

CTD corrections applied to delayed mode data

- deployment: dfo-bb046-20220707
- glider: dfo-bb046
- sensor: CTD_0329
- description: Calvert Island Line

1.0 Preamble

This document describes conductivity, temperature, and pressure data processing steps applied to delayed mode data collected using Sea-Bird Scientific Glider Payload Conductivity Temperature Depth (GPCTD) sensors mounted on C-PROOF Slocum and SeaExplorer autonomous ocean gliders. This sensor has a nominal sampling rate of 1 Hz and was designed specifically for Slocum gliders. This document covers the application of the sensor alignment correction and the thermal lag correction, as well as removal of questionable conductivity values and salinity profiles.

1.1 Set up the processing

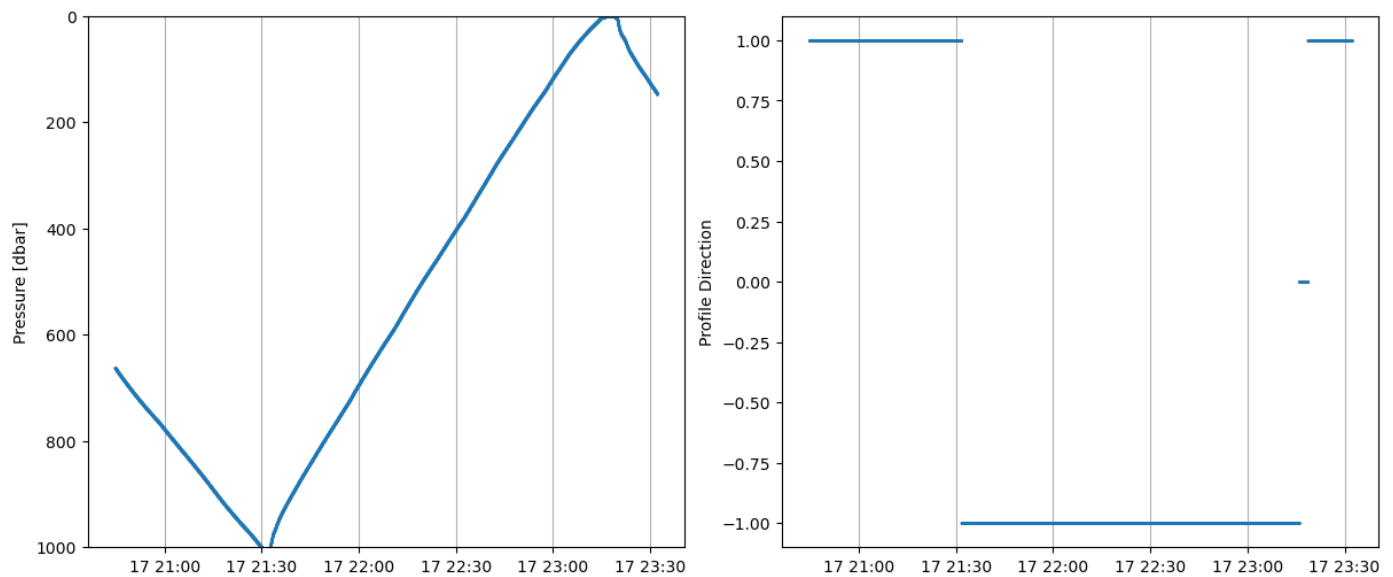
The processing steps below are applied to delayed mode data stored in a single netcdf timeseries file created using the Pyglider data processing package (<https://github.com/c-proof/pyglider>). The metadata and sensor calibration sheets are available via the deployment page on the C-PROOF website: <https://cproof.uvic.ca/gliderdata/deployments/dfo-bb046/dfo-bb046-20220707/>.

1.2 Profile Check

Check that upcasts and downcasts are being properly identified. **Negative values should be associated with upcasts**, and indeed we can see that that is the case here.

There are 1759051 data points in total.

Fig. 1: Identifying upcasts and downcasts. The left panel shows pressure vs. time and the right panel shows profile direction vs. time for a small subset of the time series.



2.0 Corrections applied to delayed mode data for this mission

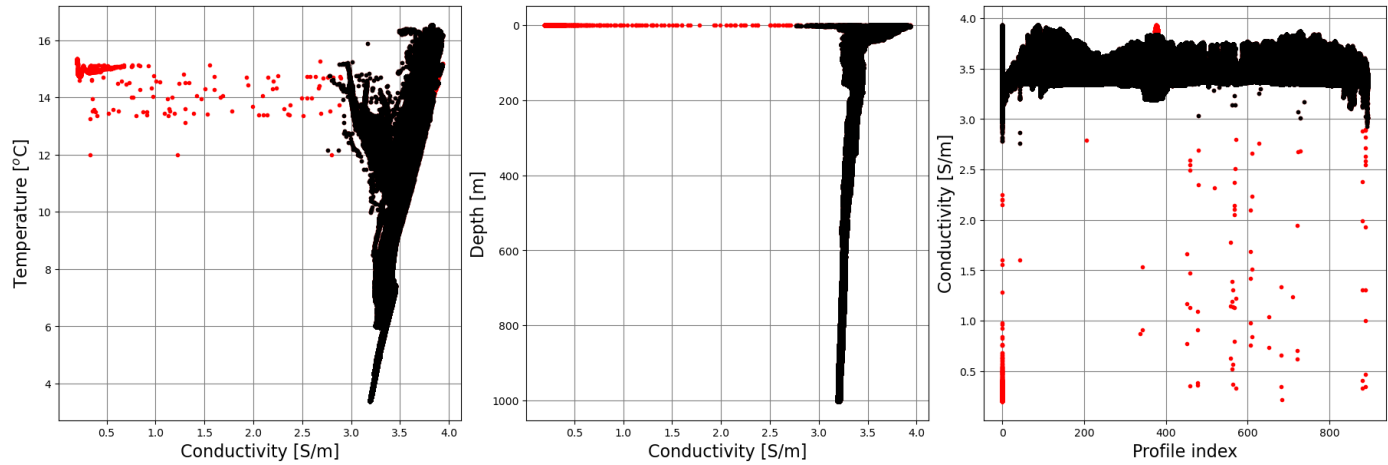
1. Identification of anomalous conductivity values
2. Sensor alignment correction
3. Identification of questionable salinity profiles
4. Thermal lag correction

2.1 Identification and removal of anomalous conductivity values

We identify and remove any conductivity values that are obviously unphysical, which is typically caused by air bubbles in the conductivity cell. We use a simple criterion applied to the raw conductivity data. The criterion temporarily flags any data points that are more than **5 standard deviations** away from the overall time series mean for a given depth, then recomputes the mean and standard deviation, excluding the temporarily flagged values. Conductivity values that still differ from the mean by more than **5 standard deviations** are flagged as 'bad' and set to NaN in the time series. If the difference between the 'bad' values and the mean is less than the accuracy of the sensor, which is 0.0003 S/m for the GPCTD, then those points are not excluded.

This criterion is applied to data binned by profile index, in increments of **100 profiles**. The use of profile index bins rather than depth or temperature bins is designed to allow for the removal of unphysical values in both **Hakai Pass** and crossing the continental shelf in **Queen Charlotte Sound**, two mission segments where the distributions of conductivity differ significantly.

Fig. 2: Temperature vs. conductivity (left), depth vs. conductivity (middle), and conductivity vs. profile index (right), with the red dots showing the unphysical values flagged as bad and removed.



Adjustments to this correction are based on examining the data and making a judgment call about which conductivity values are undeniably 'bad'. In this case, we want to exclude the **extremely low values occurring at the surface** (Fig. 2) consistent with air bubbles in the cell. Some unphysical values are missed by this correction, and may be caught during the removal of unphysical salinity profiles below.

NOTE that for this mission, the sampling rate of the GPCTD is one sample every 1s!

The mode of the sampling rate for the GPCTD is one sample every 1 seconds.

2.2 Sensor alignment correction

We now apply the sensor alignment correction. This correction is used to align the temperature and conductivity in time, relative to the pressure. This correction reduces the occurrence of salinity spikes near sharp gradients in T and S and ensures calculations are made using the same parcel of water for all variables. The misalignment between the sensors is caused by 1) the physical separation between sensors causing a transit time delay for water being pumped through the CTD and 2) different sensor response times.

