Dear Dr. Peng and Associate Editor,

We thank you and the two anonymous reviewers for constructive comments toward improving our manuscript "OBSrange: A new tool for the precise remote location of Ocean Bottom Seismometers". Our response to each reviewer comment is in red below. Based on reviewers' suggestions, the main improvements made to the manuscript and/or OBSrange code include
1) the ability to account for refraction through the water column using an automatically selected sound-speed profile appropriate for any deployment region, drawn from the 2009 World Ocean Atlas database;
2) a correction for a known horizontal shipboard GPS-transponder offset;
3) a discussion of water depth and its effects on uncertainties, optimal survey size, and the ray-bending correction;
4) a more complete reference to the ellipsoid correction and demonstrations of its importance for reducing travel-time misfit for both real and synthetic data
5) a new Figure 1 detailing coordinate systems and conventions referred to in the body of the code.
All together, the revisions entail 8 new supplementary figures. We have also uploaded for this round of reviews a preliminary User Manual (*README.pdf*), which outlines the basic structure of the code (inputs, options, outputs) and will be included in the OBSrange package upon its official release.

Regards,
J. Russell, Z. Eilon, & S. Mosher

AE's comment:
The manuscript presents a way to locate OBSs and can be useful for the OBS community. Both reviewers, however, pointed out major issues that are needed to be addressed. The most notable issues include non-constant velocity model and uncertainty analyses. Major revision is needed before it can be accepted. Please do carefully go through the detail reviews and revised the manuscript accordingly.
See our responses to the reviewers below, where we address these issues in detail.

EIC's comments:
1. This manuscript would fit SRL's electronic seismologist column well. See recent publications online at https://pubs.geoscienceworld.org/srl/search-results?page=1&f_TocHeadingTitle=ELECTRONIC+SEISMOLOGIST&fl_SiteID=129 If the authors agree, please resubmit the revised version to this column.
After discussion amongst the authors, we think the article will reach the widest audence via the regular section of SRL, and therefore, we prefer to have the article appear there.

2. You chose color online and b/w in print. Some figures may not show up properly in print. Please see guideline at https://www.seismosoc.org/publications/ssa-art-guidelines/ and change the color schemes accordingly.
We have modified some figures such that they do not rely strongly on color (primarily, Figures 1 and 4).

Reviewer #1: This manuscript presents an algorithm to located OBS with transponder data. This is a useful contribution because at present the available algorithms are either proprietary or ad hoc. I have some suggestions for how the manuscript and algorithm could be improved which I hope can be addressed before publication.

Title - The authors might want to indicate that the algorithm is for transponder locations since sometimes scientists use airgun shots.
We appreciate the suggestion. Instead of altering the title, we have modified the abstract to explicitly state that our code utilizes transponder ranging data.

line. 20-27. The water depth needs to be quoted since a lot of numbers quoted scale with depth.
This is an excellent point that is a shortcoming of our original analysis. Based on this comment and a comment from Reviewer #2, we have added to the electronic supplement the synthetic tests for survey geometry at 2000 m and 500 m water depth (Figures S7-S8). We find that uncertainties in r_xy, z, and Vp decrease with decreasing water depth, meaning that values quoted at 5000 m are upper bounds. The "optimal" radius also decreases with decreasing depth, as expected (from ~0.25 Nm at 500 m water depth to ~0.5 Nm at 2000 m depth). We now describe these tests and note the depth-dependency of our results in sections 3.3 and 4.

line 72. I would suggest citing Creager, Kenneth C., and LeRoy M. Dorman. "Location of instruments on the seafloor by joint adjustment of instrument and ship positions." Journal of Geophysical Research: Solid Earth 87.B10 (1982): 8379-8388, who developed the original method to located OBSs with shot data using similar inverse techniques
We have added the relevant citation to both the introduction and algorithm sections.

Line 92-118. The shared code would be much more generally useful is there was an option to solve for a systematic offset between the ships GPS and the ships transponder. For many research ships this offset is known and the scientists would use it to apply a heading dependent correction to the navigation data before locating the OBS. However, if it is not known or a handheld transponder is being used, solving for it is important to improve location accuracy.

