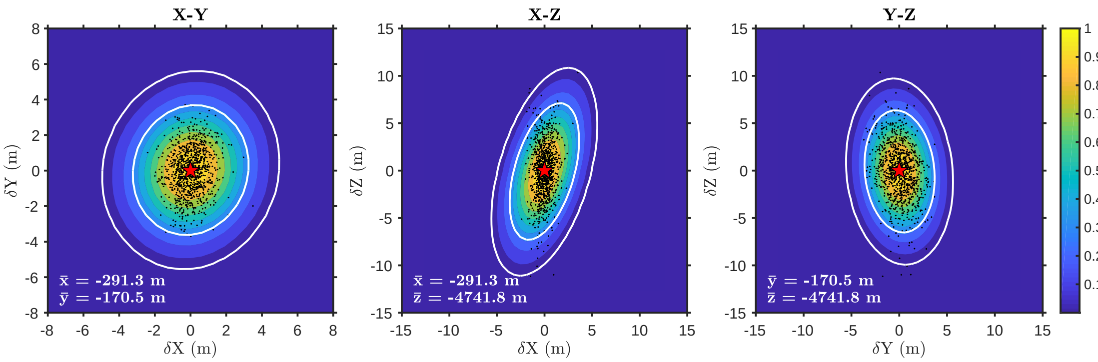

Figure 4. Three orthogonal slices through the F-test probability volume for station EC03 in the 2018 Young Pacific ORCA deployment, contoured by probability of true station location relative to the best fitting inverted location $(\bar{x}, \bar{y}, \bar{z})$, indicated by the red star. White contours show 68% and 95% contours. Black dots show individual locations from the bootstrap analysis (Figure 3).

