Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Validating field sensor performance of a CTD-SRDL tag using Slocum gliders

Caroline J Wiernicki

cwiernic@udel.edu

University of Delaware

Joseph C Gradone

Rutgers University

Carlos F Moffat

University of Delaware

Aaron B Carlisle

University of Delaware

Matthew J Oliver

University of Delaware

Method Article

Keywords: Ocean observing, gliders, sensor validation, animal telemetry

Posted Date: November 24th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11526933/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Title

Validating field sensor performance of a CTD-SRDL tag using Slocum gliders

Authors

Caroline J Wiernicki[1]*, Joseph C Gradone[2], Carlos F Moffat[1], Aaron B Carlisle[1], and Matthew J Oliver[1]

*cwiernic@udel.edu

1. School of Marine Science and Policy, University of Delaware, Lewes, DE, United States

2. Center for Ocean Observing Leadership, Department of Marine and Coastal Science, Rutgers University, New Brunswick, NJ, 08901, United States

Abstract

Background: Government and academic stakeholder groups are investing in marine animals collecting *in situ* oceanographic data as ocean observing platforms (OOPs). Proof-of-concept began with the deployment of Conductivity Temperature Depth (CTD) recording Satellite-Relayed Data Logger (SRDL) tags on Southern elephant seals (*Mirounga leonina*) in the early 2000s. The number of marine animal OOPs has since exploded across taxa globally. However, the effective utility of marine animals as OOPs requires a thorough understanding of how the operational quality of animal-collected CTD data compares to that of data collected by established platforms, such as autonomous gliders. In this study, we attached a CTD-SRDL tag to a Slocum glider to compare tag sensor measurements with those concurrently collected by the Slocum glider-borne CTD to develop a baseline understanding of tag performance in relatively homogeneous versus stratified water column conditions and as a function of platform motion.

Results: Comparisons of glider and tag CTDs showed maximum differences when sampling highly stratified water columns. We quantified the effect of the tag's clock lag on sensor error and bias and identified coefficients to correct this impact in the tag data. We identified similar coefficients to correct for offsets in the pressure sensor of the tag and demonstrate the effects of both corrections on data quality. Investigation of the effects of platform motion provided evidence of a moderate impact of incident velocity on temperature and conductivity measurements taken during stratified conditions.

Conclusions: Corrections for clock and pressure offset should be applied to animal-collected profiles in post-processing. Calculated ranges for sensor root mean square error and bias should inform the accurate application of data to oceanographic calculations, depending on the anticipated water column conditions to be sampled. Smaller spatiotemporal scale calculations

may be more sensitive to sensor error, whereas broader calculations (such as ocean heat content and practical salinity) are more robust to tag-platform error.

Key words

Ocean observing; gliders; sensor validation; animal telemetry

Background

Concurrent advances in oceanographic sensors and animal biologging techniques have generated interest by government, academic, and industry stakeholders to use tagged marine animals as ocean observing platforms (OOPs) (McMahon et al. 2021; Harcourt et al. 2019). As tagged animals navigate their local environment, collected oceanographic data are transmitted via satellite when the tag comes to the ocean's surface, permitting simultaneous collection of oceanographic and ecological data (Block et al. 2016; March et al. 2020). Widespread interest in marine animals as OOPs began with the first deployments of a novel conductivity-temperature-depth (CTD) recording Satellite-Relayed Data Logger (SRDL) tag on Southern Elephant Seals (*Mirounga leonina*) in the Southern Indian Ocean in the early 2000s. Fifty-eight tagged seals collected and transmitted 16,500 temperature-salinity profiles around Antarctica in 2004-2005 (Charrassin et al. 2008). A quarter of these profiles were taken in the sea ice zone during austral autumn-winter, a season during which traditional ship-board measurements of the water column are severely limited by weather conditions and sea ice, and most oceanographic sampling is limited to satellite measurements at the surface. These 4,000 winter sea ice profiles therefore represented a 30-fold increase in data obtained from this season and region, compared to the existing number of non-animal-collected profiles. This extreme increase in total collected

profiles emphasizes both the reliance on autonomous observations for sampling hostile environments (such as storms, icy regions, remote locations) as well as the relatively limited size of existing observations in these regions.

Since this initial proof-of-concept, the number marine animals serving as OOPs has expanded. Temperature-depth recording tags have been deployed on sea turtles to map the water column thermal structure in coastal environments, including Loggerhead (*Caretta caretta*) turtles in the Mid-Atlantic as well as on Loggerhead and Olive Ridley (*Lepidochelys olivacea*) turtles in the Southwest Indian Ocean (Patel et al. 2018; Bousquet et al. 2020). Narwhals (*Monodon monoceros*) have been similarly tagged to record and transmit water column temperatures when overwintering in the Arctic, providing valuable seasonal data on the Baffin Bay ice pack (Laidre & Heide-Jørgensen 2007). More recently, tagging efforts have shifted from targeting air-breathing marine animals toward fish. Although air breathers transmit data more reliably via satellite, as they are required to surface regularly to breathe, the taxonomic diversity of animals used as OOPs (and the range of habitats they can sample) continue to diversify. Specifically, sharks have been identified as having high potential to serve as OOPs, given their diverse life histories, biogeographies, and movement behaviors, the latter characterized by regular deep dives and surfacing events (Wiernicki et al. 2025; Andzrejaczek et al. 2022; Andzrejaczek et al. 2019). Sharks tagged with temperature-depth recorders have successfully collected and transmitted water column data off the Alaskan coast, where a single salmon shark transmitted 56 profiles over 36 days, as well as off Hawai'i, with a tagged tiger shark transmitting 520 profiles over 102 days (Holland et al. 2022; Pagniello et al. 2024).

The rise in the number of temperature-depth recording tag deployments on a growing diversity of marine animal species and habitats suggests that marine animal OOPs will

increasingly contribute to the Global Ocean Observing System (GOOS). As of March 2024, ~800,000 vertical temperature-salinity profiles have been collected and transmitted to the GOOS by globally tagged marine mammals (https://www.meop.net/index.html), ~26% of the three million profiles collected thus far by the Argos float program (https://argo.ucsd.edu/about/status/). The GOOS is a global network of systems collecting, synthesizing, and sharing oceanographic data in near-real time, and currently consists of in situ and remote instrumentation such as ship-based measurements, mooring buoy arrays, satellite, and the global drifter program (Legler et al. 2015; Liblik et al. 2016). Autonomous and semi-autonomous platforms such as Argo floats and gliders currently play a substantial role in providing subsurface oceanographic data to the scientific community. The Argo float program refers to a global array of approximately 4,000 drifting instruments that control their descent and ascent through the water column over 10-day cycles, recording temperature, salinity, and occasionally biogeochemical variables to monitor heat content in the ocean's upper 2000 m as they float passively with currents (Riser et al. 2016; Lin and Yang 2020). The float array has provided subsurface data globally, particularly in sampling regions inaccessible to ships, such as those beneath sea ice at the poles. However, while individual floats are able to resolve oceanographic features at shorter time and distance scales (seconds to days and centimeters to kilometers), and the float network collectively resolves global features over the course of months to years at current float densities, the program is unable to resolve mesoscale features and interactions and is limited in its ability to sample coastal environments and insular seas. (Liblik et al. 2016).

The increased use of gliders to sample coastal and offshore marine environments helps fill this mesoscale feature gap in the GOOS. Gliders are autonomous underwater vehicles that

use shifts in buoyancy to generate lift beneath paired wings and move horizontally through the water column. They were developed in the early 2000s based on the buoyancy-driven vertical motion of the highly successful Argo floats (Sherman et al. 2001; Eriksen et al. 2001; Webb et al. 2001). Gliders may be equipped with a variety of modular payloads that can contain physical, biogeochemical, optical, and/or acoustics sensors and can operate at depths ranging from 20 to 1250 m, transmitting collected profiles at prescribed surfacing intervals via satellite. Given that the buoyancy-driven propulsion system is energetically efficient, glider deployments may range in duration from weeks to over a year, depending on the scientific sensors and additional battery payloads equipped. Over the course of these missions, the vehicles typically travel at vertical speeds of 0.2 $m \cdot s^{-1}$ and horizontal speeds of 0.25 $m \cdot s^{-1}$, allowing them to cover distances ranging from hundreds to thousands of kilometers per deployment (Liblik et al. 2016; Rudnick 2016). These mission durations and distances allow single gliders to resolve oceanographic features at scales of meters to hundreds of kilometers over weeks to months, complementing other systems in the GOOS to resolve mesoscale features such as eddies, fronts, internal waves, and upwelling (Liblik et al. 2016, Miles et al. 2017). Gliders have shown utility for improving data collection to inform mechanisms of boundary currents, tropical storm systems, and water mass transformation phenomena such as mixing or export events (e.g., Testor et al. 2019).

While the use of autonomous subsurface platforms such as Argo floats and gliders has vastly improved the quantity, quality, and delivery of near-real time subsurface data to the GOOS, there are still gaps, costs, and limitations to both the deployment and maintenance of these devices, as well as to the types of data products they can provide. Coastal and marginal seas remain under-sampled by the Argo program compared to offshore environments (Roemmich et al. 2009). Similarly, gliders have demonstrated utility in sampling these coastal, shallow

regions, but their slow speeds create highly correlated, repeated profiles, which limits the quantity of statistically decorrelated profiles (Rudnick et a l. 2016; Wiernicki et al. 2025). The success of using marine mammals to increase profile collection in the Southern Ocean, alongside the demonstrated viability of both air- and water-breathing taxa of marine animals to transmit water column data, suggests an opportunity for marine animal OOPs to supplement the sampling of mesoscale oceanographic features in a variety of marine environments, particularly coastal and marginal sea habitats (March et al. 2020). Though there is ample evidence that marine animals can contribute to the GOOS effectively, the instrumentation behind animal OOP data collection may have impacts on measurement resolution and accuracy that merit scrutiny.

Numerous studies have cited the ability of marine animal platforms to reliably transmit profile data, and others have explored and defined quality control metrics to correct animal-collected profiles (Mensah et al. 2018; Siegelman et al. 2019; March et al. 2020; Jonsen et al. 2024). However, there have been few studies examining how the variability of platform motion as well as the presence of strong environmental gradients (such as thermoclines or pycnoclines) impact tag sensor performance on a free-swimming animal. The GOOS has established strict protocols for defining oceanographic variables of common interest to the research, management, and industrial communities, known as essential oceanographic variables (EOVs), and has strict protocols for quality-controlling EOV data prior to its integration into the global telecommunication system (GTS) for global use (NOAA-IOOS). For example, the National Oceanic and Atmospheric Administration's Integrated Ocean Observing System strongly recommends that in-situ temperature and salinity data pass thirteen quality-control thresholds and tests that target data gaps, anomalous values, and spatiotemporal misalignments caused by sensor drift, resolution limits, platform motion, thermal lag, and/or biofouling (NOAA-IOOS).

Thermal mass error (thermal lag) is a common source of error in CTD data, caused by the transfer of heat between the sensor body and the water mass being sampled, and can impact both temperature and conductivity measurements (Lueck & Picklo 1990). The effects of thermal lag are expected to be higher when sampling water columns with larger thermal stratification, as well as when sampling with a non-pumped versus pumped CTD, wherein the rate at which water samples pass through the sensor will be regulated by platform speed and incident velocity rather than a regularized sampling frequency (Garau et al. 2011; Alvarez et al. 2013; Ullman & Herbert 2014). Platform motion is also a known source of error when collecting CTD data, again related to variable interactions of platform speed, direction, and thermal mass. (Trump 1983; Wang et al. 2020).