Based on this comment and a comment from Reviewer #2, we have included an option in the code to specify a known offset in shipboard GPS position relative to the transponder and apply a heading-dependent correction to the true transponder position. The ship heading direction is directly estimated by differencing the ship GPS coordinates at each successful transponder ping. The Algorithm section has been updated to reflect these changes.

We have included additional synthetic tests exploring the importance of this correction. We find that for quasi-circular geometries such as PACMAN, the horizontal accuracy is not affected by this offset because of symmetry (see Figures 3,S6), and instead error is mapped into the depth and water velocity. Thus, uniquely solving for the GPS-transponder offset when it is not known is difficult because the offset will trade off with depth and water velocity (for some survey patterns, see Figure S9), which already strongly trade off with one another. Therefore, we have chosen not to implement the ability to solve for the offset at this time.

Line 120-146. This error analysis is very sophisticated and probably totally unnecessary for most applications. The errors shown in the figures (e.g., Figure 2) look ellipsoidal so have the authors attempted to use the covariance matrix to get errors based on a travel time uncertainty calculated from the travel time misfits? How do these compare with the bootstrap method?

The algorithm provides empirical uncertainties using bootstrap and F-test grid search methods which are robust and quick to calculate. On a standard laptop (16 GB of RAM, 2.9 GHz processor), the bootstrap and F-test gridsearch each take ~1.7 s to run, and the entire code operates in under 27 s for a single station including ray-bending corrections, outputs to the screen, and saving high-resolution PDFs.

One issue with simply utilizing the formal model covariance matrix from the inversion is that it depends directly on an estimate of the data uncertainty, (uncorrelated or otherwise), which is not known and introduces an extra assumption, in addition to the assumption of Gaussian distributed errors. Clearly, this makes direct comparison with the bootstrap difficult. The bootstrap and F-test methods do not require an explicit assumption about data uncertainties. Furthermore, this is an iterative method and therefore the iterated covariance matrix applies to the model perturbations and not necessarily to the model parameters themselves. For the above reasons, we believe that our approach for estimating confidence is superior and warranted.

Line 199. How is the reader meant to know what the PACMAN configuration is here and in the abstract. Refer to a drawing.

Figure 1 has been modified to include the PACMAN geometry. We have also modified Figure 2 to show the PACMAN survey pattern.

Line 169-217. One assumption of the method is that the rays are straight. What is the effect of refraction? That is compute synthetic data with refracted rays and relocate with straight rays and see what the errors are. My guess is that this leads to depth biases.

We thank the reviewer for pushing us to take ray bending into account. We have added this functionality to the code and included figures in the supplement showing the two-way travel-time difference between straight and refracted rays for two different sound speed profiles

(Figures S1-S2). We find that the corrections can be important (> 1 ms) for some sound speed profiles (particularly those containing a strong velocity change at the thermocline) at shallow water depths (<1000 m) and long offset. Therefore, we have implemented in the code the ability to correct for ray bending using a geographically suitable 1D sound-speed profile from the 2009 World Ocean Atlas database. The profile is automatically selected using the ship GPS coordinates and survey timestamp (see Algorithm section for details).

The ray bending correction is insignificant for the Young Pacific ORCA deployment (water depth ~4800 m). When applying the correction, the horizontal locations, z, and Vp change by ~0.1 m, ~0.4 m, and ~0.2 m/s, respectively.

Line 212-217. My standard practice has always been to get horizontal locations from the survey and depths from the multibeam because of the tradeoff between velocity and z

The user has the ability to do this if they wish by applying strong damping to perturbations in z away from the starting (multibeam) value.

Line 242-244. Is this an expected effect of a stratified ocean and of rays refracting? This leads to errors in the depth derivatives used in the inversion.

We have shown that the depth discrepancy at Pacific ORCA is not due to rays refracting (see above). After some synthetic testing (Figure 3,S5), we believe that the depth discrepancy is possibly due to an unaccounted for shipboard GPS-transponder offset. For the PACMAN survey, depending on the GPS-transponder offset geometry, this offset can yield a nearly constant travel-time bias which maps into a bias in z and Vp. This is now discussed in section 3.2 of the main text.

