Dear Dariia,

Thank you for taking the time to review our manuscript and for your thoughtful comments. We thoroughly considered your comments and are planning on making the changes as listed in bold below. We are still working on the final version of the manuscript to make sure that all new edits will fit in the corresponding sections.

Thank you again,

Claudine and co-authors

The manuscript by Hauri et al. details the integration of CO2 and CH4 sensors into the Seaglider and its performance during trials and rigorous data quality assessment. This work is an important contribution to the community effort to increase the number and type of observations in the global ocean, with a particular focus on the water column biogeochemistry and GHG fluxes.

The biggest obstacles to high-resolution carbonate system data from mobile platforms to date are the size, power consumption, and slow response time of the gas sensors. While pCO2 sensors have a relatively slow response, it is still possible to correct the glider-borne data to get decent quality, as nicely shown in this work. Please see my comment on Figure 8, though. As for the methane sensor, I've been wondering whether glider deployments are the best avenue for getting water-column data. The very slow response time of the CH4 sensor doesn't allow for full advantage of the glider-specific capabilities and allows for getting qualitative results, at best. In this respect, ROV-based surveillance or moored platforms would perhaps be a better and more economical option for methane monitoring. It all, of course, depends on the research objectives: qualitative vs quantitative assessment of CH4 distributions. The combined CH4/CO2 sensor is a great idea, too, but given the difference in response times, it will be hard to take full advantage of such a package on gliders. Perhaps an additional discussion on the complexity of the problem of ocean observing - there is no 'one size (sensor) fits all' approach when talking about autonomous sampling, sensors and platforms - would help to orient the readers in the field and help them appreciate the presented work even more.

These are some very interesting thoughts and align with Damian's (reviewer 2) comments, too. We reframed the discussion accordingly.:

"The SG HydroC CH_4 was successfully integrated into the Seaglider as part of this project. While tank experiments showed promising results, short field tests of the CH_4 Seaglider in shallow water revealed low and patchy methane concentrations near the detection limit. The CH_4 Seaglider requires further testing

in environments with strong $p$$CH_4$ gradients during longer and deeper dives (to allow for equilibration) to assess the accuracy of its response time-corrected data in the field. The sensor's slow response time also limits the glider to providing qualitative rather than quantitative results. However, due to the scarcity of oceanic CH_4 observations, deploying a CH_4 glider can help identify the location of methane sources and guide the placement of in situ observations to conduct a more quantitative assessment of CH_4 fluxes and dynamics."

Competing interests. Clearly declare the potential of a direct financial benefit to the co-authors affiliated with the private companies. That's what this section is for.

Thank you for pointing this out. We will add the following statement:

"Authors Hayes and Abdi are employed by AOOI and CSCS (respectively) and their objective is to support the ocean research community by providing innovative, cutting-edge observing technological solutions. These include autonomous platforms and related services in unique configurations. Through the support of the National Science Foundation and the National Oceanographic Partnership Program, AOOI was able to jointly develop the CO2 and CH4 gliders and prove and improve the scientific utility of this approach. Authors Kinski and Kemme are employed by -4H-JENA engineering GmbH, the manufacturer of the HydroC CO2 sensor. The objective of -4H-JENA engineering GmbH is to provide best possible accuracy of dissolved gas measurements on any platform and at any environmental condition. Intensive collaboration with scientists is essential for the development of these products."

Minor comments:

L33: 'greater CH4 activity' – either specify the location or use something general e.g. 'field mission'.

We rephrased to: "The CH_4 Seaglider passed its flight trials in Resurrection Bay, yet needs to be tested during a field mission in an area with CH_4 concentrations beyond background noise."

L40: pCO2 and/or fCO2
Done.

L61: Some of the listed obstacles are the necessary steps to ensure high data quality. Please make a distinction here: those to assure QC and those unique to the glider/mobile/profiling integration or data processing.

We rephrased: "Biogeochemical sensors deployed on autonomous platforms such as moorings and Argo floats have become more commonly used in recent years, but high power requirements, sensor size, and data quality continue to present obstacles to widespread adoption and utilization on underwater gliders."

L76: Here, at the start, some context as to why you mention pH sensors would be great.
We rephrased: "Ocean gliders autonomously collect water column data along planned waypoints, which allows for controlled exploration and adaptive sampling. To date, pH is the only carbon system parameter that has been successfully integrated into ocean gliders (Hemming et al., 2017; Saba et al., 2019; Possenti et al., 2021; Takeshita et al., 2021). The most promising results came from ISFET based pH sensors (Saba et al., 2019; Wright-Fairbanks et al., 2020; Takeshita et al., 2021)."

