Revision Notes to Reviewer's Comments

We have completed a major revision of the earlier version of the manuscript by taking into consideration of the response of each of the reviewers. We hope this revised manuscript addresses all the reviewer's comments.

Reviewer's comments and suggestions greatly improved the structure, content, and the quality of this manuscript. We thank the reviewers for their valuable time spend in careful reading and constructive comments and suggestions. Their critiques gave us an opportunity to approach this research problem from various angles and expanded our understanding about the Northwest Pacific Ocean circulation.

The following revision notes explain details of what we have done to address specific comments from the each of the reviewers. *The reviewer's comments are in red text in italics*. Authors response is in black text. *Changes in the manuscript is in blue text in italics*.

Reply to Comments of Reviewer # 1

1 Major comments

A detailed description of the acoustic travel time operators is missing, providing an impression that assimilation is weakly sensitive to constructing the ray path approximations using a relatively coarse (20 km) model grid at 300-600 km ranges and identifying the respective travel times from real data. The particular missing details are listed below: 1. What was the reason for computing the sound speed (or slowness?) in between spatial interpolations of the qS data from the coarse model grid to the fine 2d grid used for computing the ray bundles (lines 242-246)? In my experience, first mapping the background temperature and salinity on the fine along-the-range grid and then computing slowness and its derivatives provides best results, especially in the high-latitude environments.

The first paragraph in Section 3c, "Comparison of Travel-time Estimates with Observations," has been revised to provide more details on the interpolation and the ray tracing program used. An error in the original version has also been corrected. The paragraph has been replaced by:

The 50 vertical levels and $1/6^\circ \times 1/6^\circ$ *(*$\sim$ 18 km*) spherical grid used in the regional implementation of the MITgcm provide adequate resolution in the vertical and horizontal to allow using relatively straightforward interpolation schemes to obtain the sound-speed fields on the vertical sections between the acoustic transceivers needed to construct acoustic rays and compute travel times. Model temperatures and salinities were first linearly interpolated in the horizontal for each vertical level to construct profiles at 21 evenly-spaced locations*

(including the end points) on the geodesics connecting the acoustic transceivers (Tests using 41 evenly-spaced locations did not affect the results). Least-squares cubic spline interpolation was then used in the vertical to construct smooth profiles of temperature and salinity at 1-m intervals. Finally, sound-speed profiles at 1-m intervals were computed using the Del Grosso equation [Del Grosso(1974)]. The MATLAB ray tracing program ZRAY (M. Dzieciuch, personal communication) used these sound-speed profiles to compute acoustic ray paths and travel times. The program ZRAY linearly interpolates in the horizontal between the sound-speed profiles constructed from the model. It uses cubic splines to interpolate in the vertical to a finer grid as needed.

Accurate mapping of travel-time perturbations on the perturbations of model prognostic fields (theta,S) requires computation of the sound slowness derivatives wrt to temperature and salinity along the ray path. What sound speed/slowness and theta/T conversion equations were used for this purpose?

As noted in the response to the previous comment, the following sentence has been added to the first paragraph in Section 3c:

Finally, sound-speed profiles at 1-m intervals were computed using the Del Grosso equation [Del Grosso(1974)].

What ray tracing algorithm was actually used (line 247)?

As noted above, the following sentence has been added to the first paragraph in Section 3c:

The MATLAB ray tracing program ZRAY (M. Dzieciuch, personal communication) used these sound-speed profiles to compute acoustic ray paths and travel times.

Tidal filtering in your case is somewhat tricky, because travel times sampling (8 observations per day) was interrupted by 24-hour(?) intervals of silence that are longer than the periods of the key tidal constituents. Section 4a should be expanded to include the filtering details. For example, did you use prior spatial distribution of the phases and amplitudes of the tidal harmonics to constrain the fit? What was the "significantly longer" period? How did you actually filter the "extreme outliers"?

Section 4a, "Low-frequency Travel-time Variability," has been significantly expanded to more fully describe the procedure used to estimate and remove tidal variability and construct filtered time series of travel times. Several references have also been added.

2 Minor comments

Abstract: mentioning "reciprocal travel times" appears a bit misleading because one may think that the model was constrained by observations of both travel times and differential

travel times, which directly constrain the horizontal velocity fields along the rays. However, I could not find any mention of reciprocal travel time operators and the respective results in the manuscript.