Line 265. Please provide more details on how ellipsoidal shape of earth is accounted for. Do the authors means selecting the correct scaling factors to convert degrees of latitude and longitude to meters?

We use the WGS84 reference ellipsoid to convert from geodetic coordinates (latitude, longitude, h) to local ENU (x, y, z) coordinates using standard coordinate transformations from *Hoffmann-Wellenhof et al.* (2001). We have added these details to the *Algorithm* section of the paper.

Line 306-312. I do not follow. The earth's ellipticity is only ~0.3% so there is no way I can see how one can accumulate horizontal errors of 10 m in 1 NM (0.5%) even with the most silly assumptions.

We have included in the supplement two figures which demonstrate the importance of the ellipsoid correction, both are included below. First, we perform a synthetic test showing the perturbation to the ship location as it traverses around the instrument, as well as the associated travel time perturbations when failing to account for the ellipsoid. The second plot is a demonstration using real data from the Pacific ORCA deployment.

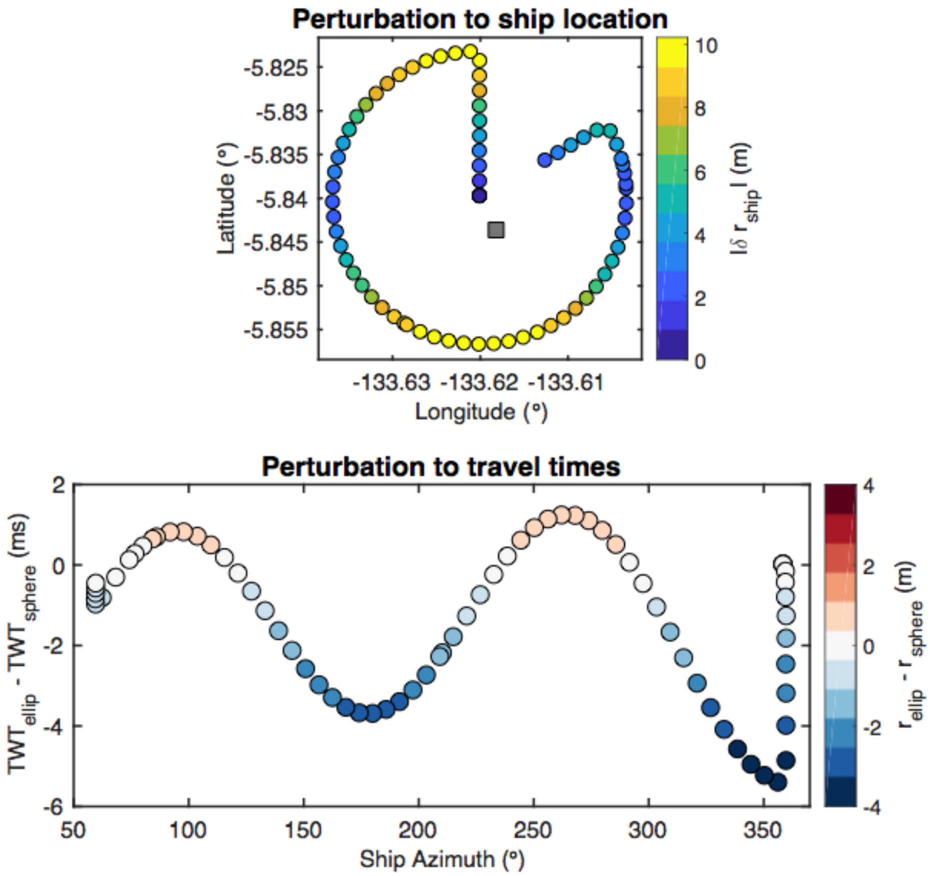

Figure S3: The synthetic test is shown above for a 1 Nm PACMAN survey. The top panel shows the survey pattern and is colored by the relative displacement of the ship position in (x,y) space due to failure to account for the ellipsoid. The lower plot shows the calculated perturbation in two-way travel time due to the difference in apparent ship position. We find that for a typical survey pattern, the ship is displaced by up to 10 meters at the north and south extremes and results in peak-to-peak TWT perturbations of ~5.5 ms.