L111: Zero correction is good, but span correction is also needed in post-deployment treatment.
Thank you for the comment, span correction was included in the post-processing when a post-calibration was conducted but was not clearly stated in the text and has been corrected:

"SG HydroC CO_2 (SG HydroC CO2T-0422-001) data from the tank experiment (Table 1, Figure 4) and rosette mounted CTD casts (Table 2 and 3, Figure 7 and S1) were post-processed to correct for baseline drift (change in the zero signal reference) and span drift (changes in the sensor's concentration dependent characteristics) using pre- and post-calibration coefficients interpolated over the deployment (Fietzek et al., 2014). For the May 2022 Seaglider integrated SG HydroC CO_2 sensor (SG HydroC CO2T-0718-001, Table 3, Figures 8 and 9), data were post-processed with pre-calibration coefficients only (no span drift correction) because the sensor was damaged during the return shipment for post-calibration."

L116: Indicate rpms of the pump used.

RPM is not stated on Seabird's documentation. Seabird states a 5M has a flow rate of 25 ml/s (https://www.seabird.com/modular/sbe-5m-mini-submersible-pump/family-cms.block?productCategoryId=54627473799) but this was not measured directly when powered by the Seaglider.

L133: Please provide more details here about the testing, it is not clear from the description.

We rephrased to: "In situ comparison of the orientation of the sensor suggested the highest data quality is achieved with this mounting design based on close examination of pCO_2 and internal pressure data from the HydroC oriented with the membrane facing forward."

L143-onward: are the drivers available for users in open-access mode should anyone decides to outfit their glider with the sensor? The same for SIRMA. Please clarify.
The CNF file and SIRMA code will be both publicly available upon submission of the final manuscript. The links will be added to the data access paragraph. We also added: "More detailed information can be found in the OceanGliders CO_2 SOP (CITE)."

L168: Please specify the wavelengths of the ecopuck channels.
We now specify in the manuscript: "... Wetlabs Ecopuck measuring chlorophyll fluorescence at 695 nm."

L178: Maybe paste a link to the glider specs here. Please provide a detailed description of the mission here (distance, duration, energy use, number of dives, depth, number of water samples collected and their depth, etc).
We don't think a link to the glider specs is necessary nor appropriate at this point. Table will be provided.

L190: on rocks??? Maybe a sketch diagram will help picture the flow-through setup.
We don't think that a sketch up is necessary
We simplified "were secured next to the Seaglider".

L197: It's not clear what you mean here.
We reworded to: "The SBE-55 and SG HydroC CO2 were powered by a SBE-33 carousel deck unit. The SG HydroC CO2 interfered with the communication stream and thereby prevented real-time data acquisition and control of the SBE-55, however HydroC data were internally logged as required."

L201: Which depths were targeted? And why? Please describe the CTD cast, including the number of stops, their duration, and the reasoning behind such an experiment design.

The SG HydroC CO_2 interfered with the communication stream and thereby prevented real-time data acquisition and control of the SBE-55, however HydroC data were internally logged as required. Depth of the rosette package was monitored directly on the winch and the timing of firing of the sample bottles, after an approximate 15 min hovering period to (allow for equilibration), was programmed in advance based on time intervals. Target depths for discrete water sample collection were 5 m, 20 m, 40 m, 60 m, and 80 m. However, only samples from the upper 20 m of the water column were usable due to issues with manually measuring the depths and sample collection.

L205-208: This paragraph probably fits better with the description of the tank experiment above.

We think that this paragraph fits best where it is at the moment, since it describes discrete water samples in general.

L283-286: Was the span correction applied too? Was the detector calibrated in post-processing beyond zero calibration?

Yes, span drift was corrected when post-calibration was carried out (see comment above and corresponding addition to paper). Span drift correction was not performed for the Seaglider mounted HydroCs because post-calibration was not possible (in one instance the sensor was damaged on transport back to 4H Jena, and the other instance the sensor was lost) so the temporal stability of the sensor response slope (span drift) could not be determined.

We clarified: "SG HydroC CO_2 (SG HydroC CO2T-0422-001) data from the tank experiment (Table 1, Figure 4) and rosette mounted CTD casts (Table 2 and 3, Figure 7 and S1) were post-processed to correct for baseline drift (change in the zero signal reference) and span drift (changes in the sensor's concentration dependent characteristics) using pre- and post-calibration coefficients interpolated over the deployment (Fietzek et al., 2014). For the May 2022 Seaglider integrated SG HydroC CO_2 sensor (SG HydroC CO2T-0718-001, Table 3, Figures 8 and 9), data were post-processed with pre-calibration coefficients only (no span drift correction) because the sensor was damaged during the return shipment for post-calibration."

L291: It’s not clear here why the real-time data is needed for post-processing. Wasn’t the sensor recording data internally? How do you expect this lack of calibration to affect data quality?

The data were recorded internally, but the glider was lost at sea during the winter mission. One of the key parameters for post processing (*p_NDIR*) was not included in the real-time data. Since the HydroC was only collecting data for ~4 days during the spring mission and ~2 days during the winter mission, we do not believe this lack of calibration affected data quality.

We clarified: “The *p*CO_2 data from February 2023 were not post-processed because a required parameter was not relayed in real-time and the glider was lost. Lack of post-calibration most likely had no negative effect on the quality of data since the HydroC was only collecting data for ~4 days during the spring mission and ~2 days during the winter mission.”