GPCTD specs are as follows:

- Nominal sampling rate is 1 Hz (0.5 Hz for this mission)
- Pumped flow-rate through the sensors is 10 ml/s

Janzen and Creed (2011) provide constants used for the GPCTD sensor alignment correction "... based on the transit time between the T-C sensors, the temperature response time, and the estimated response time of the conductivity sensor in a 10 ml/s flow." The following constants are used for their Sea-Bird GPCTD alignment corrections:

- Temperature advanced by +0.5 s relative to pressure
- Conductivity advanced by +0.4 s relative to pressure

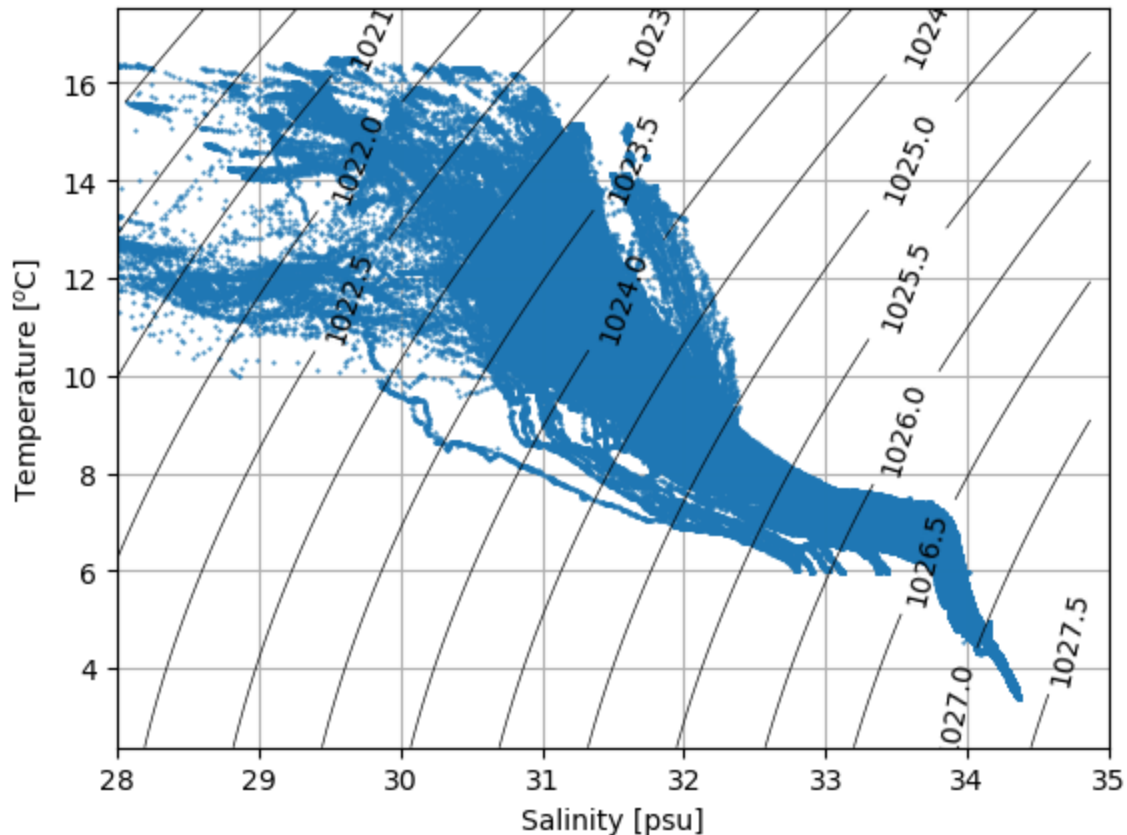
Our procedure for the alignment correction:

1. Directly estimate the sensor alignment correction constant for conductivity relative to temperature, τ_C
2. Confirm that the directly estimated value of τ_C is within 0.5s of the Janzen and Creed (2011) value of -0.1s (if the directly estimated value for a given sensor is consistently higher or lower than -0.1s, investigate further)
3. Apply the alignment correction using the alignment constants for conductivity and temperature provided in Janzen and Creed (2011)

We **apply the correction** using a linear time interpolation method, based on the one in the Sea-Bird CTD data processing manual and the Sea-Bird CTD-processing toolbox on Github (https://github.com/rejectedbanana/Sea-Bird-Toolbox/blob/master/CTD_PROCESSING/SBE_alignCTDW.m).

We determine the correction using only water below the highly variable near-surface layer, which is approximately identified using the T-S diagram below. We choose as an isopycnal cutoff a value above the main thermocline, but which excludes that highly variable surface layer:

Fig. 3: Temperature vs. salinity diagram. Black contours give density in kg/m^3 .



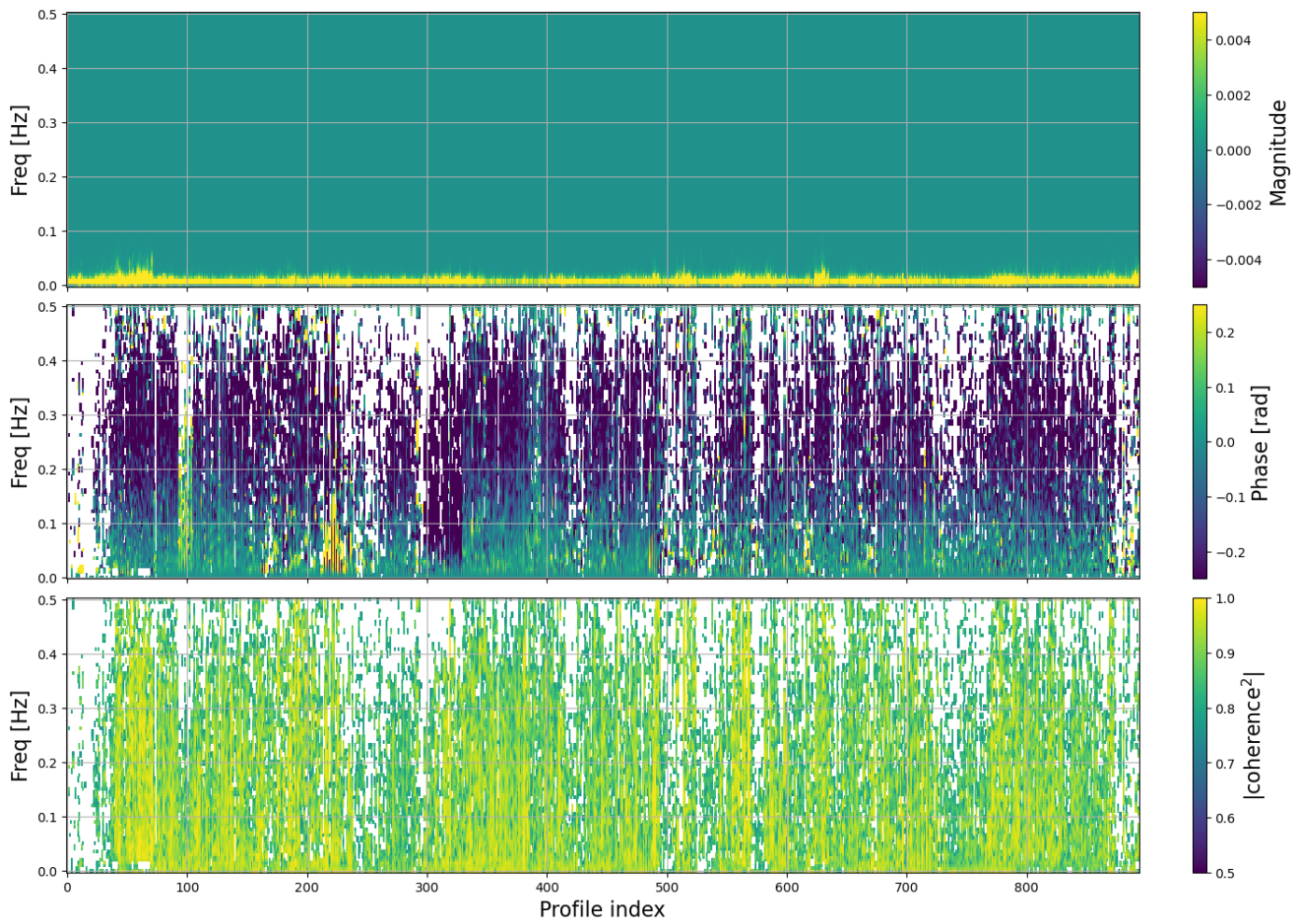
Step 1: directly estimate the sensor alignment correction constant following the methodology of Ferrari and Rudnick (2000):

1. Assume conductivity is approximately linearly related to temperature
2. Calculate the cross-spectrum between the standardized temperature and conductivity timeseries
3. Calculate the coherence, to determine where the cross-spectrum is valid
4. Fit the function $2\pi f\tau_C$ to the phase of the cross-spectrum, where f is frequency and τ_C is the correction constant we are looking for

These steps are applied to each individual profile, with fresh surface water **less dense than 1022.5 kg/m^3** excluded based on the T-S diagram (Fig. 3). The cross-spectrum is estimated using Welch's method, with a Hann window of **length 128 and 50% overlap**. The spectrogram (Fig. 4) shows only values for which the coherence is **at least 75%** and exceeds the 95% confidence level. The choice of 75% is arbitrary, and results are not sensitive to this choice within the range 50% to 90%.