The word "reciprocal" has been removed from the Abstract and Introduction. In addition, the following paragraph was added to Section 2, "2010–2011 NPAL Philippine Sea Experiment."

Acoustic signals traveled in opposite directions between the transceivers during each transmission sequence. (There were only one-way transmissions from the sources on T1–T5 to the receiver on T6 after the source on T6 failed.) Sum and difference travel times have commonly been computed from reciprocal travel times in previous tomographic experiments to separate the effects of sound-speed perturbations from those due to currents [Munk et al.(1995)]. Travel-time signals due to sound-speed perturbations cancel in the difference travel times if the oppositely traveling signals are sufficiently close in time and the ray paths sufficiently close in space for the sound-speed perturbations encountered by the oppositely traveling signals to be identical. Sum and difference travel times were not computed here, however. Rather, the reciprocal travel times were treated independently in the assimilation without assuming any correlation. This was done largely because of the 9-min intervals between the source transmissions, which gave a maximum difference between the times of the oppositely traveling signals of up to 45 minutes (Table 1). In addition, the centers of the 4-element receiving arrays were located 22.4 m above the acoustic centers of the sources, giving ray paths that differed for the oppositely traveling signals.

Line 83: Worth to mention that strong mesoscale variability in this region requires quadratic corrections to travel times [1], and 4dVar assimilation of the acoustic observations (including reciprocal times) were performed in [2, 3].

In response to the first comment, the final sentences in the third paragraph in Section 4b, "State Estimates," have been modified to read:

To first order, travel times are not sensitive to errors in the ray paths [Munk et al.(1995)]. Nonetheless, the standard approximation that the travel-time perturbations are linearly related to the sound-speed perturbations along the unperturbed ray paths [Munk et al.(1995)] can become invalid in ocean regions containing strong spatial variability if the ray paths computed for the assumed background sound-speed field differ substantially from those in the actual field [Yaremchuk and Nechaev(2001)]. Here, however, the non-acoustic state estimates used to initialize the NSE simulations and calculate the ray paths proved to be sufficiently close to the actual fields that the linear approximation was adequate. In any event, the assumption of fixed ray paths is not essential. Tests in which the ray paths were updated by re-tracing rays as the state estimate was updated at weekly intervals did not change the acoustic state estimates significantly. In addition, the assumption was checked by tracing rays and computing travel times in the final optimized state estimates. For the ray paths in this experiment, the rays were not strongly variable unless the angles with the horizontal were

less than about 8°. Travel times from these low-angle rays were not used in the assimilation.

In response to the second comment, the final sentences in the second paragraph in the Introduction have been modified to read:

In the 2010–2011 NPAL Philippine Sea Experiment, tomographic travel-time data have finally been used directly to constrain a high-resolution, mesoscale eddy permitting, regional implementation of a primitive equation ocean model, the Massachusetts Institute of Technology general circulation model (MITgcm, [Marshall et al.(1997)]), using the Estimating the Circulation and Climate of the Ocean (ECCO, [Stammer et al.(2002)]) four-dimensional variational (4DVAR) assimilation system. Two of the previous studies [Lebedev et al.(2003), Yaremchuk et al.(2004)] also used 4DVAR assimilation systems, but for simplified quasi-geostrophic dynamics, as noted above.

Line 148: An example of the ray paths between a pair of transceivers would be illustrative: judging by the ray identifiers in Figure 2, the number of turning points of a ray is close to the number of model grid steps along the ray path. A comment on this and its possible impact on the results is desirable. (e.g., discovered warm bias despite numerous in situ observations constraining NSE solution).

Section 2, "2010-2011 NPAL Philippine Sea Experiment," has been expanded to include the following paragraph and a new Figure 3 showing the ray weighting functions for the ray paths for which travel times are shown in Figure 2 has been added. Figure 3 showing the ray weighting functions is also provided with this reply.

The sampling properties of the ray paths are given by the ray weighting functions [Munk et al.(1995)]. To first order, the perturbation in travel time $\Delta\tau_n$ for each path n relative to the travel time in an assumed background sound-speed field C_0 is given by

$$\Delta\tau_n = -\int_{\Gamma_n} ds \frac{\Delta C}{C_0^2} \tag{1}$$

where Γ_n is the ray path in the background sound-speed field, s is arc length, and ΔC is the sound-speed perturbation. The ray weighting function ds/C_0^2 gives the weighting with which ΔC contributes to $\Delta\tau_n$. In general, the background sound-speed field and the sound-speed perturbations vary in both depth and range. The ray weighting functions for the paths for which travel times are shown in Figure 2 are given in Figure 3. Here the weights ds/C_0^2 have been summed in 10-m depth bins between each source and receiver to show the vertical sampling properties of the rays. The ray travel times are most sensitive to sound-speed perturbations at the upper and lower turning depths of the rays.