To maximize and fully realize the practical utility of CTD-SRDL tag data as a regular contributor to the GOOS, it is necessary to quantify baseline tag sensor performance in the field relative to a concurrently mobile control platform. This study intends to establish this baseline of field performance through the joint deployment of CTD-SRDL tags affixed to Slocum gliders, during mixed and strongly stratified oceanographic conditions. The objective of this study is to quantify tag performance in the field by comparing CTD-SRDL and glider measurements from calibrated sensors, specifically quantifying potential impacts of platform motion and thermal stratification on tag measurements. Ultimately, this approach will define a baseline of tag field performance that will then be able to inform, and potentially correct, CTD data products collected during deployment on free-swimming marine animals.

Methods

Data collection

A CTD-SRDL tag (a prototype tag body designed for deployment on sharks; Sea Mammal Research Unit [SMRU], University of St. Andrews) was deployed with a Slocum glider for two deployments during 2022 off the coasts of southern New Jersey and Delaware. The first tag-glider deployment took place in March 2022 on a G2 Slocum glider during well-mixed hydrographic conditions, and the second deployment occurred in September 2022, on a G3 Slocum glider during highly stratified conditions (Figure 1). The tag was mounted atop the glider body during each deployment, at the leading edge of the science payload, with a plastic mount hose-clamped around the body of the glider; the tag's CTD port was approximately 21.6 cm above each glider's CTD port. The G2 and G3 were equipped with pumped CTDs (Sea-Bird Scientific) and sampled at ½ Hz. The CTD-SRDL was sampled at 1 Hz with non-pumped flow (Table 1).

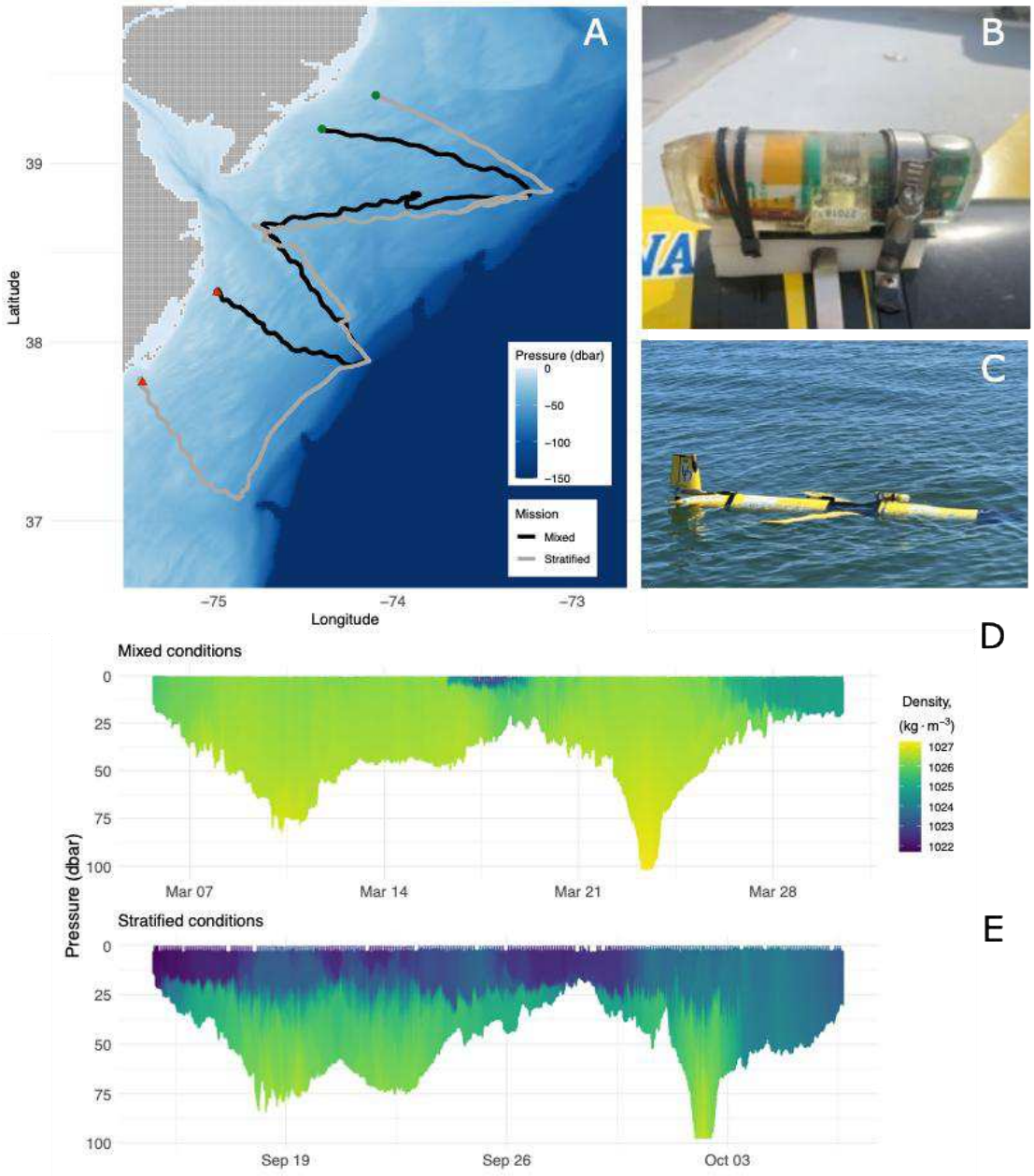

Figure 1. A: Map of March (mixed condition, black) and September (stratified condition, grey) glider-tag mission trajectories, with mission start and end locations denoted by green circles and red triangles, respectively; B: Image of the CTD-SRDL tag mounted to the top of the G2; C: Image of the CTD-SRDL tag, with repotted battery, mounted to the top of the G3; D: Density derived from glider-measured temperature, pressure, and conductivity over depth and time during the March mission; E: Density derived from glider-measured

temperature, pressure and conductivity over depth and time during the September mission. The water column became well-mixed in early October due to the passage of the remnants of Hurricane Ian.

Table 1. Manufacturer-reported accuracy and resolution values for temperature, conductivity, and pressure sensors equipped on shark CTD-SRDL and Slocum glider CTDs; bracketed values for pressure readings refer to device full scale percentile values in bars.

	± Accuracy, resolution	
Sensor	SRDL tag	Slocum glider
Temperature (˚C)	± 0.005, 0.001	± 0.002, 0.001
Conductivity (S/m)	± 0.005, 0.0002	± 0.0003, 0.0001
Pressure (bar)	± 0.15% full scale [200], 0.005 bar	± 0.1% full scale, 0.002% full scale bar [100]

Glider-tag dataset preprocessing

Following tag and glider recovery for each deployment, data from each device were preprocessed as measurements from the "mixed" (March 2022) and "stratified" (September 2022) missions. Preprocessing of glider data included estimating density from glider-measured conductivity, temperature, and pressure using the UNESCO 1983 (EOS-80) equation of state. Additional preprocessing applied to datasets from both devices included identifying individual yos (defined as a time- and pressure-ordered downcast followed by an upcast) and matching concurrent yos measured by the tag and glider by nearest timestamps associated with the beginning of the leading downcast. Initial plots of pressure over time recorded by both devices indicated a nonlinear offset in the tag's pressure sensor and a temporal lag between the devices' clocks. To address the temporal lag between the device clocks, a cross-correlation function (CCF) was applied to the bottom 2 dbar of pressure time series for each matched glider-tag yo

pair, using each yo's trough as a fixed point to compare pair offset. Prior to applying the CCF, each yo pair was interpolated to a 1 Hz time series. The lag between pressure time series identified by the CCF for each tag-glider yo pair was then used as a time correction for the tag dataset. We assumed the glider time to be correct, since it receives an updated time stamp from GPS each time it surfaces (~ 2 hrs). Cross correlation analyses were conducted using the *stats* package in R. Profile lag values were then regressed as a function of hours since deployment.

After matched tag-glider yos were corrected for tag clock lag, all yos were categorized by movement phase into upcasts, downcasts, crests, and troughs. Crests and troughs were defined as the adjacent upper and lower 3 dbar intervals of measurements surrounding yo inflection points. All tag measurements were interpolated to the gliders' times, to ensure a common time stamp for all measurements. All sensor measurements across movement phases were then aggregated into 1-dbar average intervals.

Baseline SRDL tag performance

Baseline performance of the SRDL tags' sensors in the field was evaluated by calculating root mean square error (RMSE) and bias between the distributions of vertically binned (1 dbar) yo records of temperature, conductivity, and pressure data collected by the tag compared to the respective glider CTD records during each mission. Statistics were also calculated and compared between the lag-corrected and uncorrected datasets to examine the impact of clock lag on the interpretability of tag-collected data without the correction enabled by the glider's concurrent clock. RMSE and bias were calculated for each tag-glider yo pair across the duration of both missions, as well as across individual platform movement behaviors (upcast, downcast, crest, trough) within each mission. Whole-mission percent errors were also calculated from RMSE.

Ribbon plot distributions of the 1-dbar binned differences between glider and tag sensor values across pressure over time were also examined to identify spatiotemporal patterns of sensor error throughout the course of each mission. Linear mixed-effects models were used to assess the effects of clock lag correction on RMSE and bias across glider movement behaviors. Individual models were developed for each mission (mixed or stratified) and sensor measurement (pressure, temperature, or conductivity), with RMSE or bias as the response variable and the interaction between correction status (lag corrected or uncorrected) and movement behavior (upcast, downcast, crest or trough). Yo number was incorporated as a random intercept to account for repeated measurements and reduce temporal autocorrelation. Differences in RMSE and bias between corrected and uncorrected groups were then calculated using pairwise contrasts of estimated marginal means (EMM) from each model. The EMMs represent the model-predicted mean RMSE or bias for each behavior group, calculated by averaging over the distribution of values in each movement behavior and accounting for the random effect of yo number. The pairwise contrast specifically quantifies the difference between uncorrected and corrected RMSE and bias EMMs.

Impacts of water column stratification on tag performance

The potential impacts of water column stratification on tag performance were evaluated by regressing sensor RMSE and bias against differences in the mean temperature between the upper and lower 5 dbar layers. Stratification for each yo was computed as the difference between the average temperatures from 0-5 dbar and the average temperatures from the maximum depth to 5 dbar above the maximum depth. RMSE and bias statistics were calculated for glider- and

tag-measured pressures, temperatures, and conductivities per cast, and then regressed against temperature differences per yo.

Impacts of glider motion on tag performance

The potential impact of glider motion on tag performance was evaluated by regressing glider-tag sensor differences against the outputs of a calibrated dynamic glider flight model. For the mixed mission, the dynamic model was calibrated with the initial coefficients of 57.76 kg for effective glider mass in water (mg) and 0.0563 m^3 for glider volume per unit density (Vg); for the stratified mission, the dynamic model was calibrated with initial coefficients of 58.61 kg (mg) and 0.0573 m^3 (Vg), reflecting the differences in the sizes of G2 and G3 Slocum gliders. Coefficients were derived from pre-deployment ballasting. For both missions, a zero-lift drag coefficient (C_{d0}) of 0.14 was used to account for drag induced by the mounting of the tag body atop the glider (Merckelbach et al. 2019). Models for both missions also included masking depths less than 5 m and greater than 100 m to exclude near-surface behavior as well as depths beyond the maximum operational depth. Once calibrated, the dynamic models were run a second time using model-optimized coefficients for effective glider mass in water and zero-lift drag. Flight modeling was accomplished in Python using the *gliderflight* package (Merckelbach et al.). Estimated values for horizontal glider velocity ($m\text{-}s^{-1}$), vertical glider velocity ($m\text{-}s^{-1}$), incident velocity ($m\text{-}s^{-1}$), and attack angle (rad) were extracted from the flight models. Flight datasets then underwent the preprocessing procedure of identifying yos as concurrent, time- and depth-ordered downcast upcast pairs, matching yos identified in the flight datasets to lag-corrected yos identified in the tag-glider CTD datasets, approximating flight data yos to the time steps of their tag-glider CTD counterpart yos to create a joint CTD-flight variable dataset, and averaging

matched flight and CTD variables to 1 dbar depth bins. Temperature, conductivity, and pressure differences between the glider and tag sensors were then regressed linearly against individual flight variables to determine whether a coarse relationship exists between platform movement and tag sensor performance.