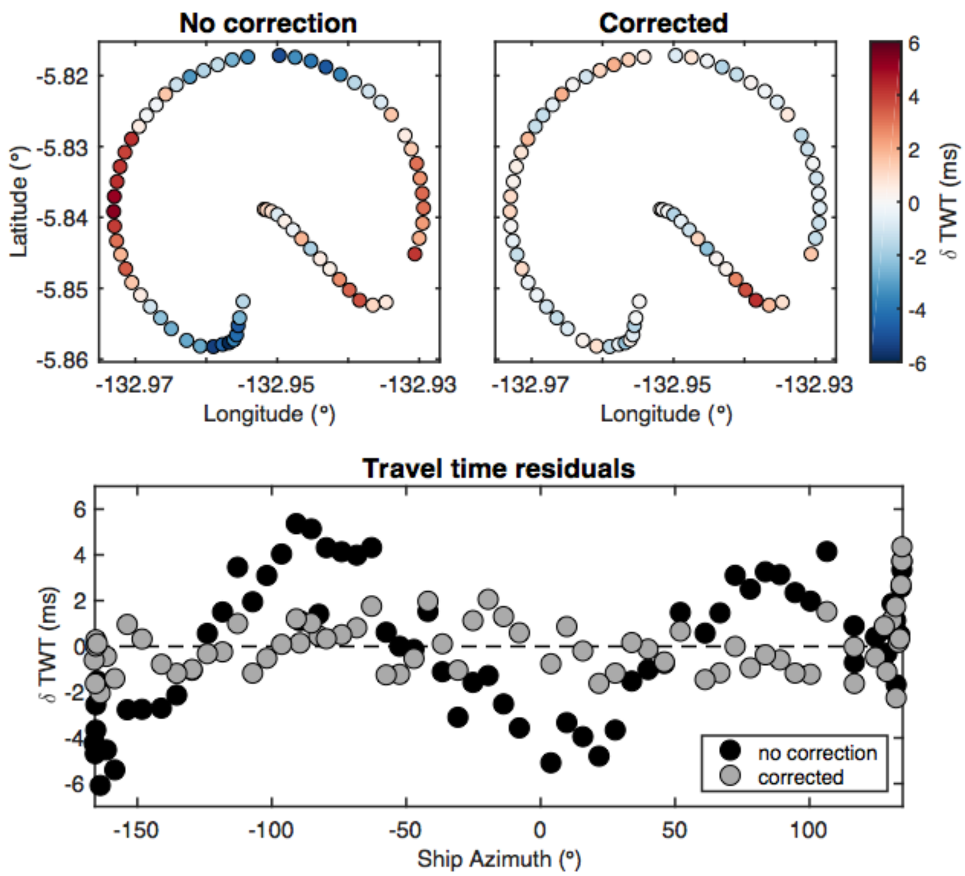

Figure S4: Above is a comparison of the travel time residuals for an inversion with (corrected) and without (no corrections) accounting for the ellipsoid for station CC06 of Pacific ORCA. We see that failing to account for the Earth's ellipsoid results in a systematic sinusoidal pattern in the residuals, similar to the synthetic case, with a peak-to-peak of 8-10 ms.

Line 332-333. Transponder times are good to microseconds so why model uncertainties that are 1000x this? This seems totally unnecessary.
We were unaware of the small magnitude of turn-around-time uncertainties and have removed these TAT perturbations from the synthetic tests.

Line 341-350. The water depth needs to be given for this to be meaningful since the survey size scales with the water depth.
See above

Line 355-359. The designation of 0.75 Nm being ideal is arbitrary because the tradeoff parameter assume arbitrarily that one wants to minimize the product of the survey time and horizontal misfit. Figure 8a-b provides a mechanism to select the quickest survey to gain the desired location accuracy which may be dependent on the experiment. Beyond 0.75-1 Nm the location errors a small but still decreasing as is clear from Figure 8a-b.