Line 210: Be more specific by changing the last sentence to "The error covariance matrices Q and R in eq. (1) are assumed to be diagonal".

We revised the sentence as follows:

The error covariance matrices $\mathbf{Q}(t)$ and $\mathbf{R}(t)$ in Equation 2 are assumed to be diagonal, i.e., uncorrelated.

Lines 228-229: change to "small-scale features in the sensitivity fields caused by the unstable modes of the adjoint propagator". Remove "Extending the duration of the linearity assumption" (sounds awkward).

We revised this sentence as follows:

In order to suppress the growth of nonlinear instabilities with longer integration times of the order of one or two months, diffusivity and viscosity coefficients in the adjoint of the background model were increased [Hoteit et al.(2005), Kohl et al.(2007)]. In addition to increased diffusivity and viscosity, the KPP mixing parameterization is disabled in the adjoint of the background model to minimize its contribution to the system nonlinearity, allowing longer integration times. A more detailed discussion about the choice of the diffusivity and viscosity coefficients in the adjoint model is provided in Appendix B in GG20.

Line 246: "to construct a dense set– how dense? Specify grid spacing in x-z. This is important because the model horizontal grid is rather coarse compared to the ray oscillation period (Fig. 2) - see my request to show a typical ray bundle.

As indicated above, the first paragraph in Section 3c, "Comparison of Travel-time Estimates with Observations," has been revised to provide more details on the interpolation. The relevant sentences are:

Model temperatures and salinities were first linearly interpolated in the horizontal for each vertical level to construct profiles at 21 evenly-spaced locations (including the end points) on the geodesics connecting the acoustic transceivers (Tests using 41 evenly-spaced locations did not affect the results). Least-squares cubic spline interpolation was then used in the vertical to construct smooth profiles of temperature and salinity at 1-m intervals. Finally, sound-speed profiles at 1-m intervals were computed using the Del Grosso equation [Del Grosso(1974)].

Lines 250-251: "state estimates were first averaged in time to provide weekly– what filter was used? Specify.

We revised this sentence as follows:

In order to compare the measured travel times with travel times computed for the non-acoustic state estimates constructed by GG20, the state estimates were first averaged in time to provide weekly averaged three-dimensional fields of temperature and salinity with time stamps centered on the middle of each week, reducing the number of model realizations and effort required to calculate acoustic ray paths and compute travel times using the model outputs. The choice of weekly averages is somewhat arbitrary but was made because one

week is short compared to the mesoscale eddy timescale of ~ 100 *days [Qiu and Chen(2010)] while not so short as to be redundant.*

Lines 262-271: This paragraph describes the "key step– (ray identification). Although you say it is "straightforward", I would recommend to extend the description, because the technical details may be of interest. For example, what were the (approximate) distributions of the launch and reception ray angles? (illustration of the aforementioned typical ray bundle would be helpful). What was the accuracy in detecting these angles at the transceiver locations? Provide more details about the "bookkeeping" mentioned in the last sentence.

The last paragraph in Section 3c, "Comparison of Travel-time Estimates with Observations," was replaced with:

A key step in the construction of Figure 2 was the identification of the tracked travel-time series with specific acoustic ray paths in the model. The identification is done by comparing the measured and predicted acoustic arrival patterns in travel-time and vertical-arrival angle space. The identification is normally unambiguous, in part because many of the arrivals are well separated in time and in part because the four-element vertical receiving arrays allow determination of the arrival angles with a precision of $\sim 1^\circ$*, depending on the SNR [Munk et al.(1995)]. The arrivals are traditionally labeled with a ray identifier* $\pm n$*, where* n *is the total number of ray turning points and* $+(-)$ *refers to rays that start upward (downward) at the source [Munk et al.(1995)]. In the Philippine Sea, however, the effects of strong mesoscale variability combined with the relatively low curvature of the sound-speed profile near the sound-channel axis leads in some cases to multiple near-axial rays that have the same ray identifier but different ray paths. The non-uniqueness of traditional ray identifiers is not a fundamental problem, but requires the use of ray identifiers that include additional information, e.g., the vertical angle of the ray at the source, to unambiguously label the arrivals.*

Lines 304-305: "Considerable care–. Specify what you did, and what were the problems requiring this care.