Results

The March, mixed condition tag-glider deployment lasted 25.1 days and traveled 489.8 km; the September, stratified condition deployment lasted 26.1 days and traveled 675.4 km. Deployment details and associated datasets are available at the Environmental Research Division Data Access Program (ERDDAP) server hosted by Rutgers University Center for Ocean Observing Leadership (RUCOOL) (https://slocum-data.marine.rutgers.edu/erddap/index.html). Stratified conditions in September were disrupted during the last week of the deployment, when the remnants of Hurricane Ian caused local mixing; the final week of the deployment (October 3rd through October 10th) was thus removed from analysis to maintain a fully stratified thermal regime in the dataset. Following data preprocessing, 4,136 glider-tag yo pairs were identified in the mixed conditions mission and 3,815 yo pairs were identified in the stratified conditions mission (Figure 2, Figure 3).

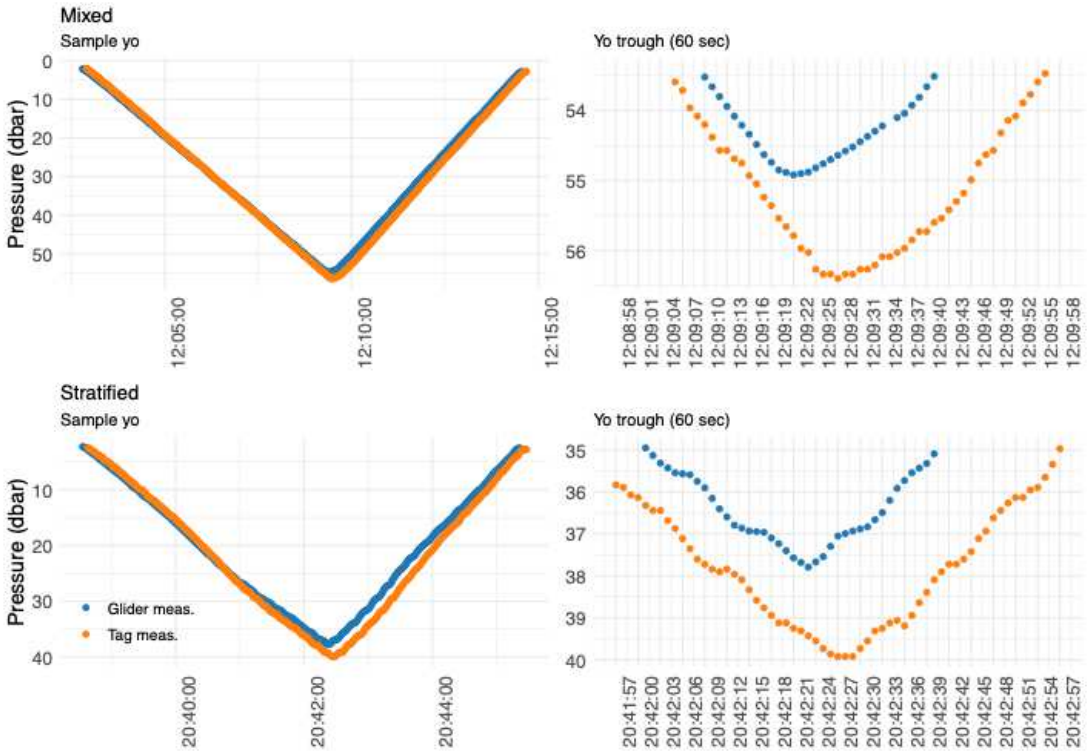

Figure 2. Sample yos taken from mixed (top row) and stratified (bottom row) missions, showing pressure values measured by the glider (blue) and tag (orange) over time. The left column plots show each yo in its entirety; the right column plots show the yos' lower inflections over the course of one minute, displaying the temporal offset in the nadir inflection of the tag and glider CTDs.

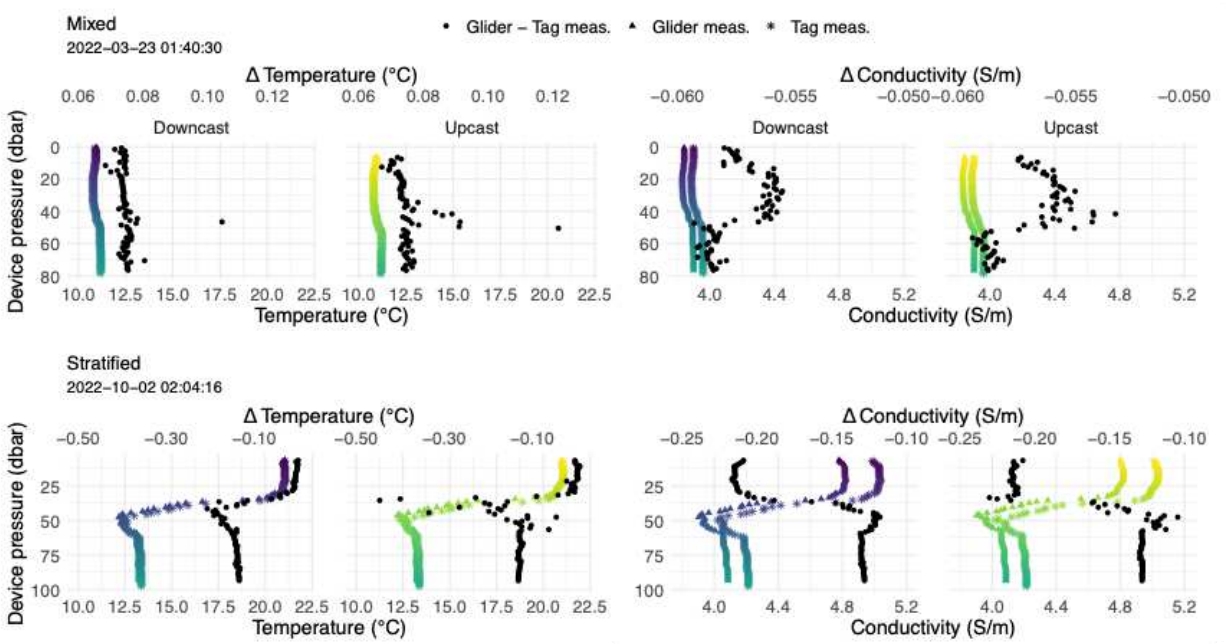

Figure 3. Distributions of glider and clock-corrected, tag-measured temperature (left column) and conductivity (right column) across the respective device's pressure measurements, for paired representative upcasts and downcasts from each mission (mixed mission casts in the top row; stratified mission casts in the bottom row); casts are colored by time, while differences between device measurements are provided in black. The upper x-axis refers to the range of device measurement differences, while the lower x-axis refers to the range of device measurements.

Application of CCF analysis to preprocessed glider-tag yo pairs identified distributions of clock lag ranging from two to eight seconds for the mixed conditions mission and from 0 to 12 seconds for the stratified conditions mission. The maximum pressure and the number of hours since deployment at which that maximum pressure occurred were also identified for each yo. The linear regression of lag (seconds) as a function of hours since the glider and tag were deployed yielded a weak but significant effect of 0.002 seconds lag per hour for the mixed mission ($p <$

2.2e-16; $R^2 = 0.143$) and a slightly larger, significant effect of 0.008 seconds lag per hour for the stratified mission ($p < 2.2e\text{-}16$; $R^2 = 0.494$) (Figure 4). In addition to the intercepts suggesting an initial offset between clock measurements, these effect sizes suggest that, for the mixed mission, there is a relatively weak expectation of a clock drift in the tag on the order of one second every 21 days; and for the stratified mission, there is a stronger expectation of tag clock drift on the order of one second every five days.

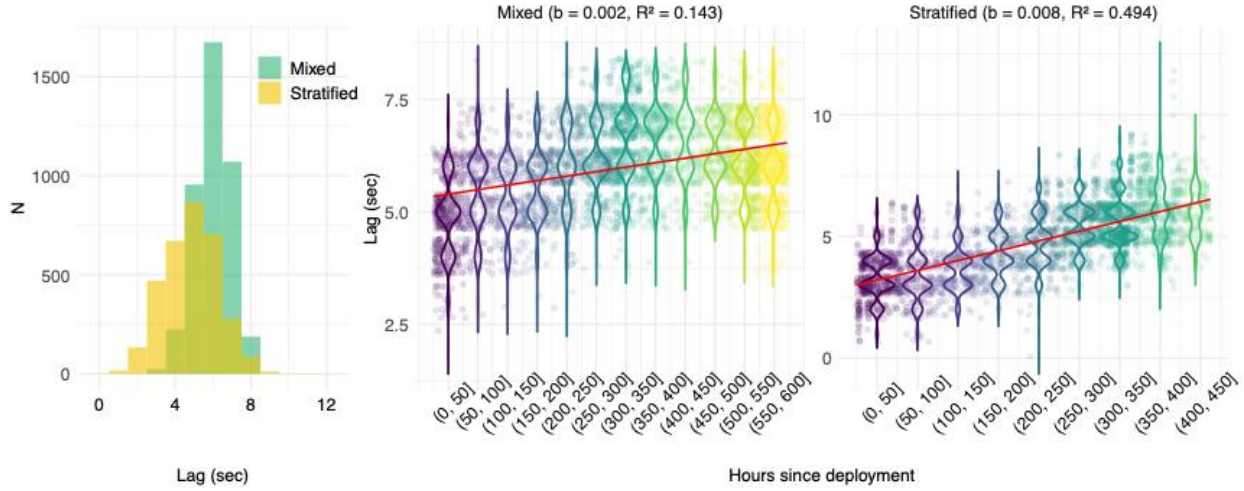

Figure 4. Left: Histograms showing frequency of lags of a given duration identified for the mixed (green) and stratified (gold) missions. Center, right: Violin distributions of lags as a function of 50-hour bins since each deployment began, overlaid with the estimated linear regression of lag as a function of hours since deployment. Effect sizes and R^2 values for each regression are provided in parentheses at the top of each scatterplot; regressions were calculated using non-binned hours.

In addition to clock offset and drift between the glider and tag, the analysis revealed a pressure offset between the glider and tag for each yo, that persevered post clock drift correction (Figure 2, Figure 3). Prior to correcting for time, distributions of differences between the tag's

reported and actual pressure were largest for the stratified (-2.08 – 7.36 dbar) versus mixed (-2.19 – 3.91 dbar) missions, across all movement phases. Applying the clock lag correction shifted modes differently for glider phases, across both missions (Figure 5). Mode differences in pressure from downcast yos in the mixed and stratified missions shifted from -0.06 dbar and -0.19 dbar to 0.94 dbar and 0.72 dbar, respectively, when comparing lag uncorrected and corrected data. Mode pressure differences for upcast yos showed the opposite response, shifting from 1.76 dbar and 1.79 dbar in the mixed and stratified uncorrected distributions to 0.6 dbar and 0.99 dbar in the corrected distributions. Regressions of pressure offset as a function of depth across lag correction status, movement phase, and mission revealed consistent associations between offset, lag, and actual pressure in all cases. For both mixed and stratified missions, applying the clock correction the regression of pressure difference as a function of true pressure resulted in smaller effect sizes, while intercept magnitudes increased for mixed downcasts but decreased for mixed upcasts and all stratified yos (Table 2). Effect sizes (β) were larger for the stratified mission, and did not vary between upcast and downcast, except for a slight increase for stratified, uncorrected downcast.