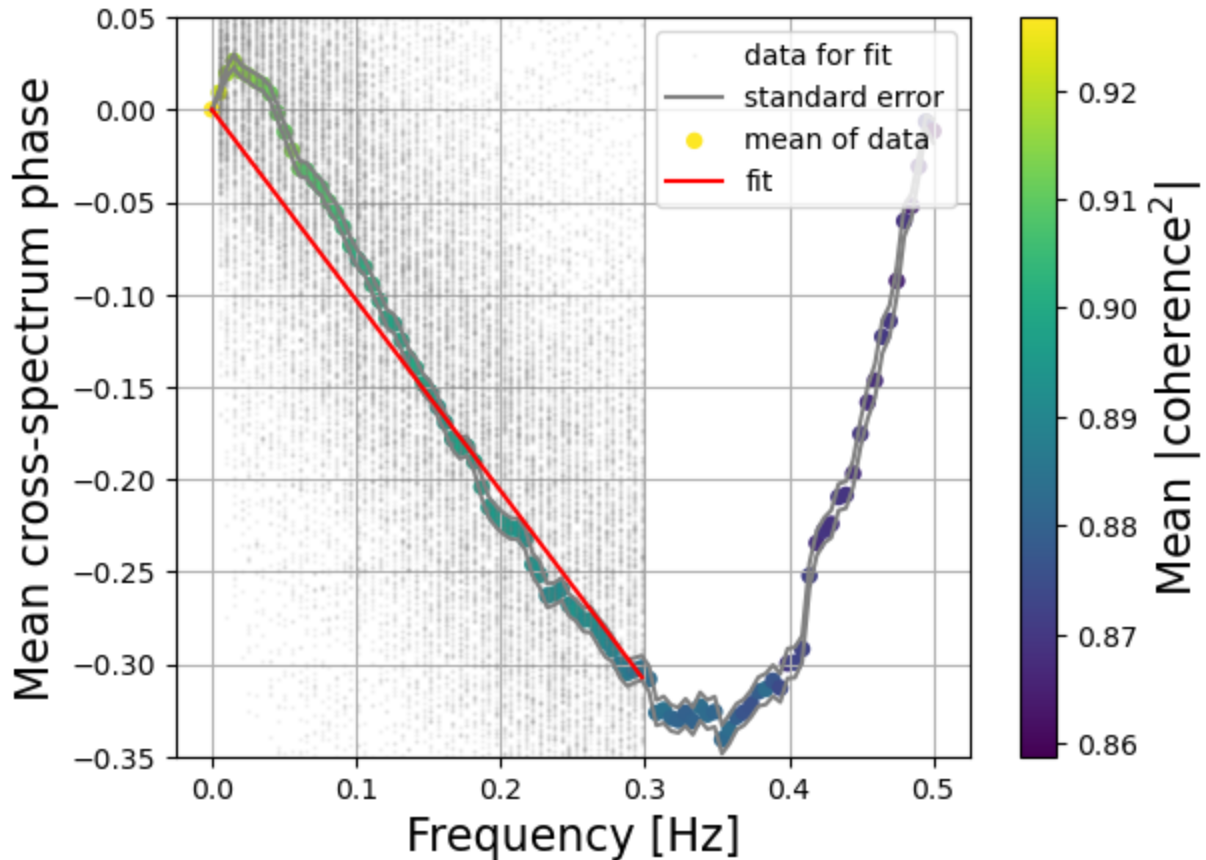
```
Applying density cutoff
Total number of profiles: 893
Ready to loop over profiles
Plotting
```

Fig. 4: Spectrogram for T and C cross-spectrum magnitude (top), phase (middle), and squared coherence (bottom) as a function of frequency and profile index. Only points exceeding the 95% confidence level and with coherence above 75% are shown.



The value of τ_C is -0.16406 with standard error 0.00078 .

Fig. 5: Cross-spectrum phase of temperature and conductivity plotted as small gray points, with the mean plotted as large dots coloured by mean squared-coherence. The gray envelope shows the standard error for the mean over all profiles. The red line shows the function fit to points for which the frequency is less than the chosen cutoff.



We fit the function only to frequencies **below 0.3 Hz** (Fig. 5). Above this frequency, we lose coherence and the slope of the cross-spectrum phase changes sign. We see that the resulting value for the **time constant** $\tau_C = -0.16 \text{ s}$ is well within the required $\pm 0.5 \text{ s}$ of the value specified for Sea-Bird GPCTDs, $\tau_C = -0.1 \text{ s}$.

Step 2: We can thus **confirm the Janzen and Creed (2011) constants** for the alignment correction are appropriate for use with this GPCTD. We preferentially use the Janzen and Creed (2011) constants over our directly estimated value for τ_C since we have found that the direct estimate is sensitive to the choices of frequency and density cutoff used, with reasonable choices producing constants that can vary by up to $\pm 0.5 \text{ s}$.

Step 3: Apply the correction using the alignment constants for conductivity and temperature provided in Janzen and Creed (2011). Specifically, advance conductivity by 0.4 s relative to pressure and advance temperature by 0.5 s relative to pressure. Figure 6 shows the magnitude of the change in temperature and salinity as a result of the sensor alignment correction. The changes in both fields are large enough to be significant for many data points.

Fig. 6: Left panel shows temperature vs. salinity diagram, where black contours give density in kg/m^3 . Right panel shows the difference between the aligned and uncorrected temperature vs. the difference between the aligned and uncorrected salinity.

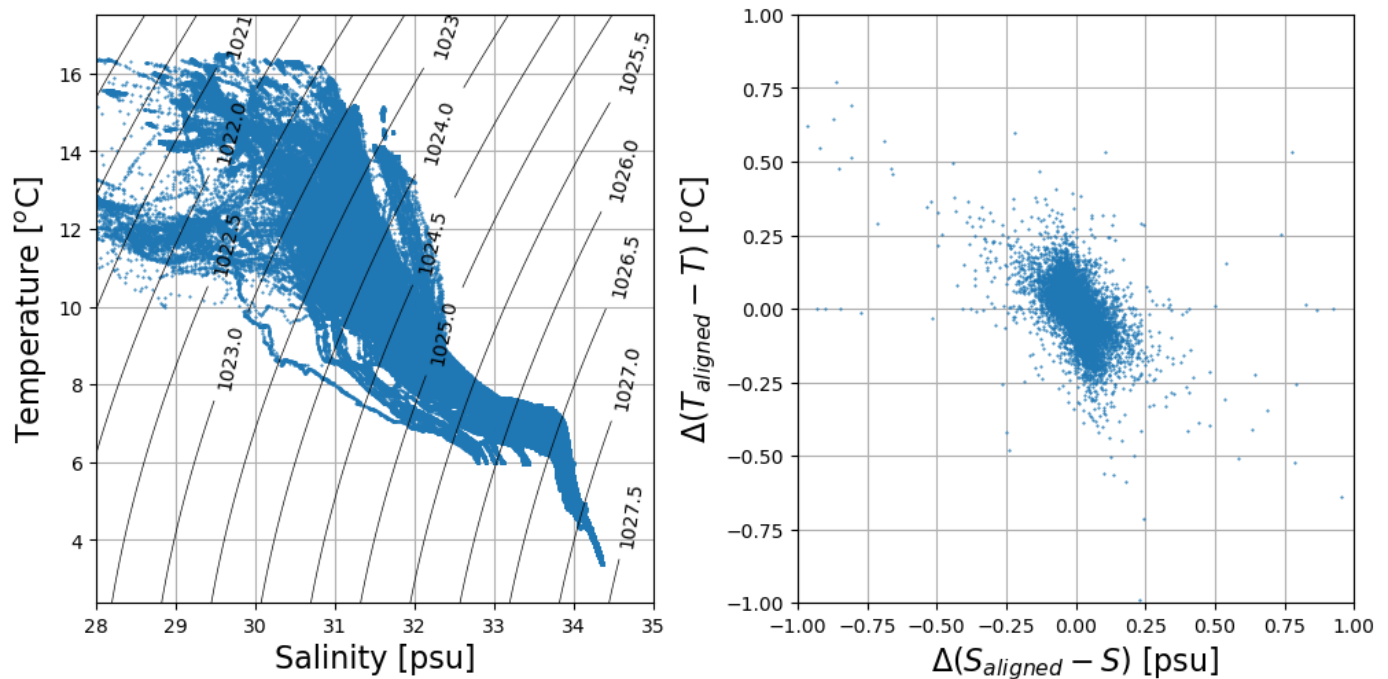
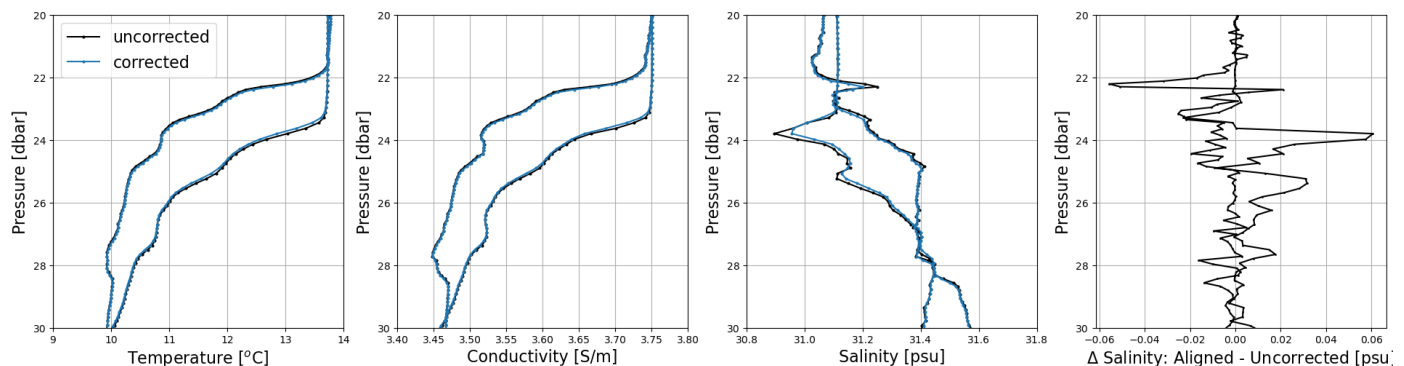


Fig. 7: Upcast and downcast of temperature, conductivity, and salinity vs. pressure, for part of the upper water column. Black indicates uncorrected fields, and blue indicates fields for which an alignment has been applied. Final panel shows the difference between the aligned and uncorrected salinity.



The salinity **does not have obvious sharp spikes** for this mission. Applying the alignment correction produces a small but visible reduction in small-scale unphysical variations in the salinity field in the region of

strong T and C gradients (Fig. 7).