Although it is true that the desired location accuracy may depend on the experiment, we believe the tradeoff parameter as we have defined it (the product of total survey length and horizontal misfit) provides a useful quantification of survey performance. The minimum in this parameter represents an optimal scenario in most cases. Of course, it is ultimately up to the practitioner to decide which survey to choose and, as the reviewer notes, we already provide all the information for an informed choice of survey via figures 8a-b.

Line 371-372. PACMAN is the best geometry of those tested. However, some ship Captains do not like steaming in circles and it looks the diamond works just as well

In the results section, we have emphasized that the PACMAN and Diamond surveys perform nearly equally well for our synthetic tests. However, we still prefer the PACMAN geometry since its quasi-circular pattern requires a smaller Doppler correction (i.e., the ship remains at a nearly constant radius from the instrument for most of the survey). In the discussion section, we offer the Diamond survey as a viable alternative to PACMAN in the case where the captain wishes to steam only along straight lines.

Line 385-399. Again, optimal radius is meaningless without water depth. The discussion does not recognize the choice of desired location survey will depend on the required location accuracy which is potentially experiment dependent and for many experiments may be >10 m.

See above

Line 400-406. So two perpendicular lines might work particularly well which raises the question why the authors have not tested the "double packman or hourglass" - ¼ circle, follow diameter over instrument, ¼ circle, follow diameter back to starting position.

The cross pattern does contain two perpendicular lines. In the discussion, we have mentioned this as a simple alternative to the line pattern. We have also added the suggested “hourglass” pattern to our synthetic tests. Overall, it performs similarly to the cross patterns, with slightly longer survey distance (Figure 4).

Line 407-418. It is interesting that the OBS motions map the mesoscale circulation but describing this as novel measurement technique is a bit farfetched in that is hard to image this being an adopted scientific technique. I think most the circulation associated with eddies is shallow so it could be mapped instantaneously with the ship's ADCP.

We agree with the reviewer – this is not a technique that should be adopted by the community for the sole purpose of measuring water column properties. There are better techniques for measuring the details of shallow ocean flow structure, especially within the upper-most few meters, such as ADCP. However, the instrument drift data are a byproduct of the inversion and do provide a novel *observation* of depth-integrated flow throughout the water column that is potentially useful for the physical oceanography community, and we believe it is important to point that out.

Line 417 chance units from "m" to "m/s" twice
Thanks – the text has been amended.

Line 419 But the failure to include the doppler term leads to bigger horizontal errors.
This is true, and the discussion has been amended to state that. Doppler corrections do decrease the horizontal errors by ~2 m (~40% reduction) in the synthetic test, and therefore, the Doppler correction is left as an option in the code.

Line 424-426. I do not quite understand this. Normally a research ship will provide an accurate offset between its GPS and hull mounted transponder which will be merged with the heading to correct the navigation prior to locations. As noted above, I think this algorithm should calculate this offset for instances when the correction is wrong, it is not available or a hand held transponder is lowered over the side.
Addressed above.

Reviewer #2:
MAJOR ISSUES

There is no comparison with a non-constant velocity model, even though the ocean velocity profile is known to be non-constant. This throws in doubt all conclusions about the uncertainties. They should make a test model with a strong but realistic velocity model and see how closely their constant velocity inversion can fit (and perhaps investigate how the constant velocity best relates to the profile velocities. [A future iteration of this code could allow a non-constant velocity with depth (for example, a starting model multiplied by a variable so they still have only one velocity variable in the inversion?)]
Addressed in Reviewer #1 comment. We have included the ability to correct for ray refraction using a 1D depth-soundspeed profile.

The authors appear to have done all of their tests assuming an ~5000m water depth and their recommended survey sizes therefore apply only for depths near this. Many OBS deployments are made at 1000-2000 m water depth, and some shallower. This is never stated and no recommendations are made for surveys at other water depths.

Addressed in Reviewer #1 comment. We repeat the synthetic tests for geometry at 2000 m and 500 m depth and now discuss these results and the depth-dependence of survey accuracy in Sections 3.3 and 4?