Table 2. Coefficient values and 95% confidence interval ranges for regressions of pressure differences as a function of true pressure, calculated across mission, cast type, and clock lag correction status.

			Effect size (dbar)		Intercept (dbar)	
Mission	Cast	Lag correction status	Value	95% CI	Value	95% CI
Mixed	Downcast	Uncorrected	0.0273	0.0272 – 0.0274	-0.457	-0.459 – -0.454
		Corrected	0.0261	0.026 – 0.0261	0.413	0.41 – 0.415
	Upcast	Uncorrected	0.0262	0.0261 – 0.0263	1.29	1.29 – 1.3
		Corrected	0.0255	0.0254 – 0.0256	0.263	0.26 – 0.265
Stratified	Downcast	Uncorrected	0.0658	0.0656 – 0.066	-1.12	-1.12 – -1.11
		Corrected	0.0597	0.0596 – 0.0599	-0.034	-0.039 – -0.029

	Upcast	Uncorrected	0.06	0.0601 – 0.06	1.36	1.36 – 1.37
		Corrected	0.0592	0.0591 – 0.0594	0.367	0.361 – 0.373

Regressions specifically examining differences in pressure measurements over actual pressure between mixed and stratified missions identified significant lower baseline offsets and larger effect sizes between stratified and mixed intercepts for uncorrected downcasts (a = -0.73, β = 0.036) and uncorrected upcasts (a = 0.016, β = 0.032). After applying the clock correction, corrected downcast intercept differences decreased in magnitude while effect size decreased marginally between stratified and mixed condition missions (a = -0.48, β = 0.031); differences in coefficients between conditions changed little for corrected upcasts (a = -0.017, β = 0.032). Thus, pressure data can be corrected, following clock lag corrections, by removing a < 1 dbar pressure-dependent offset that yields a 3-5 dbar pressure correction over 100 m.

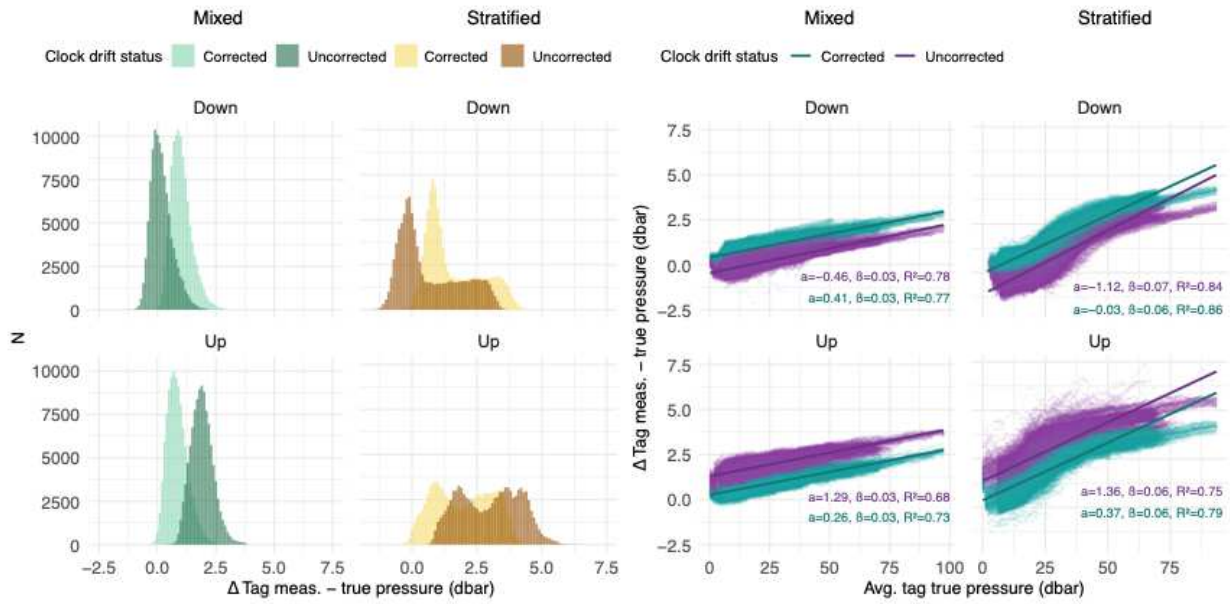

Figure 5. Left: Histograms of pressure offset between the tag's reported pressure and actual pressure relative to the glider; darker versus lighter colors denote datasets where the clock lag correction has not been applied versus where it has; green versus gold colors denote the mixed versus stratified mission. Right: Scatterplots and linear regression lines

for pressure offset as a function of the tag's actual pressure relative to its position on the glider; color denotes correction status; a, β, and R^2 report each regression's intercept, effect size, and fraction of variance explained.

Baseline SRDL tag performance

RMSE and bias values calculated with the lag-corrected yo pairs over mission duration showed that overall, the tag performed with an error of an order of magnitude higher in stratified conditions than in mixed conditions for temperature and conductivity measurements (Table 3). RMSE for pressure measurements doubled in stratified conditions. For the mixed mission, the tag measured temperature biased cooler than the glider (0.066 ˚C); pressure deeper than the glider (-0.912 dbar); and conductivities larger (-0.033 $S \cdot m^{-1}$). For the stratified mission, the biases for pressure and conductivity were still negative, albeit larger in value (-1.71 dbar and -0.162 $S \cdot m^{-1}$), and the temperature bias showed the tag measuring temperatures warmer than the glider (-0.11 ˚C). Percent errors calculated across sensors based on whole-mission RMSE values reveal that for both missions, pressure error was highest (4.58% for mixed conditions, 7.79% for stratified conditions), followed by conductivity (1.27% mixed, 3.97% stratified) then temperature (0.82% mixed, 1.68% stratified).

Table 3. RMSE, bias, standard deviation (SD), and percent error calculated across mixed and stratified missions, between glider and tag-measured pressure, temperature, and conductivity. Note that the tag and glider CTDs measured from platform sensors 0.216 m apart. Resolution and accuracy from manufacturer-reported specifications are provided for two common animal tags, along the CTD SRDL tag.

405 [Table 3 in-text location; Table 3 can be found at the end of the document, landscape orientation]

406 Distributions of RMSE and bias, calculated per movement phase (crest, downcaast,
407 trough, upcast), for both lag-corrected and uncorrected yo pairs, shows that both movement
408 phase and clock drift correction impact tag sensor performance, specifically during the stratified
409 mission. For the mixed mission, post-correction EMM RMSE differences were largest for
410 pressure, and comparatively modest for conductivity and temperature measurements across
411 movement phases (Figure 6). Correcting mixed mission pressure led to a 0.99 dbar decrease in
412 upcast EMM RMSE, a 0.56 dbar increase in downcast EMM RMSE, and a 0.5 dbar, and a 0.1
413 dbar decrease for crest and troughs. Clock corrections for mixed mission conductivity
414 measurements significantly reduced EMM RMSE by -0.001 $S \cdot m^{-1}$for downcasts, increased
415 EMM RMSE by 0.001 $S \cdot m^{-1}$ for upcasts, and increased EMM RMSE by 0.003 $S \cdot m^{-1}$ for crests.
416 For mixed mission temperature measurements, EMM RMSE differences post-corrections were
417 all positive, significant, and relatively small in magnitude (0.018 ˚C for crests to 0.006 ˚C for
418 upcasts and downcasts). Stratified mission RMSE ranges per movement phases followed similar
419 patterns with the lag corrected and uncorrected pressure measurements, albeit over a slightly
420 higher range of error values. Applying the time correction reduced RMSE median values for all
421 stratified mission movement phases substantially, as well as reduced the variation in RMSE.
422 Temperature RMSE ranges were largest and most variable for uncorrected downcasts, with
423 EMMs decreasing by 0.37 ˚C post correction, followed by ranges for uncorrected upcasts, with
424 post-correction EMMs decreasing by 0.09 ˚C. Changes to EMM RMSE for stratified
425 conductivity measurements were largest for corrected downcasts (decreased by 0.029 $S \cdot m^{-1}$) and
426 upcasts (increased by 0.021 $S \cdot m^{-1}$). All stratified mission RMSE contrasts between corrected and

uncorrected behavior groups were statistically significant, except those for temperature and conductivity measurements taken during crests.

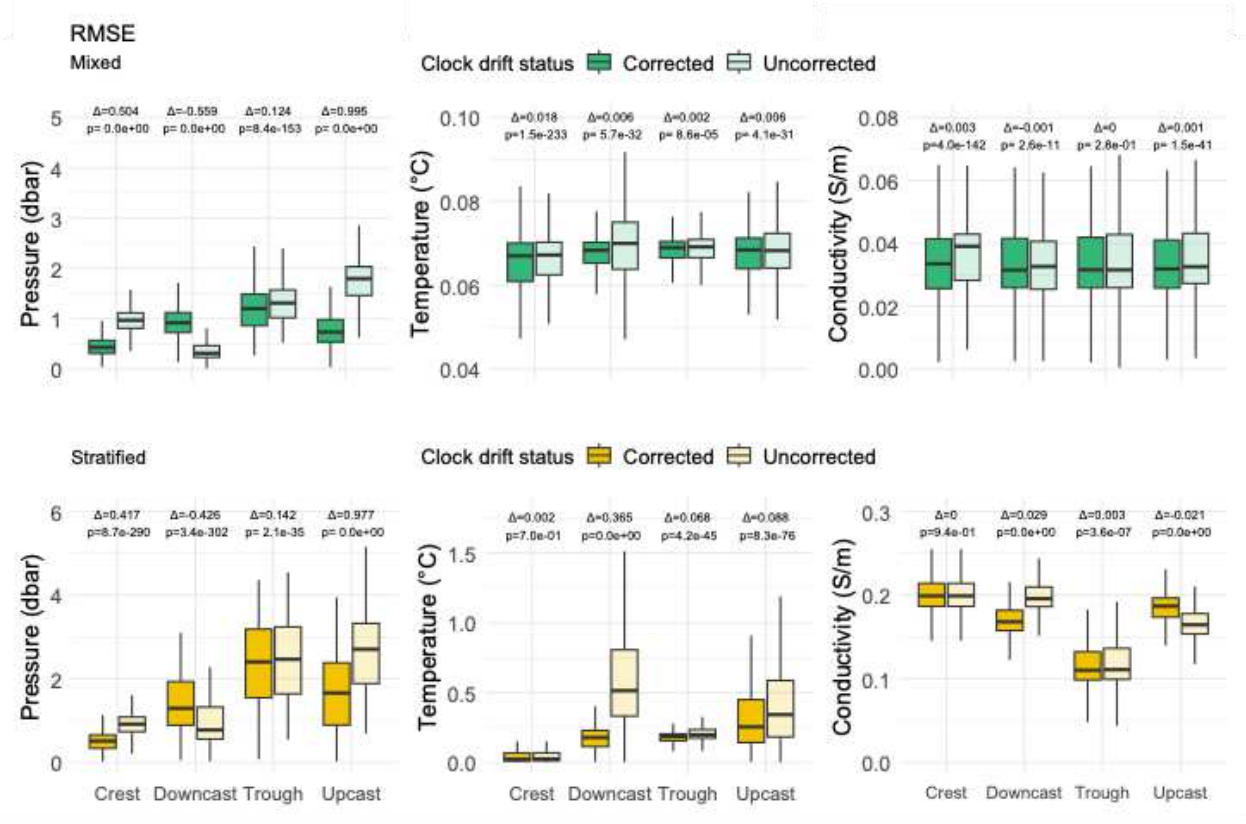

Figure 6. Boxplots of RMSE values calculated per cast, across mission (row, color), clock lag correction status (shade), and glider movement phase. Outliers were omitted from plotting. Estimated marginal mean (EMM) differences (Δ) between modeled uncorrected and corrected means are represented above each glider movement phase, along with p-values produced by each model.