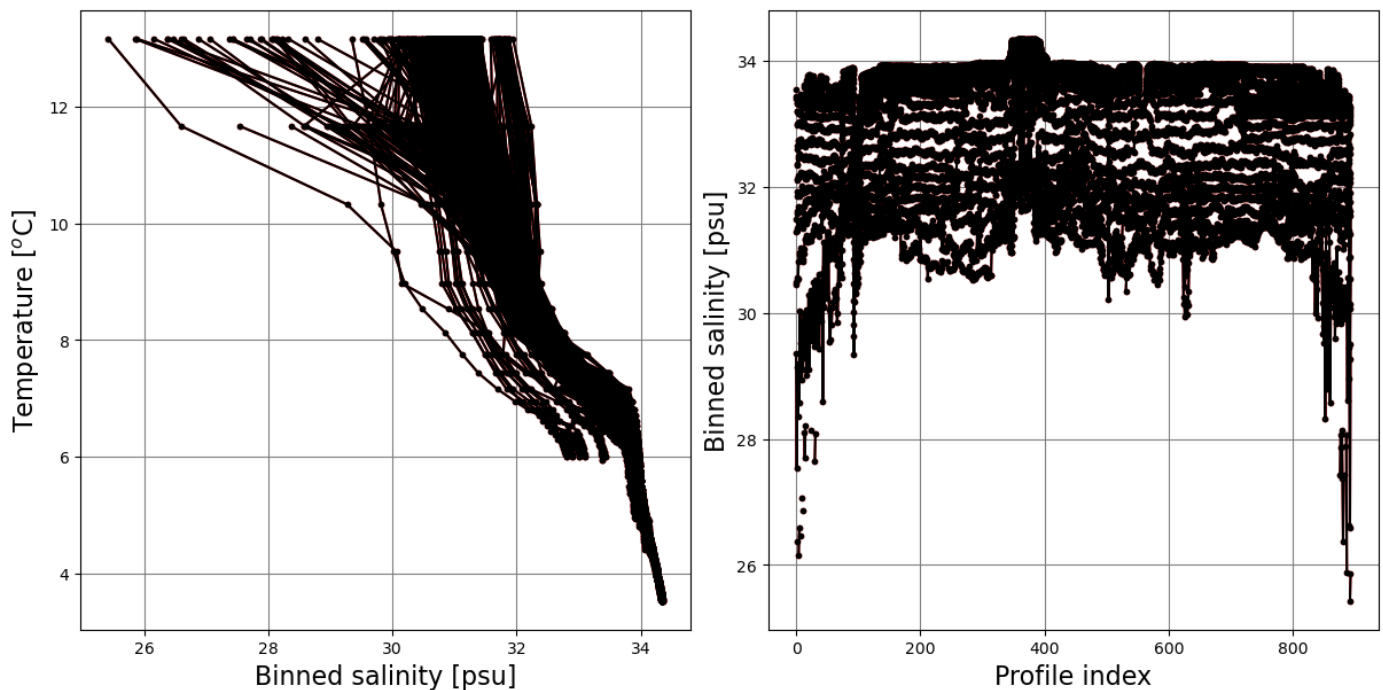
2.3 Identification and removal of questionable salinity profiles

We identify any salinity profiles that are obviously unphysical, which is typically caused by something (usually biology) getting caught in the conductivity cell, and set all values within those profiles to NaN. We use a simple criterion applied to the salinity data, binned by temperature, with bin sizes based on the time series mean temperature profile. The criterion temporarily flags any data points that are more than **4 standard deviations** away from the overall mean for the salinity time series within a given temperature bin, then recomputes the mean and standard deviation, excluding the temporarily flagged values. Salinity values that still differ from the mean by more than **4 standard deviations** are flagged as 'bad'. Finally, any profile where more than **10% of the salinity values** have been flagged as 'bad' using this criterion is removed. The number of standard deviations used and the percent of flagged required to flag a profile as 'bad' can be adjusted.

```
Profiles flagged as bad due to questionable salinity values: []
```

This mission on the Calvert Line began and ended in Hakai Pass, so that the glider was deployed and recovered close to shore in a coastal region where it is not uncommon to have freshwater pooled in a thin layer at the surface. In order to avoid excluding these low salinity values, and to exclude a possible eddy or anomalous watermass in the outbound trip, the **first and last 110 profiles** were excluded from the time series to which the correction was applied. Of the remaining profiles, **zero were flagged as bad** (Fig. 8) and had their values set to NaN.

Fig. 8: Binned salinity plotted as a function of temperature (left) and vs. profile index (right), with the salinity profiles identified as bad due to questionable values and set to NaN shown in red and indicated by the red arrows at the top of the panel on the right.



2.4 Thermal lag correction

The thermal lag effect is caused by the thermal inertia of the conductivity cell affecting the temperature of the water as it passes through the cell. To determine the thermal lag correction, the temperature inside the conductivity cell is estimated, then salinity is recalculated using the estimated temperature and the measured conductivity.

To estimate the temperature, a recursive filter is applied to the temperature field with parameters α (the amplitude of the error), and τ (the time constant for the thermal lag). The values of α and τ that minimize the area between pairs of profiles (each dive and subsequent climb along the glider path) are determined using a brute force minimization scheme. The quantity to be minimized is the root-mean squared difference (RMSD), which is calculated as the square root of the sum of the squared areas between pairs of salinity profiles (binned by temperature), normalized by the number of pairs of profiles.

Sea-Bird GPCTDs are pumped with a constant flow rate. As such, we expect the thermal lag to be approximately constant over the full mission, and it is sufficient to find a single value of α and τ for the entire mission. Janzen and Creed (2011) determined a cell thermal mass correction for the GPCTD using data from a prototype CTD that sampled twice as rapidly as the GPCTD nominally samples, with a pumped flow rate of 10 ml/s. **They found $\alpha = 0.06$ and $\tau = 10s$.**

Our procedure for the thermal lag correction:

1. For each mission, directly estimate the thermal lag parameters α and τ for that specific sensor
2. Confirm that the directly estimated value of τ is within $\pm 10s$ of the Janzen and Creed (2011) value of 10s
3. Quantify the improvement when using the Janzen and Creed (2011) parameters and the directly estimated parameters
4. Apply the thermal lag correction with the parameters that result in the greater improvement, and do not result in an over-correction
5. If a given sensor has a directly estimated value of τ that is significantly higher or lower than 10s, investigate further

Note that the thermal lag correction parameters are more likely to vary slightly between individual GPCTD sensors than the alignment correction constants.

This mission on the **Calvert Line** occurred in a highly energetic environment, so near-surface differences between a downcast and the subsequent upcast are likely to be caused by spatiotemporal variability. As such, we exclude segments of each profile in the upper water column for which the **density is $< 1023 \text{ kg/m}^3$** from the minimization routine. We also exclude the last **50 profiles**, which were primarily collected in the highly energetic environment of Hakai Pass. We use a subset of the remaining data, consisting of **100 pairs of profiles** equally spaced in time, to determine the correction.

Total number of profiles: 893

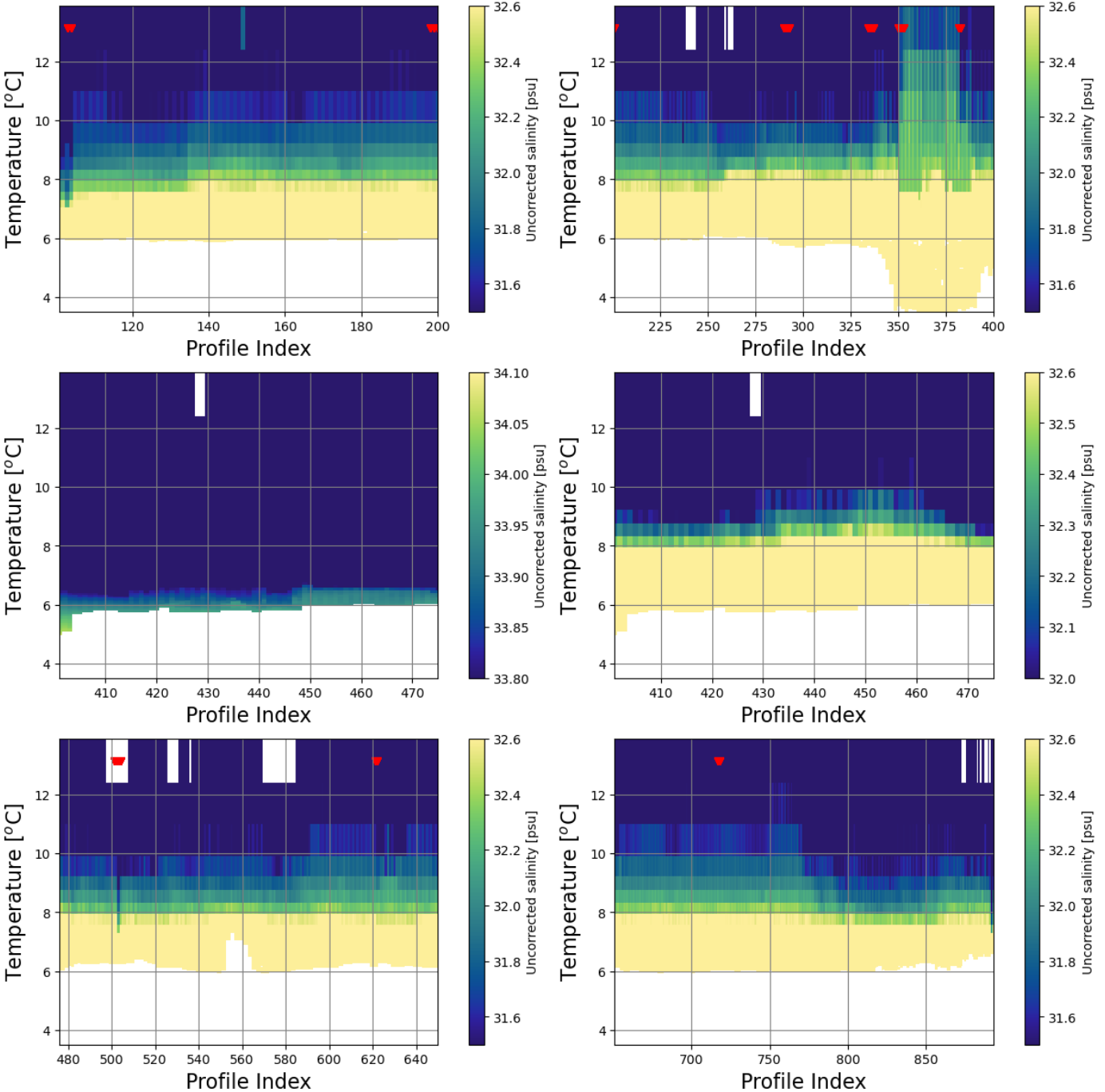
RMSD for uncorrected data = 0.4862111368516285

We exclude profiles from the correction for which the area between subsequent downcasts is more than one standard deviation from the mission mean. This ensures that no data crossing fronts or intrusions is included in the correction, in line with the key assumption that a downcast and the subsequent upcast be identical (Fig. 9).