SMALLER, BUT STILL IMPORTANT:

What would be the effect of a "biased" survey? (lots of returns on one side, few on the other)? Should returns be weighted to reduce bias?
We find that even back-azimuthally biased surveys do a good job of locating the station, as long as the measurements are made on approximately three points well-spaced around the station. We disagree with the idea of up-weighting regions of sparse returns, as if travel time errors are gaussian distributed, these errors will cancel out (albeit leading to greater RMS) at back-azimuths with lots of data, whereas up-weighting single returns (which themselves have error) might actually make the location worse by exaggerating errors at azimuths where no other data are available to constrain the inversion.

What is the time-accuracy tradeoff? (if you have only X hours to do a survey, what survey pattern should you use)? See also line 354 comment below.
We prefer to quote the distance-accuracy tradeoff, as different ships and sea conditions permit travel at different speeds. However, based on these comments, we have modified Figure 4a-c to plot *total survey distance*, instead of radius. It is then straightforward for the reader to divide the distances quoted by known/projected ship velocity and make a decision about optimal survey pattern.

I think that Section 3.4 "Exploration of survey pattern geometries" should be put before the current section 3.2, so that synthetic tests are together. Moreover, this would allow the authors to present the different survey geometries earlier in the article, avoiding some hunting around by the reader
We agree with the reviewer’s suggestion and have rearranged the results section, moving the application to the Young Pacific ORCA deployment to the end of the section.

It would be useful to also evaluate surveys using 3-4 cardinal points plus an overhead point, as this often has to be done when the ship does not allow reliable ranging when it is in motion. I'd be interested to see if it's any worse than the PACMAN survey (assuming you get a reliable fix at each of the points: one generally does multiple interrogations at each point), and how 3 points + overhead compares to 4 points + overhead.

Following this suggestion, we added a synthetic test for 4 cardinal points plus overhead to Figure 4. It does a relatively poor job at recovering horizontal positions (~20 m uncertainty for 5000 m water depth) but estimates depth and water velocity comparable to the others.

MINOR COMMENTS/CORRECTIONS:

L114: The values chosen for epsilon and γV_p should be explained: is epsilon simply chosen to be small? If it doesn't work, should another value be tried? Is the value of γV_p (50x larger than epsilon) explicitly linked to epsilon? Or just the result of trial and error? Similarly, e should be a variable parameter with a recommended value and a recommendation on how to change it if the inversion diverges or asymptotes before reaching this value (maybe include a TABLE of such values, with names and explanations?)
Both epsilon and gamma_vp are chosen by trial and error, and their default values have been tested on many different survey geometries and should work in most cases – we have amended the text to make this clearer. The value e is calculated and cannot be a variable. The inversion terminates once the RMS of e is reduced by less than 0.1 ms (this is a variable set by the user) from the previous iteration. This means the inversion terminates when the solution asymptotes or if it begins to diverge.

L160-166: Maybe too much technical detail here (if it's in Menke, just reference the article/page)
This equation is not in Menke. He only defines the covariance matrix, which is not unitless.

188: is .02 m really within the study precision?
We have amended the text to report one decimal place.

206: Precision seems too high (again).
See above.

199: The geometry of the PACMAN configuration should be explained (reference to section 3.4 is too general). My recommendation for reordering sections could help resolve this problem.
Similar to Reviewer #1 comment. We have also modified Figure 1 to contain the PACMAN geometry.

242-244: Errors suggest that the constant velocity model is causing problems. The authors should try a using the ship's XBT model (and not inverting for velocity): do they get a better result? My first impression is that the sound speed should be underestimated rather than overestimated by the constant velocity assumption, because the straight ray paths corresponding to a constant velocity model are always shorter than the true (curved) rays. The reason for overestimating the velocity should be tested, quantified and stated.

Similar to Reviewer #1 comment. We have tested the effects of ray bending on travel times. For stations at depths similar to the Young Pacific ORCA study region (~4700 m), we find that the two-way travel time differences between bent and straight rays is < 0.1 ms for offsets less than or equal to 1 Nm. These errors are more than an order of magnitude smaller than the RMS errors, suggesting that they are in fact negligible. The details of the shallowest few hundred meters (i.e., the depths constrained by the XBT) does not change this conclusion for deep stations.