Distributions of movement phase-specific bias calculated for corrected and uncorrected mission data followed a similar pattern of variation range and median size, with mixed mission distributions spanning substantially smaller ranges and with phase-effects most pronounced in

439 the pressure measurements (Figure 7). Like RMSE distributions, downcast and upcast bias
440 EMMs for both missions showed the largest shift with corrections applied, across all sensors in
441 the stratified mission, most broadly with pressure in the mixed mission. Temperature EMMs
442 calculated from the mixed mission showed lag corrected groups had lower RMSE by 0.001-
443 0.003 ˚C; differences were significant for all groups except mixed mission troughs. Conductivity
444 EMMs from the mixed mission were relatively small and significant for all groups. Mixed
445 mission pressure EMMs were largest between correction groups and statistically significant
446 across all movement behaviors, ranging from -1.02 dbar for upcasts (indicating an increase in
447 RMSE post clock correction) to 0.88 dbar for downcasts (indicated a decrease in RMSE post
448 clock correction).
449
450 Stratified mission bias distributions prior to applying the clock lag correction were
451 broadly negative across movement behaviors for all three sensor measurements, with the post-
452 correction shifting nearly all groups closer to zero. One exception to this occurs with pressure
453 bias distributions for downcasts, where an EMM of 0.99 dbars signifies an increase in bias
454 magnitude post correction, and with pressure bias distributions for upcasts, where an EMM of -
455 1.01 signifies a decrease in bias magnitude post correction. Temperature bias was larger and
456 negative for troughs versus crests, with relatively small, statistically significant EMMs indicating
457 that clock correction had little impact in reducing bias within these groups. Temperature bias for
458 downcasts improved by -0.23 ˚C but worsened by -0.27 ˚C for upcasts. Conductivity biases in
459 the stratified mission followed a similar pattern as temperature biases for upcasts and downcasts,
460 with corrected downcast improving -0.025 $S \cdot m^{-1}$ and upcasts worsening by 0.029 $S \cdot m^{-1}$. EMM
461 contrasts for conductivity bias were minimal for crests (0 $S \cdot m^{-1}$) and troughs (-0.001 $S \cdot m^{-1}$)

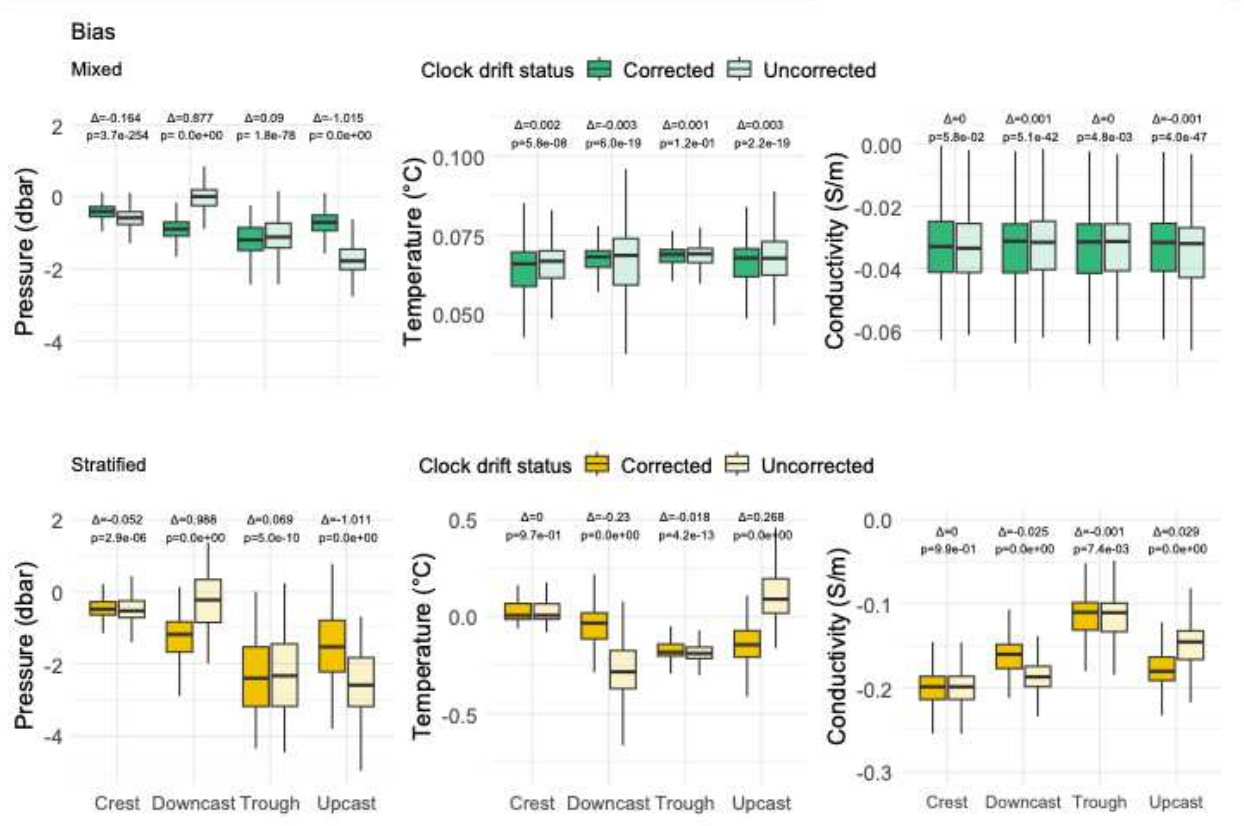

Figure 7. Boxplots of bias values calculated per cast, across mission (row, color), clock lag correction status (shade), and glider movement phase. Outliers were omitted from plotting. Estimated marginal mean (EMM) differences (Δ) between modeled uncorrected and corrected means are represented above each glider movement phase, along with p-values produced by each model.

Ribbon plots of differences in glider and tag measured temperature and conductivity for both missions emphasize the impact of seasonal stratification and platform movement phase on the tag's performance. Plots of the difference in glider-tag temperatures show that, during mixed conditions, the glider consistently measured temperatures slightly warmer than the tag. During stratified conditions, however, the glider measured consistently warmer than the tag above the

thermocline and cooler below the thermocline (Figure 8). Both the positive and negative temperature differences appear greatest within meters of the thermocline but conserve their sign (despite lessening in magnitude) on either side of the cline. Differences appear largest overall for upcasts, with depths immediately below the thermocline measuring up to one degree Celsius cooler.

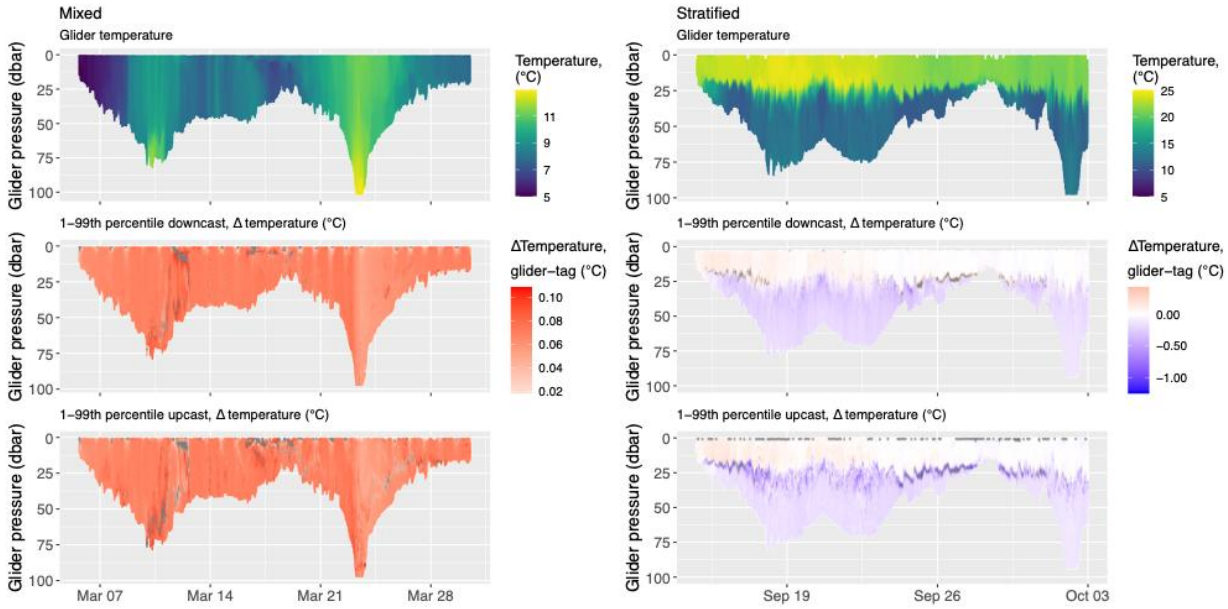

Figure 8. Ribbon plots of glider-measured temperatures (top row), 1st-99th percentiles of temperature differences between the glider and tag during downcasts (middle row), and 1st-99th percentile temperature differences between the glider and tag during upcasts (bottom row) for mixed (left column) and stratified (right column) missions; glider-measured times and pressures are reported on the x- and y-axes, respectively, and colors report glider-measured temperatures in the top row and temperature differences in the middle and bottom rows; grey colors indicate measurements > 99th percentile values. Note that the well-mixed conditions due to the passage of Ian occurring after early October during the stratified mission were removed from this analysis.

Plots of differences in conductivity measurements reflect broader patterns in glider-measured conductivity (Figure 9). Throughout both missions, the glider consistently measured lower conductivity than the tag did. While there appears to be an expected dependency to local water temperatures, likely caused by thermal lag in the conductivity cell, there does not seem to be a movement- or time-related pattern in the tag's conductivity measurements deviating from those of the glider. Ribbon plots of pressure differences over time and depth reflect a consistent pattern of the glider measuring lower pressures than the tag, exacerbated with increased local depths, and higher during stratified conditions, particularly below the thermocline during upcasts.

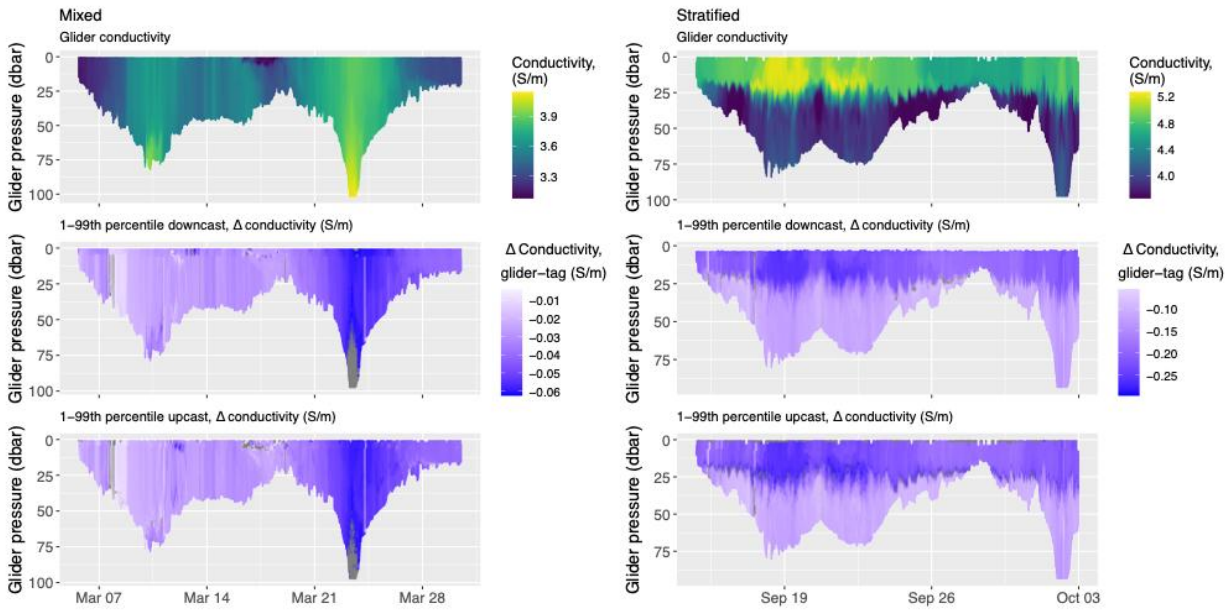

Figure 9. Ribbon plots of glider-measured conductivities (top row), 1st-99th percentiles of conductivity differences between the glider and tag during downcasts (middle row), and 1st-99th percentile conductivity differences between the glider and tag during upcasts (bottom row) for mixed (left column) and stratified (right column) missions; glider-measured times

and pressures are reported on the x- and y-axes, respectively, and colors report glider-measured conductivities in the top row and conductivity differences in the middle and bottom rows; grey colors indicate measurements > 99th percentile values. Note that the well-mixed conditions due to the passage of Ian occurring after early October during the stratified mission were removed from this analysis.