Thank you for this comment. We overlooked this detail about the travel time cost function. We included a paragraph briefly describing how the model equivalent travel times and cost function are computed.

There are 30 acoustic paths between the six acoustic transceivers, when reciprocal transmissions are included. For each acoustic path, there are multiple ray paths, giving over 341 observed travel times that are included in the cost function. Considerable care was executed to generate sound speeds, construct ray paths, and compute accurate travel times using the model solutions. For the computation of model equivalent travel times, pre-specified ray paths were considered. The steps involved in the travel time cost computation are briefly described below. For each ray path of 341 observed travel times, a ray length of 7000 points were considered and the travel times from the model fields were computed for each ray path

by integrating over the ray length, for each day of observation. Each ray path include details of incremental arc length (ds), longitude, latitude, azimuth angle, and water depth. As a first step, three-dimensional fields of model temperature, salinity, and pressure are used to compute the sound speed (C) following [Del Grosso(1974)] formulation. The sound speed routines are calibrated for a constant density of rho $= 1033kg\ m^{-3}$ *when computing the pressure term. The model equivalent travel times are computed for each day of travel time observation and for each ray path by integrating the term (ds/C) along the ray length. For that, daily averaged fields of sound speed, zonal and meridional velocities were linearly interpolated, first horizontally and then vertically, for each point along the ray length. The azimuth angle information is used to compute the zonal and meridional velocity contributions to the sound speed. The travel time differences between model and observations for each day and for each ray path were used to compute the acoustic travel time cost function, which is weighted sum of squared model–data differences. An assumed travel time uncertainty of 20 ms is used in the travel time cost function.*

Line 321: "travel times were not sensitive to errors in ray paths". Probably worth mentioning the Kuroshio Extension [1] where the second order approximation is important.

As noted in response to the comment on Line 83, the final sentences in the second paragraph in Section 4b, "State Estimates," have been modified to include a reference to [*Yaremchuk and Nechaev(2001)*].

Line 325: "unless the [arrival?] angles were less than 8o" Does that mean that you did not use these "low-grazing" rays? Again, the illustration of a ray bundle would be helpful.

As noted in response to the comment on Line 83, the final sentences in the third paragraph in Section 4b, "State Estimates," have been modified to include:

For the ray paths in this experiment, the rays were not strongly variable unless the angles with the horizontal were less than about 8°. Travel times from these low-angle rays were not used in the assimilation.

Line 330: 0.2s translates to (an approximate) potential temperature bias of 0.2o - 0.3oC, which appears considerable, given that the assumed rms error R1/2 in temperature ranges within 0.1-0.2oC at the depth range 500-1000m (Fig. 8). Does it mean that NSE solution (and HYCOM solution used as the first guess for NSE) were not properly constrained by the TS-profiles at these depths? Please comment.

Thank you for your comment. We assume you are referring to Figure 7 showing the comparison of NSE, ASE and HYCOM solutions to DVLA temperature data at single location. The NSE solution assimilated satellite derived SSH and SST and Argo temperature and salinity observations. But, there were only 22 Argo profiles located within a $2° \times 2°$ region centered at the DVLA over the entire NPAL experiment period, which might be not sufficient to constrain the NSE solutions. We included this detail in the manuscript.

The discrepancy could also be due to the small number of observations in the vicinity of the DVLA mooring that are available to constrain the model.

This misfit can occur because SSH in this region is relatively insensitive to the ocean interior and the Argo data were relatively sparse, with only 22 Argo profiles over the yearlong experiment within a $2° \times 2°$ *region centered on the DVLA.*

Line 371: I would remove "due to non-linearity" since the necessity to employ diffusive damping is already caused by the non-linearity of the background trajectory.

We revised this sentence as follows:

For some of the assimilation periods, there is a rise in the standard deviation of differences toward the end of the assimilation window, likely due to loss of controllability by the initial condition controls caused by removal of small-scale structures from the gradients by using increased diffusion in the adjoint model.

Line 413: Could you comment why the adjustments were centered at the transceiver locations (not more or less evenly distributed along the ray paths)? For example, the zonal section also passes not far from T6 (showing these two sections in Fig. 1 would be helpful), but there is no sign of T6 presence in Fig. 7.