Furthermore, we impose a cutoff for the area between pairs of profiles that will be included in the subset used to estimate the parameters. Any pair of profiles whose area is more than 3 standard deviations away from the mean will be excluded from the determination of the RMSD. This ensures that a small number of anomalous profiles do not bias the results.

Max and min area between downcasts = 10.950801621961073 8.338651828412877e-05
List of profiles to exclude: [71. 72. 73. 74. 91. 92. 93. 94. 95. 96. 99. 100. 103. 104.
198. 199. 200. 201. 290. 291. 292. 293. 334. 335. 336. 337. 350. 351.
352. 353. 382. 383. 501. 502. 503. 504. 621. 622. 717. 718.]

Fig. 9: Salinity, gridded in temperaturespace, for different segments of the deployment. The top panels show the outgoing leg of the deployment, the middle panels show profiles in water deeper than 1000m, with two different sets of colorbar limits, and the bottom panels show the return leg of the deployment. Red arrows indicate the profiles identified for exclusion using restrictions based on the area between subsequent downcasts.



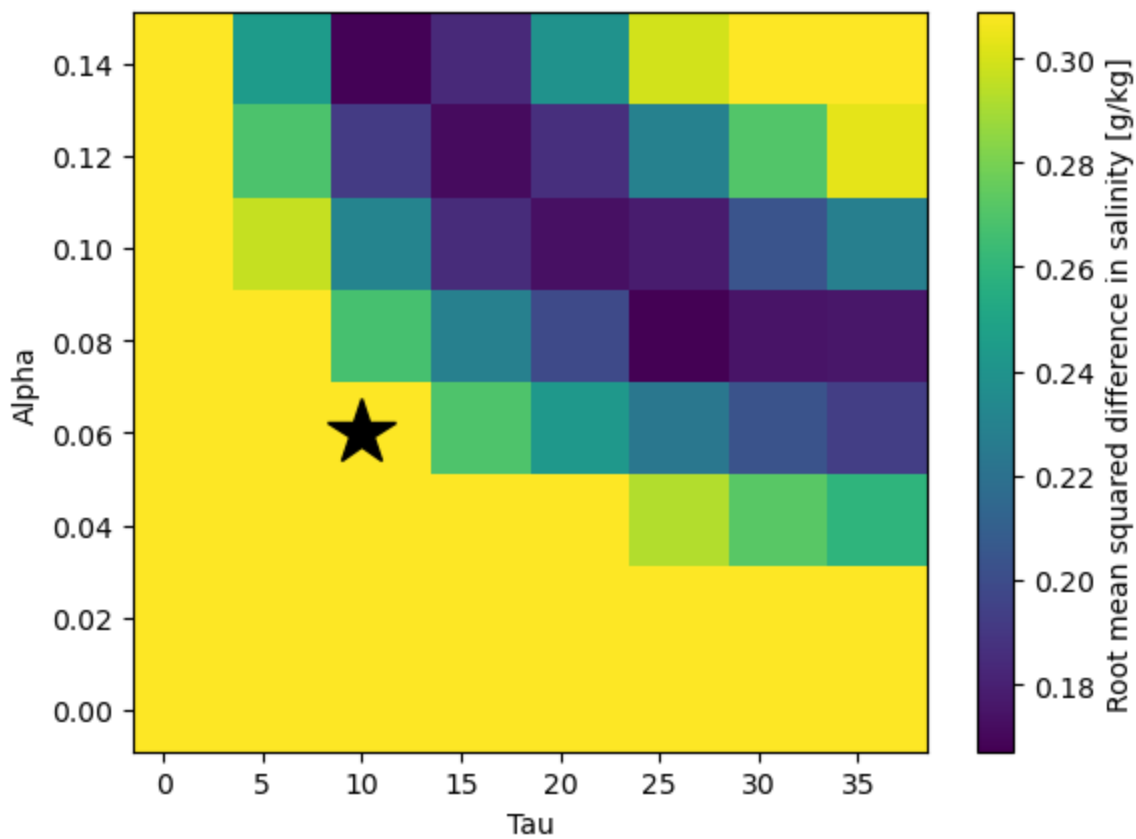
We now determine our filter coefficients α and τ to apply the correction. Given the recommended values of $\alpha = 0.06$ and $\tau = 10s$, we try values of τ between 1s and 35s, and values of α between 0.001 and 0.150 for our parameter space (Fig. 10).

```

Ranges of alpha and tau being tested are: slice(1, 152, 20) slice(1, 37, 5)
Tau = [25]
Alpha = [0.08]
RMSD = 0.16712811643628886

```

Fig. 10a: Values of the root-mean squared difference for each parameter value used during the grid search to minimize the area between pairs of profiles. Note that high values of the RMSD have been saturated (yellow) to better show the minimum (dark blue). The black star shows the Janzen and Creed (2011) values.

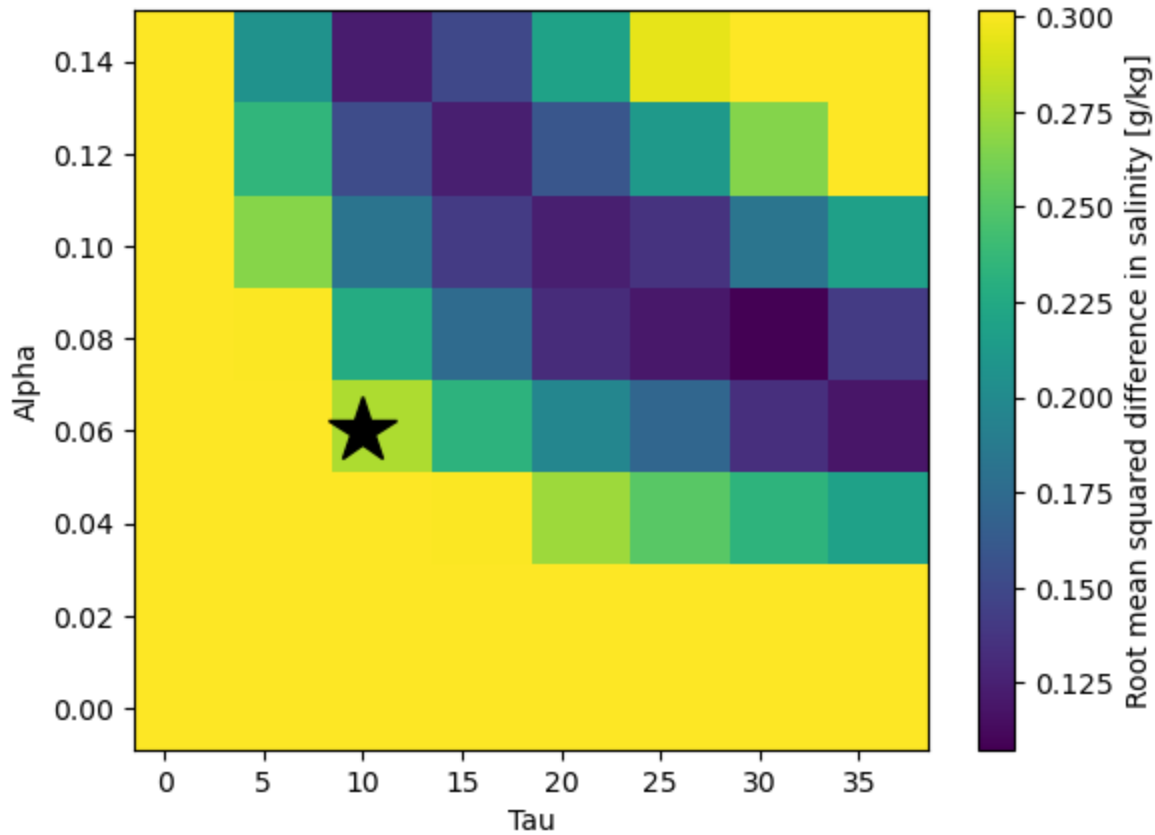


We see that there is a fairly wide band of 'good' values where the RMSD is low, with **the best directly determined parameter values being $\alpha=0.08$ and $\tau=25$ s**. This value of τ is larger than Janzen and Creed (2011), while α is similar, however the Janzen and Creed (2011) values fall just outside the band of low RMSD.

We now confirm that this band of 'good' values is the same for profiles taken from 1) the early part of the mission, using profiles from the first quarter of all profiles collected, and 2) the later part of the mission, using profiles from the last quarter of all profiles collected.