Impacts of stratification on tag performance

Glider and tag pressure, temperature, and conductivity sensor RMSE generally showed a positive relationship with water column temperature difference, while sensor bias generally showed a negative relationship (Figure 10). Note that the average maximum yo depth was 35.8 m for the mixed mission and 41.3 m for the stratified mission. Across all sensor types, the effect sizes were larger for RMSE calculated from upcasts than downcasts. Effect sizes and R^2 values were largest for pressure RMSE, with β values of 0.143 and 0.190 dbar increases in RMSE per degree temperature difference between the upper and lower water column, for downcasts and upcasts respectively. Temperature RMSE effect sizes for downcast and upcasts were approximately 0.017 and 0.024 °C. Conductivity RMSE effect sizes were negative with increased estimated temperature difference, approximately -2.1e-3 and -1.1e-3 S-m^{-1} per degree temperature difference in the upper and lower water column, for downcasts and upcasts. Regressions of bias between devices as a function of estimated stratification also generally produced larger effect sizes for upcasts than downcasts, but reversed directions across sensor types. For bias between pressure sensors, effect sizes were approximately -0.12 and -1.8 dbar, for downcasts and upcasts, per degree difference; for bias between temperature sensors, respective

effect sizes were -8.00e-3 and -1.15e-2 °C; and for bias between conductivity sensors, effect sizes were 3.02e-3 and 2.34e-3 S-m^{-1}.

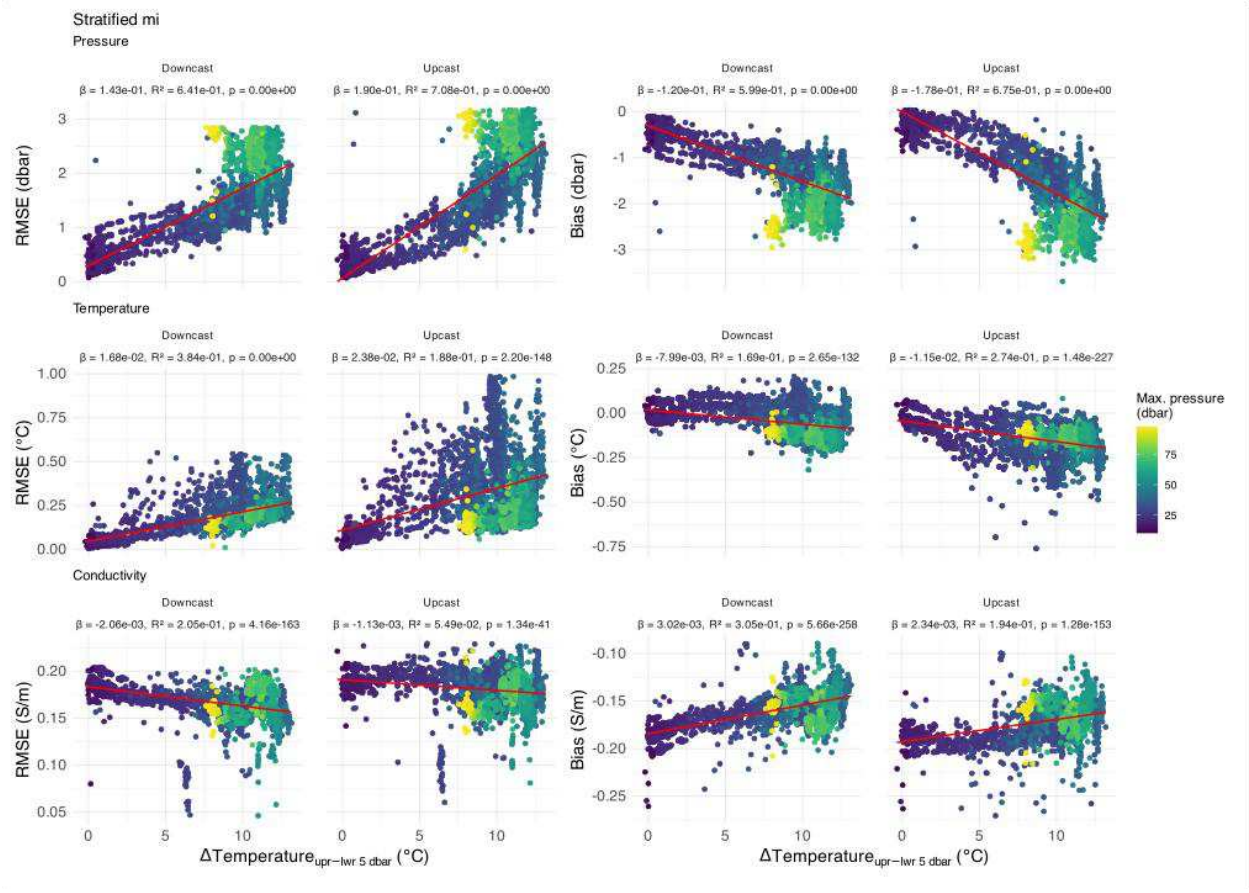

Figure 10. Distributions of RMSE (left column) and bias (right column) values calculated across individual upcasts and downcasts against the casts' associated differences in mean upper and lower five dbar temperatures for the stratified mission; point color refers to the maximum pressure associated with a given cast; the linear regression line is presented in red; β, R^2, and p report each regression's effect size, fraction of variance explained, and p-value; regressions for RMSE and bias calculated for glider-tag pressure measurements are presented in the top row, temperature measurements in the middle row, and conductivity measurements in the bottom row; note that values greater than the 99th percentile of each

yo group were defined as outliers and omitted from plotting for visualization purposes, though kept in regression calculations.

Impacts of platform motion on tag performance

Initial runs of both missions' dynamic models with manually selected coefficients calibrated to *mg* values of 57.78 kg and 58.71 kg, for the mixed and stratified missions, respectively, and C_{d0} values calibrated to 0.14 and 0.12, respectively. Distributions of 1 dbar averaged horizontal glider velocity, vertical glider velocity, and incident velocity fell within anticipated ranges. A small quantity of modeled average angle of attack values (n=128) calibrated to anomalously high values (mean=75.2˚) and were excluded from subsequent analyses. Distributions and linear regressions of glider-tag sensor measurement differences (Δ temperature and Δ conductivity) showed that variations in platform movement had greater impacts on measurements taken under stratified conditions, and that those impacts themselves were of moderate, albeit statistically significant, strength. Regressions of pressure differences against flight variables for both missions reflect the impacts of vertical motion rather than sensor deviations and are disregarded.

For both missions, all regressions were statistically significant, however many were characterized by very small effect sizes and percentages of variance explained. Despite their statistical significance, all regressions of temperature and conductivity differences against flight variables for the mixed mission yielded small magnitude effect sizes (≤ -4.63e-2 °C and -1.72e-2 ≤ $S\text{-}m^{-1}$) and negligible R^2 values (≤ 2.49e-3 and ≤ 1.13e-4, for all temperature and conductivity regressions, respectively). Regressions of temperature and conductivity differences on incident velocity, taken for the stratified mission, yielded substantially larger effect sizes and more

meaningful percentages of variance explained (Figure 11). The regression of temperature differences against averaged modeled incident velocity explained 23% variance, with a 1.71 °C increase in difference per m/s increase in velocity. The regression of conductivity differences against average modeled incident velocity explained 24% variance with a 0.46 decrease per m/s increase in velocity. These proportions of unexplained variance are likely related to the different positions of the glider and tag CTD intake ports on the glider, with each port thus experiencing unique hydrographic conditions.

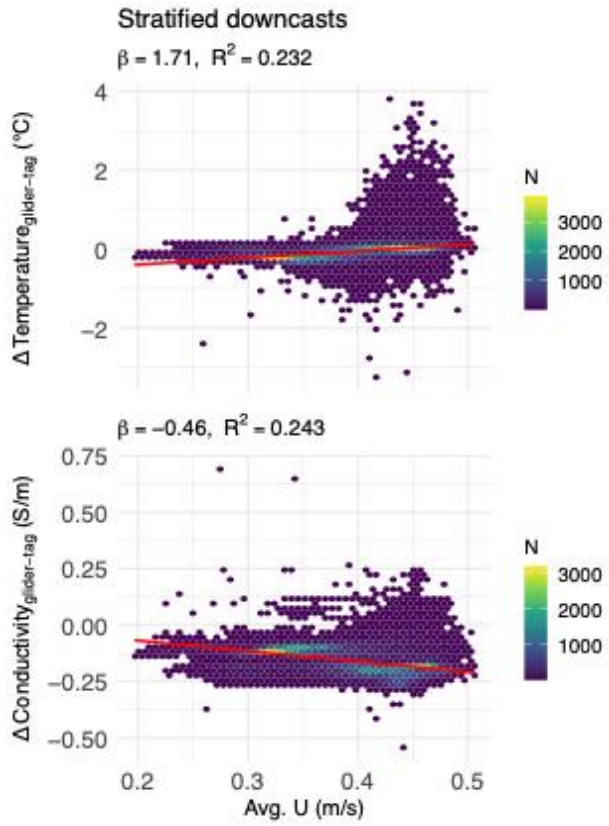

Figure 11. Distributions of differences in glider and tag temperature (top panel) and conductivity (bottom panel) measurements as a function of average modeled incident velocity (U) for the downcasts of the stratified mission; hexagons are colored by density; the linear

regression is represented by the red line, with effect size (β) and percentage of variance explained (R^2) provided at the top of each plot.

Discussion

This study characterizes the field performance of a CTD-SRDL tag in mixed and stratified conditions, using Slocum gliders as a controlled moving tag platform and sensor baseline for comparison to an established OOP. Our results quantify error and bias for tag measurements across a variety of conditions, movement phases, and corrections. Highest error and bias, relative to the glider's pumped CTD, occur in stratified conditions, neglecting clock drift. For pressure measurements under these conditions, highest error and bias occurred during upcasts; for temperature, during downcasts, and for conductivity, error was highest in crests and downcasts. When corrections for clock drift are applied, during mixed conditions, error and bias are substantially smaller across all sensors and movement behaviors. Corrected values for stratified conditions behave more variably across sensors and movement type, with pressure error and bias decreasing for upcasts but increasing for downcasts. Temperature and conductivity error and bias, on the other hand, worsened slightly for upcasts and improved for downcasts, likely related to the dependence of conductivity measurements on ambient temperature and thermal inertia. Regarding pressure, the deviations between the tag's measured position versus it's true position (estimated from the distance between the tag's CTD port and the glider's CTD port) are within manufacturer accuracy (3 dbar) for the mixed mission (~0-3 dbar) but exceed it for the stratified mission (~0-6 dbar). Regressions of pressure difference as a function of true pressure for both missions show an increase in bias with increased ambient pressure, that cannot be solely attributed to the tag's pressure sensor specifications alone. Rather, the regressions of

pressure difference taken for both missions show that correcting clock drift does reduce pressure bias, demonstrated by a drop in the regression intercept. Both regressions also demonstrate an effect size that is conserved between clock-corrected and uncorrected datasets, with an increase in pressure difference of ~0.03 dbar per dbar increase in true pressure under mixed conditions, and an increase in pressure difference of ~0.06 dbar per dbar increase in true pressure under stratified conditions. This demonstrates that while applying the clock lag correction reduces pressure bias, a 0.03-0.06 dbar increase in bias with increased true pressure can be expected over the course of tag deployment.