We revised the paragraph discussing initial condition control adjustments for temperature and salinity. We also included a new figure showing spatial distribution of adjustments for initial condition temperature and salinity controls at selected depth levels (Figure 8). We also marked the transceiver locations in Figure 9 showing the zonal and meridional sections centered on DVLA location. Figure 8 shows the adjustments distributed along the ray paths depending on the ray depth. Since the density of the rays are more at the transceiver locations T1 through T5, the adjustments were centered on those locations. We included the following revised paragraph in the manuscript. Figure 8 is also provided with this reply.

The mean adjustments to the temperature and salinity initial condition controls from seven acoustic state estimations covering the entire NPAL experiment period are examined for selected depth levels in Figure 8 and for the zonal and meridional sections across the DVLA mooring ($21°$ *21.7418 N,* $126°$ *00.7867 E) in Figure 9. The adjustments are confined between 250 and 1750* m *with multiple maximum values centered on depths of about 500 m and 1400 m, for both temperature and salinity. This is expected as the ray path depths are mostly confined in the depth range of 250 to 2500 m for the most of the rays between the transceivers T2, T3, T4, and T5, whereas some ray paths between the transceivers T1, T2, T5, and T6 traces deeper depths in the range 2500 to 4000 m. The spatial adjustments are distributed along the ray paths and are confined to the geometry of the ocean acoustic tomographic array. The adjustments are more or less centered on the locations of acoustic transceivers T1 through T5, for both temperature and salinity,*

perhaps due to increased ray density at those locations (Figure 8). The temperature and salinity adjustments near the location of T6 transceiver are moderate when compared to other five transceivers, likely due to its premature failure. For temperature, the vertical adjustments between 250 and 600 m *are more or less uniformly distributed. The adjustments between 600 and 1750* m *are centered on the locations of acoustic transceivers T5 and T2 for the zonal section, and T1 and T4 for the meridional section, for both temperature and salinity. The adjustments tend to decrease the NSE temperatures and salinities, compensating for the subsurface warm bias in the NSE estimates between 500 and 1500* m *depth.*

Lines 424-424: remove "the first guess solutions of the acoustic state estimates, which are the– appears distractive (redundant). Same in line 798 (caption to Fig. 8).

We removed this redundancy. Thank you

Line 448: Again, could you comment on these large (0.5C) differences with regard to the constraints imposed by the TS-profiles in the NSE (and HYCOM) solutions, which was virtually unchanged after acoustic assimilation? Is it partly due to the extreme sparsity of the Argo data in the NSE, so respective contributions were too small compared to, say, SSH/SST contributions?

This discussion in the manuscript refers to forecast comparison with DVLA temperature data at single location. This differences can be due to two factors: 1) climatological forcing of forecasts, 2) few Argo profiles (22) within a $2^\circ \times 2^\circ$ region centered at the DVLA over the entire NPAL experiment period, which might be not sufficient to constrain the NSE solutions. The acoustic assimilation slightly improves the forecasts reducing the differences from $0.5^\circ C$ for F–NSE to $0.3^\circ C$ for F–ASE at about 600 m.

Line 450: changing "measurements" to "acoustic observations" will improve clarity.

We included this change. Thank you

You could consider abbreviating Gopalakrishnan et al (2020) to, say, G20: the manuscript cites this paper more than a dozen of times.

We replaced Gopalakrishnan et al (2020) to GG20. Thank you.

Fig. 3: I could not find any reference to the stars in the upper panels.

The magenta and blue stars marks the minimum and maximum differences. We included this detail in the Figure 4 caption. Thank you.