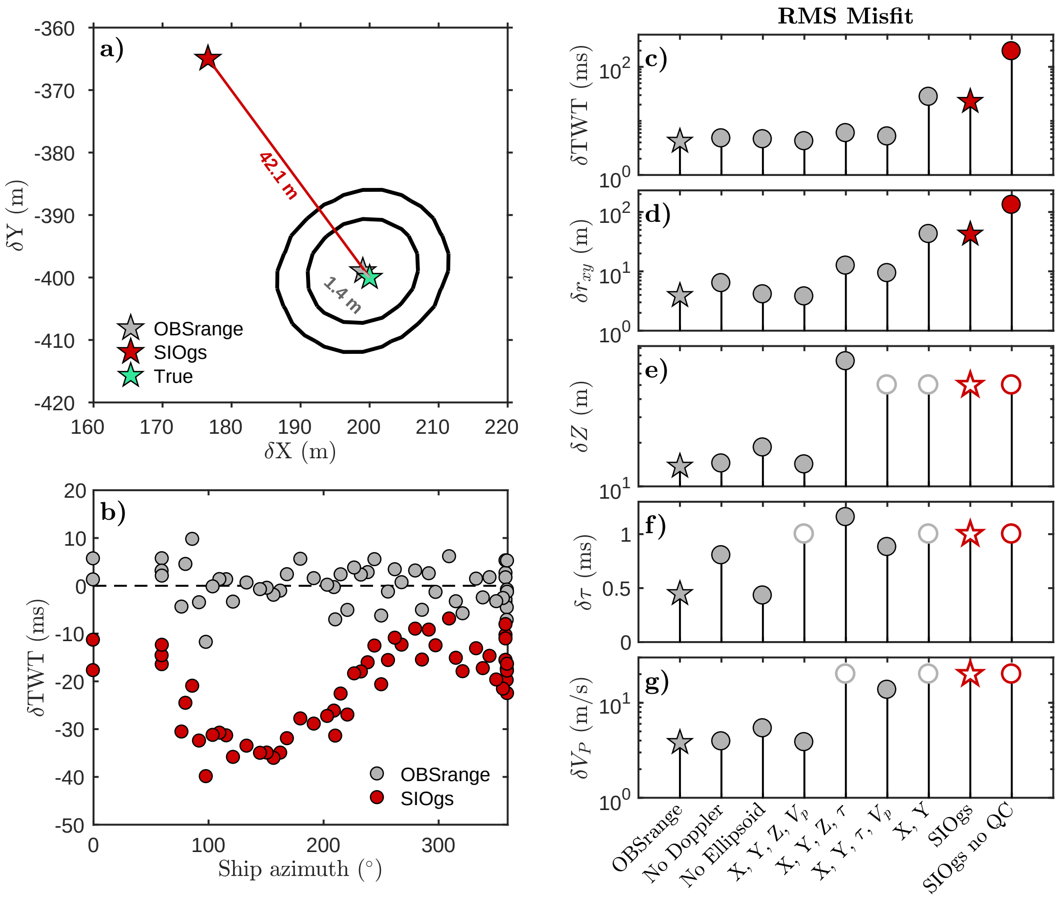

FIGURE 5. Synthetic test of OBSrange performance (gray symbols) compared with the SIO tool (red symbols) for a *PACMAN* survey of radius 1 Nm. a) Map view comparing the OBSrange and SIO inverted instrument locations with the true location in green. Black contours show the 68% and 95% confidence from the OBSrange F-test. b) Two-way time (TWT) residuals for both methods as a function of ship azimuth from the true station location. c) TWT and d–g) model parameter RMS misfits for 9 inversions, where closed symbols represent parameters that are solved for in the inversion and open symbols are parameters that remain fixed throughout the inversion. The mislocation in x–y is given by $\delta r_{xy} = \sqrt{\delta x_O^2 + \delta y_O^2}$. Stars in c–g mark the inversions shown in a) and b). See table 1 for details of the 9 inversions.