This study also quantifies the impacts of stratification and platform motion on tag measurement performance, given that the longer-term application of CTD-SRDL tags on free-swimming marine animals will likely involve higher variability in OOP speed and angle of attack. Given that many OOP applications are interested in mapping water column heat structures, among other features, how this variability in motion might interact with stratification (and thus measurements might be impacted by thermal inertia of the tag itself) is of great interest. Pressure and temperature error increased significantly with larger values of temperature differences between the upper and lower water column, with negative bias values reflecting the tag's lag in recording warmer temperatures than the glider as it moves through a more stratified upcast or downcast. For the tag's conductivity measurements, error and bias showed respective increases and decreases with increased thermal differences per profile, albeit at a much smaller magnitude. Lastly, potential impacts of glider motion to tag performance were relatively limited. For the mixed mission, all glider velocity components (horizontal, vertical, and incident) and attack angle explained minimal variance (> 20%) when regressed against all glider – tag sensor differences, for upcasts and downcasts. For the stratified mission, incident velocity demonstrated

620 moderate (23-24%) impact on temperature and conductivity biases for downcasts, with relatively
621 small but still consequential effect sizes (an increase in temperature bias of 1.7 °C and decrease
622 in conductivity bias of ~ -0.5 $S \cdot m^{-1}$, per 1 m/s increased in average incident velocity.)
623 Collectively, these results demonstrate that while in situ deployments of the CTD-SRDL
624 tag on a moving platform in the field deviate slightly but significantly from the resolution and
625 accuracy of both tag bench tests and the specifications of an established OOP platform, the
626 impacts of clock lag, platform motion, and thermal inertia can be estimated and satisfactorily
627 corrected for post deployment. Depending on the presence of predominantly mixed or thermally
628 stratified conditions, an estimated clock correction on the order of 0.002-0.008 seconds per hour
629 post deployment substantially improves sensor performance. Similar corrections for pressure
630 drift with increased ambient pressure can be expected and applied—specifically in thermally
631 mixed conditions, when (in this case) differences between measured and true tag pressure
632 exceeded the accuracy limits of the tag's pressure sensor, suggesting mixed effects of
633 temperature and depth. Quantified ranges in RMSE and bias across glider movement phase,
634 clock lag correction status, and mission condition provide margins of anticipated data product
635 resolution uncertainty. While these margins are relatively small during mixed conditions, the
636 increase in the range of RMSE and bias across movement phase during stratified conditions have
637 implications for the quality of data products one might expect if seeking to resolve thermally
638 stratified structures, such as fronts or eddies, and whether the tag is sampling upcast or downcast
639 behaviors.
640 These RMSE and bias ranges are probably lower estimates, given that the glider's speed
641 and attack angle are likely more consistent than what the tag might encounter when attached to a
642 free-swimming animal. One major limitation of this study is the lack of increased variability in

643 speed and platform motion. While there is evidence that the glider's incident velocity had an
644 impact on conductivity and temperature sensor error under stratified conditions, the still
645 relatively low R^2 (>25%) and the absence of an effect detected from regression with the mixed
646 condition flight model variables suggests that the interaction between stratification and platform
647 motion merits more rigorous investigation than the coarse field estimate ranges of impact we
648 illustrate here. Previous studies have modeled and applied post-hoc corrections for spikes and
649 anomalies in salinity profiles collected from marine mammals deployed with CTD-SRDL tags,
650 as well as from CTD-SRDL tags cast concurrently with a reference CTD (Nakanowatari et al.
651 2017; Mensah et al. 2018). Both studies have acknowledged the joint impacts of platform speed
652 and location stratification intensity and are still able to provide high-resolution spike corrections
653 to individual profiles. Although this study applies broader scale investigation over the course of
654 multi-week deployments, it would be valuable to examine short- and long-term impacts of more
655 variable speeds through stratified environments to jointly model the impacts of irregular speed
656 and thermal inertia on both measured temperature and conductivity, as well as derived salinity
657 and density.

658 The results of this study have implications for the post-processing and utility of CTD-
659 SRDL data collected by marine animal OOPs for the GOOS. First, the presence of an impact of
660 clock drift on CTD sensor performance over time suggests that the time series of profile data
661 collected over longer deployments will likely degrade over time but may be corrected in the
662 absence of a reference CTD time series with a regularly updating timestamp, such as that on
663 Slocum gliders. Second, improvements in pressure error during upcasts but worsening during
664 downcasts following the application of a clock correction imply, at shallow depths (<100 m), a
665 nonlinear relationship between the pressure sensor and depth, despite the relatively high R^2

values provided by the linear regressions in this study. While the CTD-SRDL tags are programmed to detect the nadir of a dive (or inflection in pressure measurements) and turn on the CTD after an ascent of a certain pressure change, the presence of individual or multiple inflections (or "wiggles") over a short period of time may inflate errors in the pressure response, particularly in shallow environments such as coastal shelf systems. Lastly, the apparent range in biases for pressure, temperature, and conductivity measurements across mission conditions suggests that in highly stratified conditions, particularly in shallow environments, tag performance is not consistent. While these inflated ranges in conductivity readings may result from the presence of spikes in the profile caused by thermal lag in the conductivity cell, interquartile ranges for temperature measurements recorded during stratified downcasts and upcasts are ~ 0.25 °C for clock lag-corrected casts and ~ 0.5°C for uncorrected casts. This effect, which shifts the tag's in situ resolution, may have implications for how the data can be utilized by the GOOS.

The GOOS supplies near-real time oceanographic data for a myriad of applications and scales. This study quantifies how CTD-SRDL tag-measured oceanographic data compares to those datasets collected by a platform that are regularly and reliably integrated into the GOOS, and the subsequent impacts of tag error and bias can be coarsely estimated from our results. For example, two of the many applications of GOOS data are the calculations of global ocean heat content (OHC) and of practical salinity. Global OHC is a critical component of Earth's energy budget, which in turn forms the backbone of measuring and predicting the planet's climate (Meyssignac et al. 2019; Abraham et al. 2013). The accuracy of global OHC measurements has increased significantly with the development of CTD profiling platforms, such as Argos floats, gliders, CTD-SRDL tagged elephant seals, and mooring buoys (Meyssignac et al 2019; Abraham

et al. 2013). The practical salinity scale, on the other hand, is the ratio of seawater conductivity at 15 °C to the conductivity of a potassium chloride standard at the same temperature and facilitates the calculation of comparable salinity levels from conductivity, temperature, and pressure measurements (Perkin & Lewis 1980; Lewis 1982). Accurate calculations of salinity are important for understanding local and global processes with physical and ecological implications, ranging from the habitat quality of estuarine salt wedges to coastal stratification to global ocean overturning and atmospheric circulation (Woody et al. 2000). Both variables represent important applications of GOOS data, scaling individual CTD profiles to global processes. While both types of measurements depend on and are inherently shaped by the diversity of platforms that supply water column temperature data, the impact of increased CTD-SRDL data contributions to OHC and the practical salinity scale can be estimated using the in-situ temperature sensor bias values calculated in this study.

If one were to calculate an estimated, local OHC ($J \cdot m^{-2}$) for the upper 95 m sampled by the glider during the mixed and stratified mission, where:

$$OHC = \rho c_p \int_0^{95} (T(z) - T_{ref}) dz$$

Here, OHC ($J \cdot m^{-2}$) is estimated as the integrated temperature (° C) at pressure-depth (where 1dbar pressure estimates to 1 m depth), relative to a reference temperature (°C), for pressure from surface to 95 dbar. Assuming a heat capacity of seawater (c_p) of 3985 $J \cdot (kg \cdot K)^{-1}$, and densities ($\rho$) of 1026.15 $kg \cdot m^{-3}$ and 1024.15 $kg \cdot m^{-3}$ (mean glider-derived densities for the mixed and stratified missions, respectively), one could calculate glider- and tag-derived OHC estimates and observe that the subsequent calculations in OHC follow illustrated relationships in statistical bias between upcast vs. downcast, stratification, and depth. Differences in glider-tag

711 OHC are ultimately conserved across mission and cast type. OHC values calculated from upcasts
712 and downcasts from the mixed mission, and from stratified upcasts all show an estimated 0.8%
713 median difference between glider- and tag-derived OHC, regardless of clock correction.
714 Stratified downcasts show an even lower median difference of 0.2% for uncorrected and
715 corrected datasets.

716 Similarly, glider and tag CTD values can be used to calculate unitless practical salinity
717 from conductivity (mS/cm), temperature (°C), and pressure (dbar). When these conversions are
718 applied to measurements from both devices and compared across mission and cast type,
719 estimated practical salinity values again follow patterns of statistical bias identified in this study.
720 Median tag-derived practical salinities are approximately 1.2% higher than median glider-
721 derived salinities for them mixed mission, across cast types and whether the data was corrected
722 for clock lag or not. Stratified mission salinities are higher, with median tag-derived salinities are
723 approximately 3.7% and 3.8% higher than glider derivatives, regardless of clock correction, for
724 both downcasts and upcasts, respectively. While these exercises are limited to a relatively small
725 dataset, the conservation of an incredibly small, translated impact of transmitted data bias and
726 RMSE between tag and glider to an operational data product is promising.

727

728 **Conclusions**

729 This study demonstrates significant impacts of clock lag, stratification, and depth, as well
730 as the meaningful conservation of sensor accuracy across platform movement for data products
731 collected by an in situ CTD-SRDL on a movement platform. Sources and magnitudes of bias are
732 identified, with post-processing estimated correction values similarly quantified and proposed for
733 the effects of clock lag, depth, and stratification across sensors. While tag error and bias relative

to glider sensor measurements are substantially higher in magnitude and range during stratified conditions, radiating impacts are estimated to be minimal, when applied in operational settings such as OHC and salinity calculation. Scaling from a single deployed CTD-SRDL to a larger sample size will likely increase percent differences from glider-derived metrics, and this warrants additional scaled research.

List of abbreviations

CCF	Cross-correlation function
CTD	Conductivity, temperature, depth
EMM	Estimated marginal means
EOV	Essential oceanographic variables
ERDDAP	Environmental Research Division Data Access Program
GOOS	Global Ocean Observing System
GTS	Global telecommunication system
OHC	Ocean heat content
OOP	Ocean-observing platform
RMSE	Root mean square error
RUCOOL	Rutgers University Center for Ocean Observing Leadership
SD	Standard deviation
SRDL	Satellite-relayed data logger

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

Glider deployment details and datasets are available at the Environmental Research Division Data Access Program (ERDDAP) server housed with Rutgers University Center for Ocean Observing Leadership (RUCOOL) (https://slocum-data.marine.rutgers.edu/erddap/index.html). CTD-SRDL tag data will be shared on reasonable request to the corresponding author.

Competing interests

The authors declare that they have no competing interests.

Funding

Funding for this study was provided by the National Oceanographic Partnership Program (NOAA Award – NA20NOS0120221), as well as through the Mid-Atlantic Regional Association Coastal Ocean Observation System (NOAA Award – NA21NOS0120096).