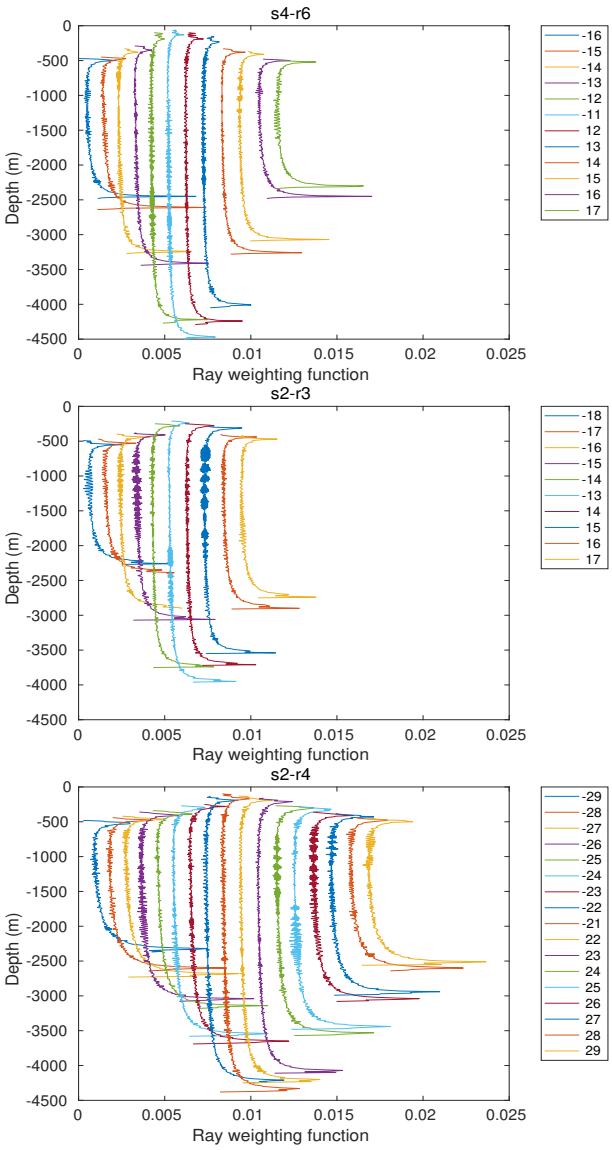

Figure 1: Ray weighting function binned every 10 m depth for different source-receiver pair ray paths shown in Figure 2. The weighting function for each ray path is offset by a value of 10^{-3} in the x-axis to make the rays distinguishable.

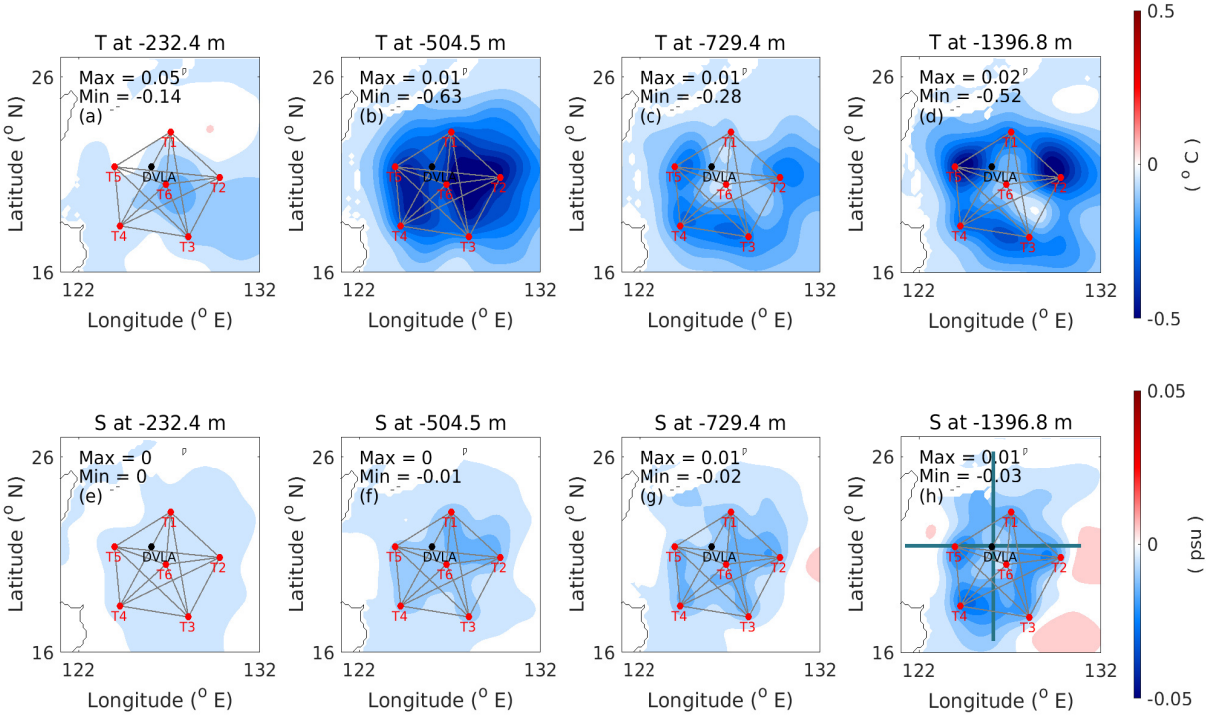

Figure 2: Mean corrections over seven state estimation experiments to the initial conditions (controls) of temperature (top panels) and salinity (bottom panels) for selected depth levels. The location of acoustic transceivers (T1 – T6) and DVLA are marked. The green lines centered on the DVLA location in panel (h) marks the zonal and meridional sections shown in the next Figure. The plots were zoomed to show the geometry of the tomography array.