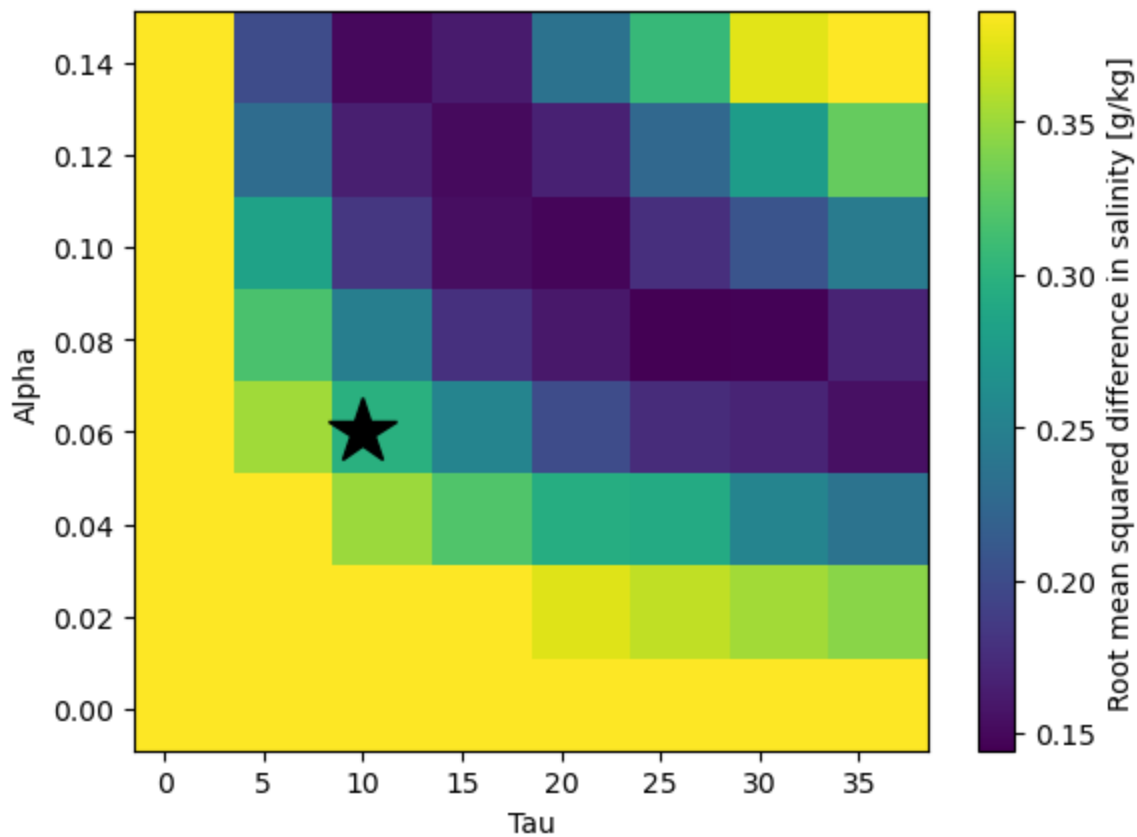
```
Ranges of alpha and tau being tested are: slice(1, 152, 20) slice(1, 37, 5)
Tau = [30]
Alpha = [0.08]
RMSD = 0.10746050096949732
```

Fig. 10b: Values of the root-mean squared difference for each parameter value used during the grid search to minimize the area between pairs of profiles for the first quarter of the mission. Note that high values of the RMSD have been saturated (yellow) to better show the minimum (dark blue). The black star shows the Janzen and Creed (2011) values.



Ranges of alpha and tau being tested are: `slice(1, 152, 20) slice(1, 37, 5)`
 Tau = [25]
 Alpha = [0.08]
 RMSD = 0.14429309512057711

Fig. 10c: Values of the root-mean squared difference for each parameter value used during the grid search to minimize the area between pairs of profiles for the last quarter of the mission. Note that high values of the RMSD have been saturated (yellow) to better show the minimum (dark blue). The black star shows the Janzen and Creed (2011) values.



For the first and last quarter of the mission, the values of $\tau=30s$ and $25s$ and $\alpha=0.08$ are similar to those for the mission as a whole, and the location of the band of low values is the same.

Since our directly determined parameters $\tau=25s$ and $\alpha=0.08$ fall just outside of the band of low RMSD values as the Janzen and Creed (2011) parameters, i.e., they result in only marginally lower RMSD overall compared to the Janzen and Creed (2011) parameters, we will quantify the improvement for both our directly determined values and the Janzen and Creed (2011) parameters before deciding which to use for the final correction.

Quantifying improvement for our directly determined parameters:

0.08 25

The effect of applying our directly determined thermal lag correction on the salinity data is shown in Figure 11. It's clear that the correction has **significantly reduced the up-down asymmetry** between pairs of profiles, throughout the water column.

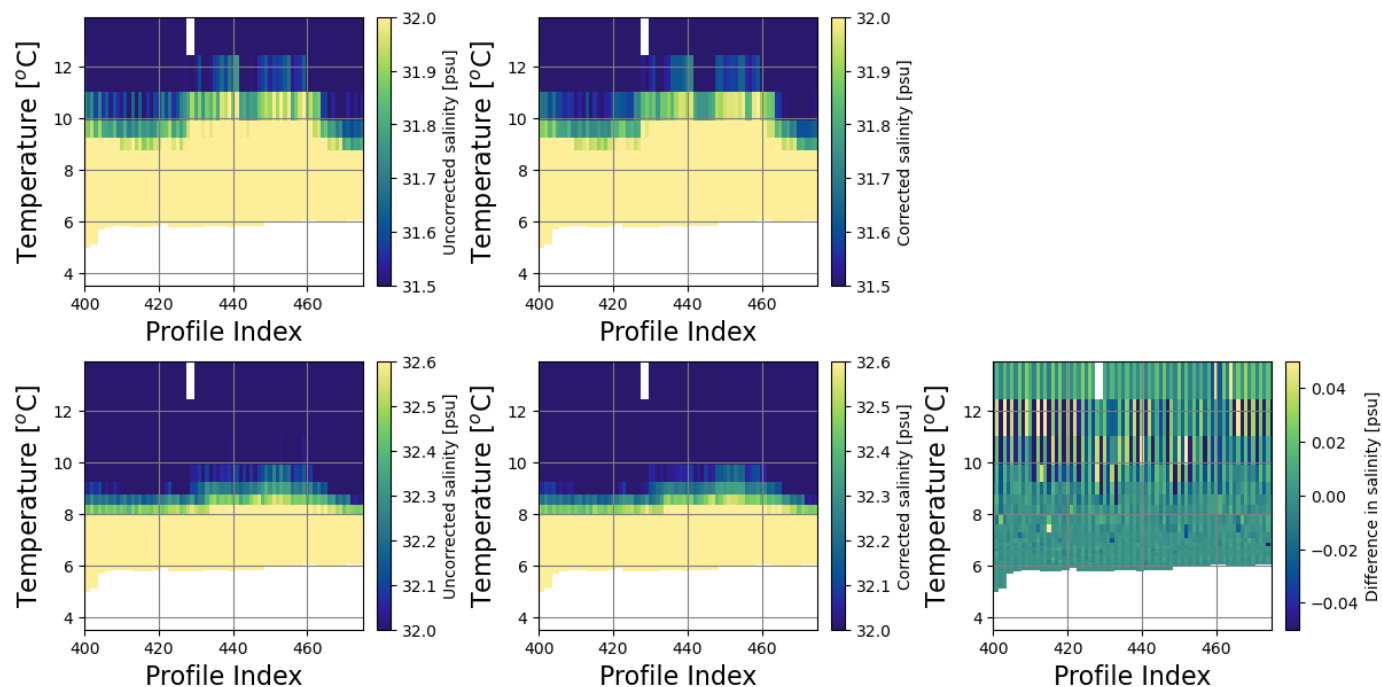


Fig. 11: Clean salinity (with bad profiles removed), plotted vs. temperature and profile index. Profiles deeper than 1000m are shown. The uncorrected salinity is shown in the left panels, the corrected salinity is shown in the middle panels, and the difference between the two is shown in the right panel. The bottom row is identical to the top row, but with the colorscale adjusted to show depths within the thermocline.

To quantify the magnitude of the improvement, Figure 12 shows the area between all pairs of profiles considered, plotted versus profile index and as a histogram, as well as the difference in area between the corrected and uncorrected data.