Authors' contributions

MJO and ABC conceptualized the use of Slocum gliders to validate CTD-SRDL tag performance and acquired funds. CJW deployed tags and collected data piloting glider missions,

780 with supervision from MJO. CJW performed statistical analysis, with supervision from MJO and
781 CFM. CJW developed glider flight models, with supervision from JCG. CJW wrote original
782 manuscript draft, with review and editing provided by all authors.
783
784 Acknowledgments
785 Not applicable.
786
787 **References**
788 McMahon CR, Roquet F, Baudel S, Belbeoch M, Bestley S, Blight C, Boehme L, Carse F, Costa
789 DP, Fedak MA, Guinet C. Animal borne ocean sensors–AniBOS–An essential component
790 of the global ocean observing system. Frontiers in Marine Science. 2021 Nov
791 5;8:751840.
792 Harcourt R, Sequeira AM, Zhang X, Roquet F, Komatsu K, Heupel M, McMahon C, Whoriskey
793 F, Meekan M, Carroll G, Brodie S. Animal-borne telemetry: an integral component of the
794 ocean observing toolkit. Frontiers in Marine Science. 2019 Jun 26;6:436861.
795 Block BA, Holbrook CM, Simmons SE, Holland KN, Ault JS, Costa DP, Mate BR, Seitz AC,
796 Arendt MD, Payne JC, Mahmoudi B. Toward a national animal telemetry network for
797 aquatic observations in the United States. Animal Biotelemetry. 2016 Feb 3;4(1):6.
798 March D, Boehme L, Tintoré J, Vélez-Belchi PJ, Godley BJ. Towards the integration of animal-
799 borne instruments into global ocean observing systems. Global change biology. 2020
800 Feb;26(2):586-96.
801 Charrassin JB, Hindell M, Rintoul SR, Roquet F, Sokolov S, Biuw M, Costa D, Boehme L,
802 Lovell P, Coleman R, Timmermann R. Southern Ocean frontal structure and sea-ice

formation rates revealed by elephant seals. Proceedings of the National Academy of Sciences. 2008 Aug 19;105(33):11634-9.

Patel SH, Barco SG, Crowe LM, Manning JP, Matzen E, Smolowitz RJ, Haas HL. Loggerhead turtles are good ocean-observers in stratified mid-latitude regions. Estuarine, Coastal and Shelf Science. 2018 Nov 30;213:128-36.

Bousquet O, Dalleau M, Bocquet M, Gaspar P, Bielli S, Ciccione S, Remy E, Vidard A. Sea turtles for ocean research and monitoring: overview and initial results of the STORM project in the Southwest Indian Ocean. Frontiers in Marine Science. 2020 Oct 29;7:594080.

Wiernicki CJ, Curtis TH, Block BA, Shivji MS, Vaudo JJ, Wetherbee BM, Holland KN, Pinti J, Oliver MJ, Carlisle AB. Considerations for using sharks as ocean observing platforms. ICES Journal of Marine Science. 2025 Feb;82(2):fsaf011.

Andrzejaczek S, Lucas TC, Goodman MC, Hussey NE, Armstrong AJ, Carlisle A, Coffey DM, Gleiss AC, Huveneers C, Jacoby DM, Meekan MG. Diving into the vertical dimension of elasmobranch movement ecology. Science Advances. 2022 Aug 19;8(33):eabo1754.

Andrzejaczek S, Gleiss AC, Pattiaratchi CB, Meekan MG. Patterns and drivers of vertical movements of the large fishes of the epipelagic. Reviews in Fish Biology and Fisheries. 2019 Jun 1;29(2):335-54.

Holland K, Meyer C, Potemra J, Holland M. Ocean depth–temperature profiles for operational oceanography from a shark-borne transmitter. Animal Biotelemetry. 2022 Nov 19;10(1):34.

824 Pagniello CM, Castleton MR, Carlisle AB, Chapple TK, Schallert RJ, Fedak M, Block BA.
825 Novel CTD tag establishes shark fins as ocean observing platforms. Scientific Reports.
826 2024 Jun 15;14(1):13837.
827 Legler DM, Freeland HJ, Lumpkin R, Ball G, McPhaden MJ, North S, Crowley R, Goni GJ,
828 Send U, Merrifield MA. The current status of the real-time in situ Global Ocean
829 Observing System for operational oceanography. Journal of Operational Oceanography.
830 2015 Aug 18;8(sup2):s189-200.
831 Liblik T, Karstensen J, Testor P, Alenius P, Hayes D, Ruiz S, Heywood KJ, Pouliquen S, Mortier
832 L, Mauri E. Potential for an underwater glider component as part of the Global Ocean
833 Observing System. Methods in Oceanography. 2016 Dec 1;17:50-82.
834 Sherman J, Davis RE, Owens WB, Valdes J. The autonomous underwater glider" Spray". IEEE
835 Journal of Oceanic Engineering. 2001 Oct 31;26(4):437-46.
836 Eriksen CC, Osse TJ, Light RD, Wen T, Lehman TW, Sabin PL, Ballard JW, Chiodi AM.
837 Seaglider: A long-range autonomous underwater vehicle for oceanographic research.
838 IEEE Journal of oceanic Engineering. 2001 Oct 31;26(4):424-36.
839 Webb DC, Simonetti PJ, Jones CP. SLOCUM: An underwater glider propelled by environmental
840 energy. IEEE Journal of oceanic engineering. 2001 Oct 31;26(4):447-52.
841 Miles T, Seroka G, Glenn S. Coastal ocean circulation during Hurricane Sandy. Journal of
842 Geophysical Research: Oceans. 2017 Sep;122(9):7095-114.
843 Riser SC, Freeland HJ, Roemmich D, Wijffels S, Troisi A, Belbéoch M, Gilbert D, Xu J,
844 Pouliquen S, Thresher A, Le Traon PY. Fifteen years of ocean observations with the
845 global Argo array. Nature Climate Change. 2016 Feb;6(2):145-53.

Lin M, Yang C. Ocean observation technologies: A review. Chinese Journal of Mechanical Engineering. 2020 Dec;33(1):32.

Rudnick DL. Ocean research enabled by underwater gliders. Annual review of marine science. 2016 Jan 3;8(1):519-41.

Testor P, De Young B, Rudnick DL, Glenn S, Hayes D, Lee CM, Pattiaratchi C, Hill K, Heslop E, Turpin V, Alenius P. OceanGliders: a component of the integrated GOOS. Frontiers in Marine Science. 2019 Oct 2;6:422.

Roemmich D, Argo Steering Team. Argo: the challenge of continuing 10 years of progress. Oceanography. 2009 Sep 1;22(3):46-55.

Mensah V, Roquet F, Siegelman-Charbit L, Picard B, Pauthenet E, Guinet C. A correction for the thermal mass–induced errors of CTD tags mounted on marine mammals. Journal of atmospheric and oceanic technology. 2018 Jun;35(6):1237-52.

Siegelman L, Roquet F, Mensah V, Rivière P, Pauthenet E, Picard B, Guinet C. Correction and accuracy of high-and low-resolution CTD data from animal-borne instruments. Journal of Atmospheric and Oceanic Technology. 2019 May;36(5):745-60.

Jonsen I, McMahon C, Harcourt R. Best Practice Manual for SMRU CTD Satellite Relay Data Loggers: Instrument Calibration, Near Real-Time and Delayed Mode Data QA/QC. Version 1.0.[ENDORSED PRACTICE].

U.S. Integrated Ocean Observing System (IOOS). *Quality Assurance / Quality Control of Real Time Oceanographic Data* [Internet]. Silver Spring, MD: National Oceanic and Atmospheric Administration. Available from: https://ioos.noaa.gov/project/qartod/

U.S. Integrated Ocean Observing System. Manual for Real-Time Quality Control of In-situ Temperature and Salinity Data: A Guide to Quality Control and Quality Assurance of In-

situ Temperature and Salinity Observations. Version 2.1. Silver Spring, MD: U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Ocean Service, Integrated Ocean Observing System; 2020. Available from: https://cdn.ioos.noaa.gov/media/2020/03/QARTOD_TS_Manual_Update2_200324_final.pdf

Lueck RG, Picklo JJ. Thermal inertia of conductivity cells: Observations with a Sea-Bird cell. Journal of Atmospheric and Oceanic Technology. 1990 Oct;7(5):756-68.

Garau B, Ruiz S, Zhang WG, Pascual A, Heslop E, Kerfoot J, Tintoré J. Thermal lag correction on Slocum CTD glider data. Journal of Atmospheric and Oceanic Technology. 2011 Sep;28(9):1065-71.

Alvarez A, Stoner R, Maguer A. Performance of pumped and un-pumped CTDs in an underwater glider. In2013 OCEANS-San Diego 2013 Sep 23 (pp. 1-5). IEEE.

Ullman DS, Hebert D. Processing of underway CTD data. Journal of Atmospheric and Oceanic Technology. 2014 Apr;31(4):984-98.

Wang Y, Luo C, Yang S, Ma W, Niu W, Liu H. Modified thermal lag correction of CTD data from underwater gliders. Journal of Coastal Research. 2020 May 1;99(SI):137-43.

Trump CL. Effects of ship's roll on the quality of precision CTD data. Deep Sea Research Part A. Oceanographic Research Papers. 1983 Nov 1;30(11):1173-83.

Merckelbach L, Berger A, Krahmann G, Dengler M, Carpenter JR. *GliderFlight: Python tools for estimating underwater glider flight parameters* [software]. Available from: https://github.com/smerckel/gliderflight

Nakanowatari T, Ohshima KI, Mensah V, Mitani Y, Hattori K, Kobayashi M, Roquet F, Sakurai Y, Mitsudera H, Wakatsuchi M. Hydrographic observations by instrumented marine mammals in the Sea of Okhotsk. Polar Science. 2017 Sep 1;13:56-65.

Meyssignac B, Boyer T, Zhao Z, Hakuba MZ, Landerer FW, Stammer D, Köhl A, Kato S, L'ecuyer T, Ablain M, Abraham JP. Measuring global ocean heat content to estimate the Earth energy imbalance. Frontiers in Marine Science. 2019 Aug 20;6:432.

Abraham JP, Baringer M, Bindoff NL, Boyer T, Cheng LJ, Church JA, Conroy JL, Domingues CM, Fasullo JT, Gilson J, Goni G. A review of global ocean temperature observations: Implications for ocean heat content estimates and climate change. Reviews of Geophysics. 2013 Sep;51(3):450-83.

Perkin R, Lewis E. The practical salinity scale 1978: Fitting the data. IEEE Journal of Oceanic Engineering. 1980 Jan 31;5(1):9-16.

Lewis EL. The practical salinity scale of 1978 and its antecedents. Marine geodesy. 1982 Jan 1;5(4):350-7.

Woody C, Shih E, Miller J, Royer T, Atkinson LP, Moody RS. Measurements of salinity in the coastal ocean: A review of requirements and technologies. Marine Technology Society Journal. 2000 Jan 1;34(2):26-33.

908 [Table 3]
909

	CTD SRDL tag, in situ						Manufactured Specifications					
	Mixed Mission			Stratified Mission			MiniPAT		SPLASH-F		CTD SRDL	
Variable	RMSE	Bias ± SD	% Error	RMSE	Bias ± SD	% Error	Resolution	Accuracy	Resolution	Accuracy	Resolution	Accuracy
Pressure (dbar)	1.03	-0.913 ± 0.48	4.58	2.06	-1.71 ± 1.1	7.79	0.49 – 0.98	± 1%	0.061	± 1%	2 ± (0.03+0.035%)	± 0.15%
Temperature (˚C)	0.069	0.066 ± 0.02	0.82	0.294	-0.11 ± 0.27	1.68	0.05	± 0.1	0.02	± 0.1	0.001	± 0.005
Conductivity (S/m)	0.045	-0.033 ± 0.03	1.27	0.173	-0.162 ± 0.06	3.97					0.0002	± 0.005

910