References

[*Del Grosso(1974)*] Del Grosso, V. A., 1974: New equation for the speed of sound in natural waters (with comparisons to other equations). *The Journal of the Acoustical Society of America*, **56 (4)**, 1084–1091.

[*Gopalakrishnan et al.(2020)*] Gopalakrishnan, G., B. D. Cornuelle, M. Mazloff, P. Worcester, and M. Dzieciuch, 2020: State estimates and forecasts of the eddy field in the subtropical countercurrent in the northern philippine sea. *Journal of Atmospheric and Oceanic Technology (in review)*.

[*Hoteit et al.(2005)*] Hoteit, I., B. Cornuelle, A. Kohl, and D. Stammer, 2005: Treating strong adjoint sensitivities in tropical eddy-permitting variational data assimilation. *Quarterly Journal of the Royal Meteorological Society*, **131 (613)**, 3659–3682.

[*Kohl et al.(2007)*] Kohl, A., D. Stammer, and B. Cornuelle, 2007: Interannual to decadal changes in the ECCO global synthesis. *Journal of Physical Oceanography*, **37 (2)**, 313–337.

[*Colosi and Munk(2006)*] Colosi, J. A., and W. Munk, 2006: Tales of the venerable honolulu tide gauge. *Journal of physical oceanography*, **36 (6)**, 967–996.

[*Lebedev et al.(2003)*] Lebedev, K. V., M. Yaremchuk, H. Mitsudera, I. Nakano, and G. Yuan, 2003: Monitoring the Kuroshio Extension with dynamically constrained synthesis of the acoustic tomography, satellite altimeter and *in situ* data. *Journal of Oceanography*, **59 (6)**, 751–763, *10.1023/B:JOCE.0000009568.06949.c5*.

[*Munk et al.(1995)*] Munk, W. H., P. F. Worcester, and C. Wunsch, 1995: *Ocean Acoustic Tomography*. Cambridge University Press.

[*Pawlowicz et al.(2002)*] Pawlowicz, R., B. Beardsley, and S. Lentz, 2002: Classical tidal harmonic analysis including error estimates in matlab using t_tide. *Computers & Geosciences*, **28 (8)**, 929–937.

[*Qiu and Chen(2010)*] Qiu, B., and S. Chen, 2010: Interannual variability of the North Pacific Subtropical Countercurrent and its associated mesoscale eddy field. *Journal of Physical Oceanography*, **40 (1)**, 213–225.

[*Yaremchuk et al.(2004)*] Yaremchuk, M., K. Lebedev, and D. Nechaev, 2004: A four-dimensional inversion of the acoustic tomography, satellite altimetry and in situ data using quasigeostrophic constraints. *Inverse Problems in Science and Engineering*, **12 (4)**, 409–431, *10.1080/10682760310001633689*, `https://doi.org/10.1080/10682760310001633689`.

[*Yaremchuk and Nechaev(2001)*] Yaremchuk, M. I., and D. A. Nechaev, 2001: Simulations of quasigeostrophic currents derived from satellite altimetry and acoustic tomography of an open ocean region. *Journal of Atmospheric and Oceanic Technology*, **18 (11)**, 1894–1910, *10.1175/1520-0426(2001)018¡1894:soqcdf¿2.0.co;2*.

[*Stammer et al.(2002)*] Stammer, D., and Coauthors, 2002: Global ocean circulation during 1992–1997, estimated from ocean observations and a general circulation model. *Journal of Geophysical Research-Oceans*, **107 (C9)**, 3118.

[*Marshall et al.(1997)*] Marshall, J., A. Adcroft, C. Hill, L. Perelman, and C. Heisey, 1997: A finite-volume, incompressible Navier Stokes model for studies of the ocean on parallel computers. *J. Geophys. Res*, **102 (C3)**, 5753–5766.