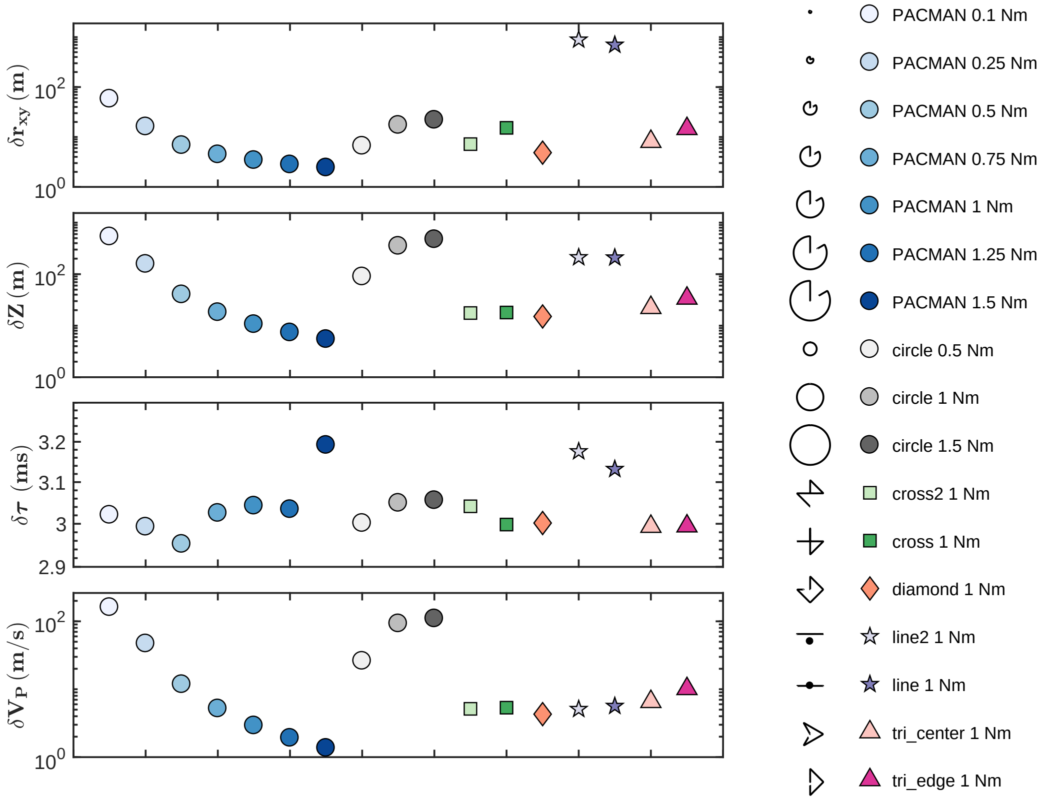

Figure 6. Model parameter RMS misfits for 6 synthetic survey geometries of varying radii: *PACMAN*, circle, cross, diamond, line, and triangle. Each survey geometry is shown to the left of its respective legend entry. Panels show the RMS misfit for each model parameter and survey type for 10,000 synthetic survey realizations, where the mislocation in x–y is given by $\delta r_{xy} = \sqrt{\delta x_O^2 + \delta y_O^2}$. We find that model parameters are most accurately recovered using the *PACMAN* survey pattern with radius $\geq$1 Nm, and the line surveys perform worst.

Table 1. Details of the synthetic tests from Figure 5. Final model parameters are the median of 1000 bootstrap iterations and are omitted if held fixed during the inversion. Parameters x and y are displayed as distance from the drop location.

Model Name				**x (m)**	**y (m)**	**z (m)**	**τ (ms)**	**V_p (m/s)**
OBSrange	**method**	Newton's	**initial**	0	0	-5000	13.0	1500
	δT correction	Yes	**final**	199	-399	-5055	13.8	1521
	ellipsoid correction	Yes	**true**	200	-400	-5050	14.0	1520
	remove bad data	Yes	**RMS**	3.0	2.4	13.8	0.4	3.8
No Doppler	**method**	Newton's	**initial**	0	0	-5000	13.0	1500
	δT correction	No	**final**	201	-395	-5050	13.4	1520
	ellipsoid correction	Yes	**true**	200	-400	-5050	14.0	1520
	remove bad data	Yes	**RMS**	3.2	5.5	14.4	0.8	3.9
No Ellipsoid	**method**	Newton's	**initial**	0	0	-5000	13.0	1500
	δT correction	Yes	**final**	200	-398	-5063	14.0	1524
	ellipsoid correction	No	**true**	200	-400	-5050	14.0	1520
	remove bad data	Yes	**RMS**	2.9	2.8	18.5	0.4	5.4
X, Y, Z, V_p	**method**	Newton's	**initial**	0	0	-5000	13.0	1500
	δT correction	Yes	**final**	199	-399	-5057	-	1522
	ellipsoid correction	Yes	**true**	200	-400	-5050	14.0	1520
	remove bad data	Yes	**RMS**	2.9	2.4	14.2	1.0	3.8
X, Y, Z, τ	**method**	Newton's	**initial**	0	0	-5000	13.0	1500
	δT correction	Yes	**final**	192	-391	-4977	12.8	-
	ellipsoid correction	Yes	**true**	200	-400	-5050	14.0	1520
	remove bad data	Yes	**RMS**	8.7	8.9	72.9	1.2	20.0
X, Y, τ, V_p	**method**	Newton's	**initial**	0	0	-5000	13.0	1500
	δT correction	Yes	**final**	194	-394	-	13.1	1506
	ellipsoid correction	Yes	**true**	200	-400	-5050	14.0	1520
	remove bad data	Yes	**RMS**	6.6	6.6	50.0	0.9	13.7
X, Y	**method**	Newton's	**initial**	0	0	-5000	13.0	1500
	δT correction	Yes	**final**	173	-371	-	-	-
	ellipsoid correction	Yes	**true**	200	-400	-5050	14.0	1520
	remove bad data	Yes	**RMS**	29.0	30.8	50.0	1.0	20.0
SIOgs	**method**	Grid Search	**initial**	0	0	-5000	0.0	1500
	δT correction	No	**final**	177	-365	-	-	-
	ellipsoid correction	No	**true**	200	-400	-5050	14.0	1520
	remove bad data	Yes	**RMS**	23.4	35.0	50.0	1.0	20.0
SIOgs no QC	**method**	Grid Search	**initial**	0	0	-5000	0.0	1500
	δT correction	No	**final**	320	-453	-	-	-
	ellipsoid correction	No	**true**	200	-400	-5050	14.0	1520
	remove bad data	No	**RMS**	120.1	53.4	50.0	1.0	20.0