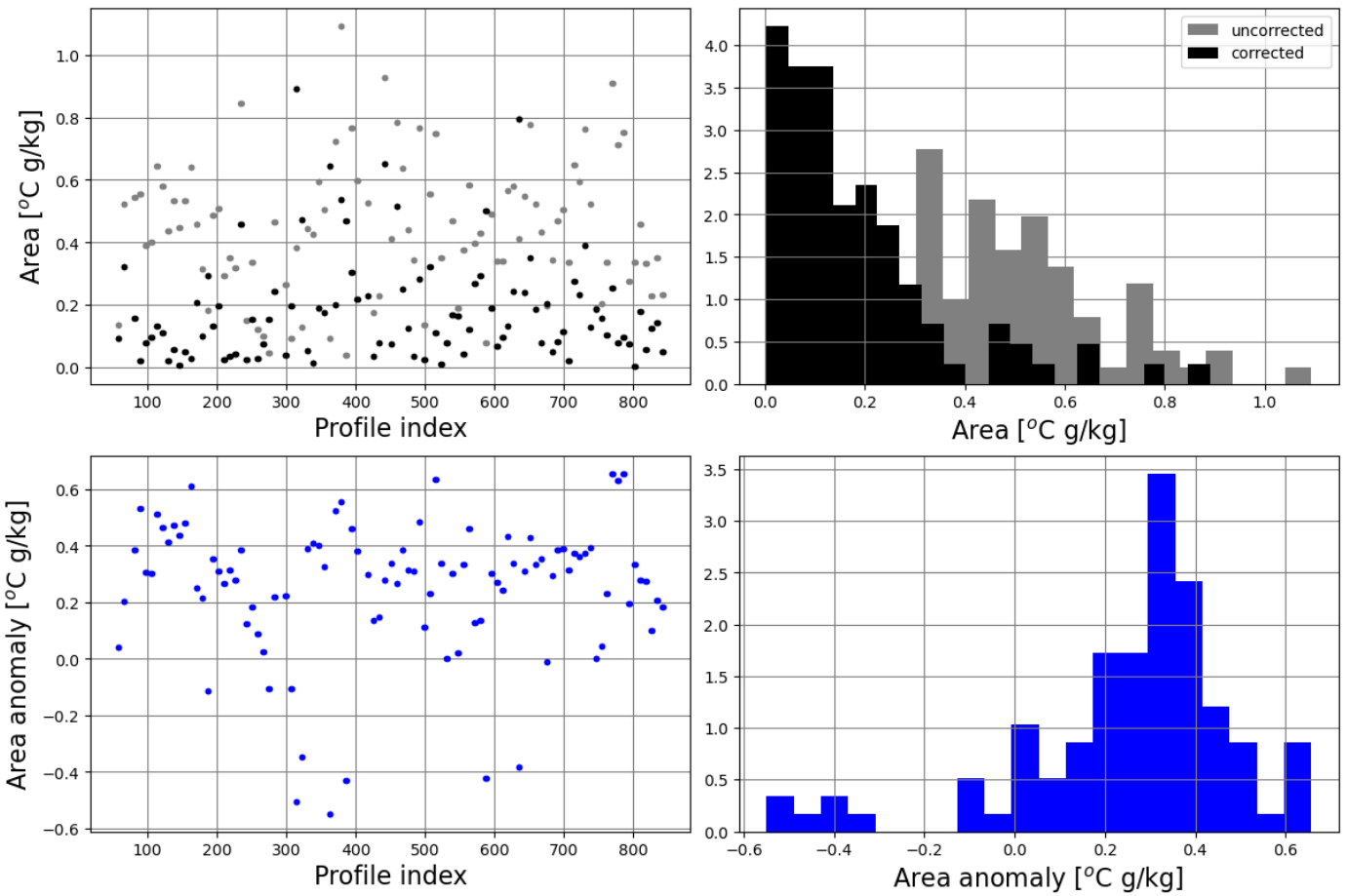


Fig. 12: Area between pairs of salinity profiles, calculated in temperature-salinity space, plotted vs. profile index number (left) and as a histogram (right), for the uncorrected salinity field (grey) and the corrected salinity field (black), and as an anomaly between the uncorrected and corrected fields (bottom row, blue).

Total area mean = 0.1831788197858375

Total area anomaly mean = 0.2528230133096317

Quantifying improvement for the Janzen and Creed (2011) parameters:

0.06 10

The effect of applying the Janzen and Creed (2011) thermal lag correction on the salinity data is shown in Figure 13. It's clear that the correction has **somewhat reduced the up-down asymmetry** between pairs of profiles, throughout the water column.

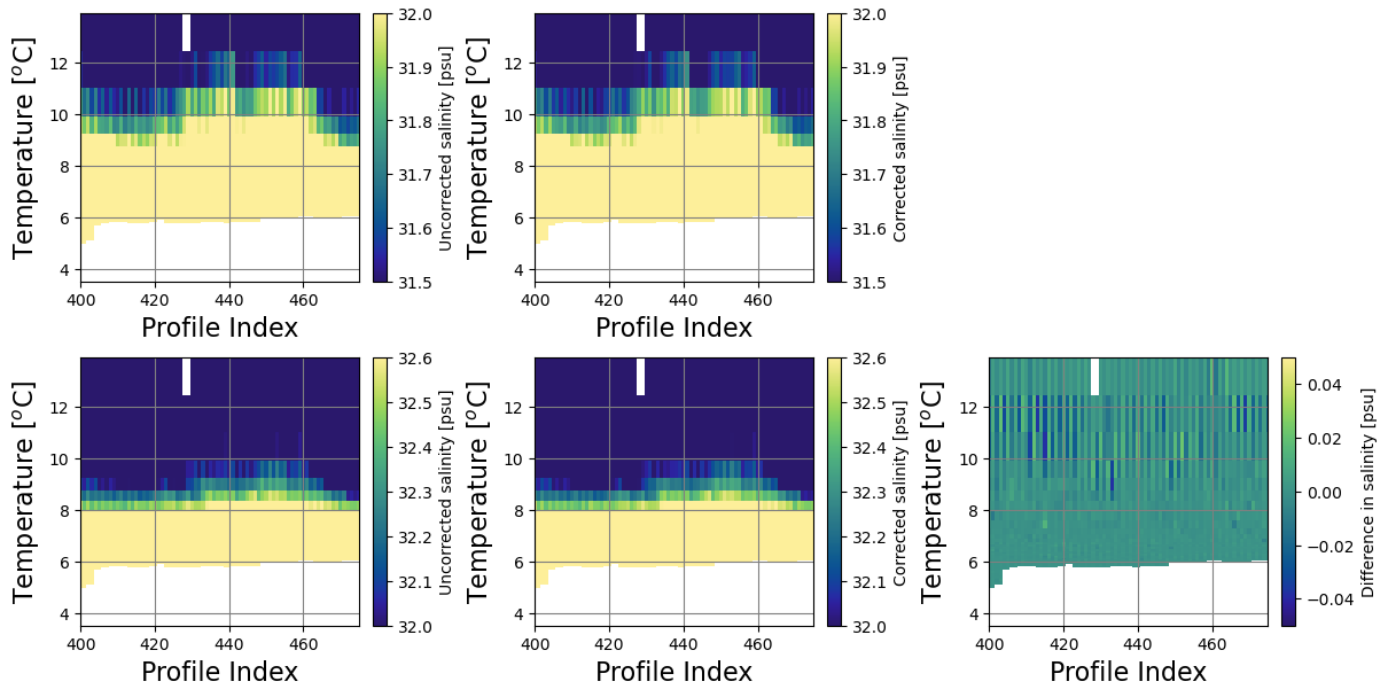


Fig. 13: Clean salinity (with bad profiles removed), plotted vs. temperature and profile index. Profiles deeper than 1000m are shown. The uncorrected salinity is shown in the left panels, the corrected salinity is shown in the middle panels, and the difference between the two is shown in the right panel. The bottom row is identical to the top row, but with the colorscale adjusted to show depths within the thermocline.

To quantify the magnitude of the improvement, Figure 14 shows the area between all pairs of profiles considered, plotted versus profile index and as a histogram, as well as the difference in area between the corrected and uncorrected data.

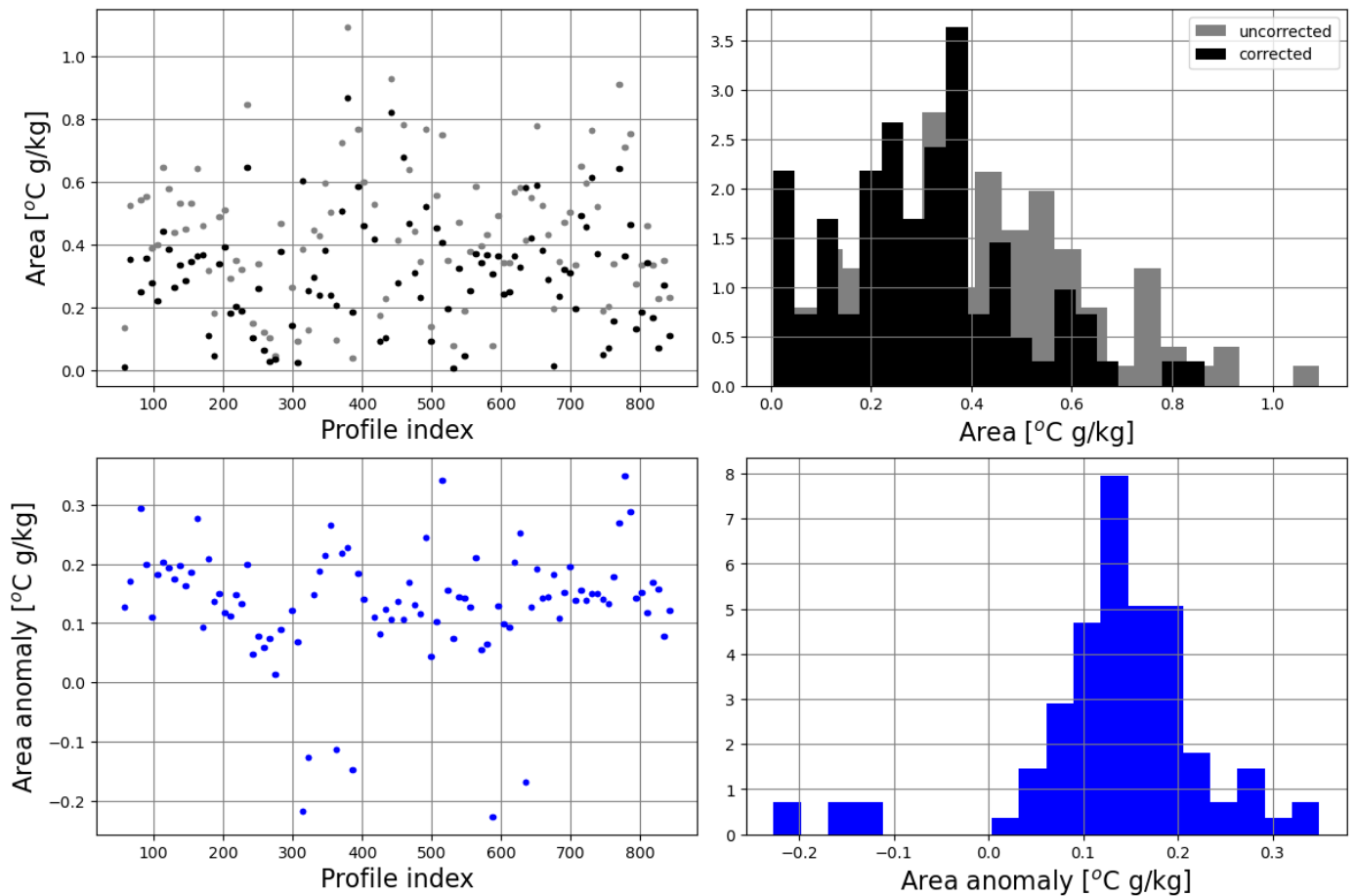


Fig. 14: Area between pairs of salinity profiles, calculated in temperature-salinity space, plotted vs. profile index number (left) and as a histogram (right), for the uncorrected salinity field (grey) and the corrected salinity field (black), and as an anomaly between the uncorrected and corrected fields (bottom row, blue).