We have added plots to the supplement demonstrating the differences between straight and curved rays (Figures S1-S2). We have also added an option to correct for ray bending in the code by applying a travel-time correction to the data using ray shooting through a velocity profile from the World Ocean Atlas database appropriate for the region (and there are instructions in the code for a user to easily swap in their own velocity model, for example using XBT data).

265 and 302: I did not see where you show how your code accounts for the Earth's ellipsoidal shape. This should be presented fairly early on in the description of the algorithm, along with a statement of the input parameters. It is not clear what the input to the code is: (time, lat, lon, traveltime, [cog],[sog]?). Time is at send? Receive? Is information input on the position of the pinger w.r.t. the GPS antenna? This should be fairly easy
Similar to Reviewer #1 comment. We have added additional details to the first paragraph of the algorithm section, including the ellipsoid model used (WGS84) and a reference containing the coordinate transformation equations. We also provide new supplementary figures demonstrating the impact of the ellipsoid correction (Figures S3-S4). We have also added the ability to account for a known GPS-transponder offset. We have written a User Manual that summarizes the inputs, options, and outputs of the code, a preliminary version of which is now submitted as a supplemental file for the reviewer/editor's consideration.

321: This part is not quantitative enough: how much does this improve w.r.t. overall location error, and how do you determine/reconstruct the ships' radial velocity (would a ship's measure of SOG/COG be better than an interpolation between pings?). Did you test the effect of having the antenna not collocated with the pinger (could make a plot of the error as a function of the x,y offset)
We assume that the user does not have access to the ship's SOG/COG estimates, as these are not collected by default during acoustic surveys, and making collection of these a requirement to use the code would lead to much less uptake among the community. We agree that these might be better than ping interpolation (although as the reviewer probably knows, these measurements significantly fluctuate in real time, requiring even more onerous processing to get average values at the time of the survey), but incorporation of this data is beyond the purview of this code. Accounting for non-co-located GPS and transponder is also a point raised by the other reviewer, and has been addressed above.

354: I think survey TIME times misfit would be a more valid trade-off parameter than survey RADIUS times horizontal misfit.
Survey time depends also on ship velocity, which depends on sea conditions, ship type, mates preference, etc. See response to comment above - we now specify total survey length, which can be easily converted to time by dividing by average ship velocity.

377-391: The first two paragraphs of the discussion seem more like conclusions than discussion.
We are not exactly sure to what the reviewer is referring here, but prefer to keep this text as it provides context framing the discussion to follow.

418: The decadal average is for the region? The globe?
The decadal average for the region. Text amended.

419: "only slightly" -> "XX%"
The Doppler corrections improve RMS travel-time misfit by only ~0.3 ms (~7% reduction). This has been added to the text.

432: Last sentence is too qualitative.
We removed this part of the discussion and wrote a paragraph on the importance of the GPS-transponder correction.

Figure 8: This important plot could be made clearer. Symbols could be improved so that they don't rely so much on color: give PACMAN a real PACMAN shape and make the symbol size correspond to the survey size. Crosses could be '+' and lines 'x' (or '-'). Tri-center could have a dot in the middle (as could PACMAN if you don't find a PACMAN symbol). Does the tri_edge really only cover one side of the instrument (plus cross over), or does it make an equilateral triangle around the instrument and not cross over? Also, instead of having a huge legend, you could just provide an x-axis with the survey names, and a summary box of the survey shapes (PACMAN, circle, cross2, cross, diamond, line2,line,tri_center, tri_edge. Or you could make the x axis the radius/length and plot the different symbols overlapping. You could also make a second set of plots with estimated survey time (or survey path distance) on the x-axis: might be the best way to compare survey methods/efforts.
(Now Figure 4) Based on this comment, we have updated the figure to show total survey length on the x axis as this can be easily converted into survey time. We have replaced the array of colored symbols and now only color symbols which change in survey size (blue); otherwise, they appear black and gray. We have increased the size of the survey geometry legend for clarity and include Cardinal and Hourglass patterns.