Total area mean = 0.30339069098796495

Total area anomaly mean = 0.13261114210750424

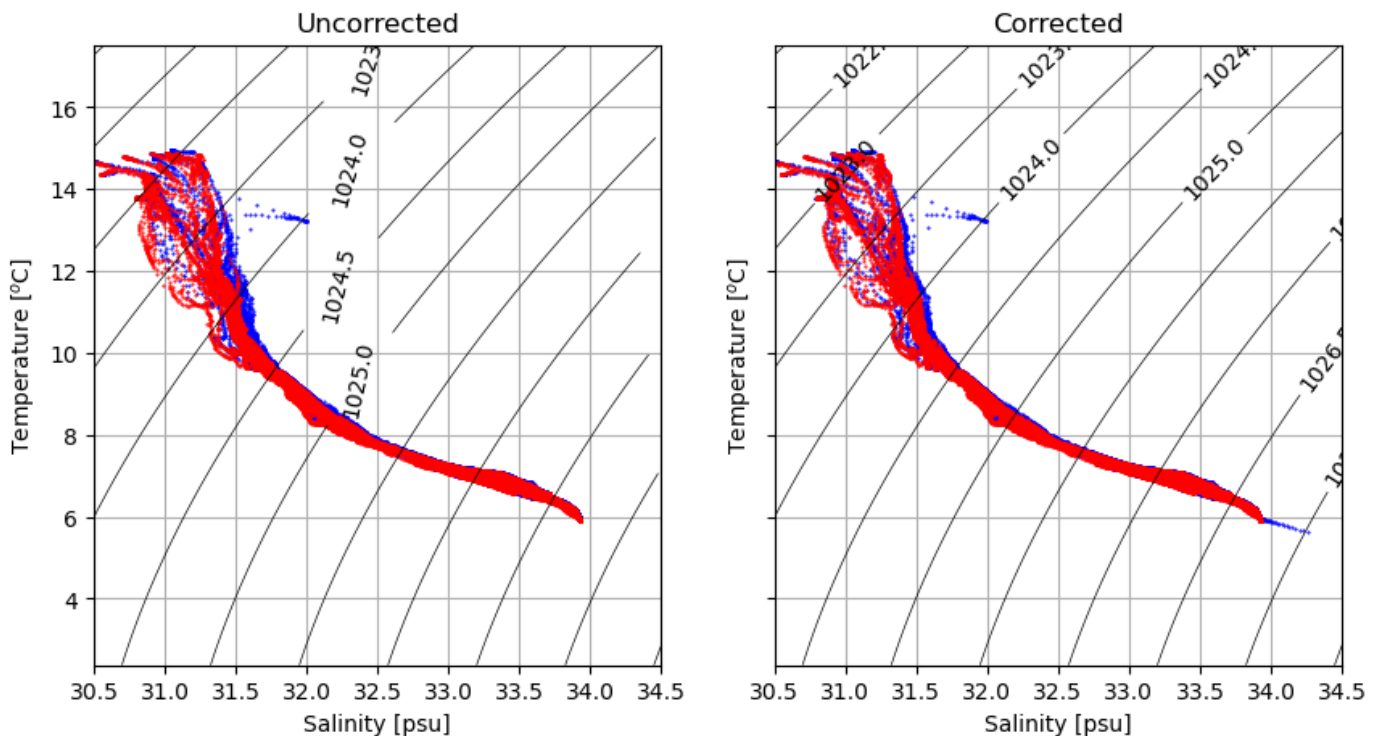
Summary of thermal lag correction

While the quantitative improvement is larger for our directly determined parameters, the large value of τ (and previous large values of τ for CTDs on this platform) raises concerns. It is likely that there is an issue with the CTD configuration or installation on this glider, and it is flagged for further investigation. For now, we use the **Janzen and Creed (2011)** parameters for our final thermal lag correction. We note that there will be some residual up-down asymmetry, however using the directly determined values of α and τ would result in undesirable smoothing of the data due to the large coefficient values.

0.06 10

0.06 10

Fig 15: Temperature-salinity diagrams for a subset of all profiles, showing the difference between upcasts (red) and downcasts (blue), for the data without the thermal lag correction applied (left panel) and the data with the thermal lag correction applied (right panel).



3.0 Summary of corrections applied to delayed mode data for this mission

1. Identification of anomalous conductivity values:

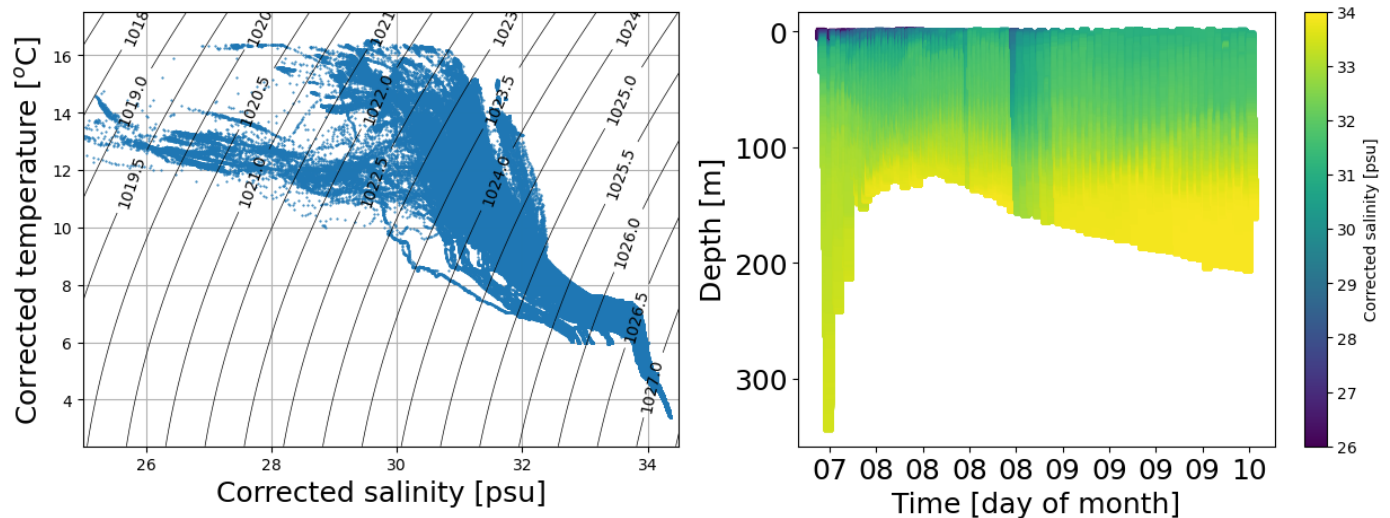
- Anomalous conductivity values at the surface caused by air bubbles in the cell were set to NaN.

2. Sensor alignment correction

- Following Janzen and Creed (2011), temperature was advanced by +0.5 s relative to pressure, while conductivity was advanced by +0.4 s relative to pressure.
- These values were confirmed by directly estimating the time constant using the methodology of Ferrari and Rudnick (2000).

- The correction produced a small but quantifiable reduction in small-scale unphysical variations in the salinity field in the region of strong temperature and conductivity gradients.
3. Identification of questionable salinity profiles
 - Zero salinity profiles were flagged as 'bad' and their values set to NaN. We note that 110 profiles were excluded from this correction at the start of the mission due to the presence of a possible eddy or anomalous water mass feature.
 4. Thermal lag correction
 - The directly determined values for the thermal lag correction produced an improvement that was larger than the recommended values from Janzen and Creed (2011), however the values of the correction coefficients were concerningly high and the platform has been flagged for further investigation.
 - The final thermal lag correction was applied using the Janzen and Creed (2011) values of $\alpha=0.06$ and $\tau=10s$.
 - The correction reduced the root-mean squared difference for the area between between pairs of profiles by 60%, and adjusted the salinity field by at least 0.05 psu in the thermocline.

Fig. 16: The corrected temperature and salinity fields shown in a T-S diagram with density contours (left) and the corrected salinity field plotted vs. depth and time (right).













Corrected data saved to file: ./dfo-bb046-20220707/dfo-bb046-20220707_adjusted.nc

Out[40]: xarray.Dataset

► Dimensions: (time: 1759051)

▼ Coordinates:

time	(time)	datetime64[ns]	2022-07-07T16:45:19.204999936	 
latitude	(time)	float64	...	 
longitude	(time)	float64	...	 
depth	(time)	float64	...	 
pind	(time)	float64	-0.0 0.0 0.0 0.0 ... -0.0 -0.0 -0.0	 

► Data variables: (22)

► Indexes: (1)

► Attributes: (67)