We addressed the loss of the glider in more detail at the end of section 2.3 - Spring and winter CO2 Seaglider missions:
“Before the February mission the on board modem was replaced with a different model, with different input voltage requirements, which were probably not met as the mission evolved. As a result, the glider could not communicate anymore and was lost. While this was an unfortunate mistake, the loss of the glider had nothing to do with the HydroC CO2 integration.”

L305-307: you refer here to the detector response time, not the sensor response time, correct? Please be clear. Because in the next paragraph, you give 106,108 and 109s as the response times. This could be confusing to the reader.
We clarified what we mean by response time:
“The ability to determine the in situ response time (t_63 of the HydroC, including membrane and SBE-pump contributions) of the sensor enables the user to correct for hysteresis through data post processing, critical for a sensor operating on profiling platforms or anywhere where strong environmental gradients are encountered.”

L315: Was it 1min running average or? Why was 1min was chosen? With a response time of ca 2 min, 1 min sampling resolution is not sufficient to apply RTC reliably, especially when resolving gradients. I think Miloshevish (2004) touches on that.

Thank you for this thoughtful comment. After carefully reviewing the Miloshevich et al. (2004) and Dølven et al. (2022) papers again, HydroC CO2 data were

reprocessed accordingly. Instead of 1-minute averaging, smoothing was applied to keep the original 2 second HydroC resolution before RTC was applied. The RTC resolution was changed to 8 seconds following the L-curve analysis included in the RTC code by Dølven et al. (2022). Figures were updated accordingly.

Correction: Lines 311-321
Response times determined during calibration at -4H-JENA were used for response time correction (RTC) and found to be 106 seconds for the HydroC mounted on the rosette in May 2022 and 108 seconds when it was integrated into the Seaglider in February 2023 (HydroC CO2T-0422-001). The response time of the HydroC integrated into the Seaglider in May 2022 (HydroC CO2T-0718-001) was 109 seconds. Since field verification of the response time was recommended to ensure the highest quality post-processed data product, we verified the sensor response time at deployment. After the glider was stationary for approximately 15 minutes, a zeroing interval was performed with the HydroC CO_2. The response time was determined by reviewing the time it took for the signal to recover to the ambient concentration. Our in situ response time tests suggested to be within 5 seconds of the response time found during calibration (not shown). Before RTC was applied, HydroC CO_2 data were smoothed using a quadratic regression (MATLAB's smoothdata.m function with the loess method) over a 2 minute window. This was done to eliminate erroneous spikes in the RTC data and retain the original 2 second resolution of the pCO_2 data. A response time correction resolution of 8 seconds was determined with the L-curve analysis included in the publicly available code from Dølven et al. (2022). The Dølven et al. (2022) RTC method produced more realistic profiles than an RTC method (Miloshevich et al., 2004, not shown) previously used for HydroC CO2 correction from a profiling float (Fielder et al. 2013), so we opted to use the Dølven et al. (2022) algorithm. In addition, Dølven et al. (2022) developed their algorithm with equilibrium-based sensors in mind and was proven with a sensor with a long response time (HydroC CH4 $\tau 63 \cong 23$ minutes). HydroC data at the original resolution (2 s) and RTC resolution (8 s) were linearly interpolated onto the Seaglider timestamp and 1-meter binned data were calculated by first averaging 1 meter (+/- 0.5 m) upcast and downcast data independently, linearly interpolating over gaps, then averaging the interpolated 1-meter binned upcast and downcast together.

L327-329: did the pumps both have the same rpms? This result is surprising. There should be no difference between 5T and 5M, only the pressure rating.

RPM for 5M is not stated on Seabird's documentation. Seabird states a 5M has a flow rate of 25 ml/s (https://www.seabird.com/modular/sbe-5m-mini-submersible-pump/family-cms.block?productCategoryId=54627473799) but this was not measured directly when powered by the Seaglider.

L359: Check Equation 1, must be % diff = (delta/pCO2 disc) *100%
Thank you for catching this! The denominator should be divided by two so it should read:
% diff = (delta/(pCO2HydroC - pCO2 disc)/2) *100%
The values are within ~0.1% using the equation you recommend.

L366: Sensor-pCO2 or discrete pCO2?
We will add "... *p*CO_2 measured with the HydroC …"

L375: The dashed black line shows the downcast data. I think you should be comparing both upcast and downcast to the discrete samples, especially given the fact that it seems to be an unresolved time lag if using lab-derived time constraints (see my comment to Fig. 8 below for some ideas). How far apart in time were glider upcast and downcast?
It is standard practice to give differences based on a ~1 meter depth averaged profile (e.g. Saba et al., 2019) and since we were outside the recommended distance for comparison (~100 m Thompson et al., 2021) we believe this is sufficient.

L386: It would be easier to read this paragraph if the May glider mission is compared to the May Sunny Cove data and the Feb glider data to the Feb buoy data first. The agreement is quite impressive, and you should emphasize this fact more clearly.
We rewrote to: "NOAA's moored sensor located in Sunny Cove (59.911 °N, -149.35 °W), near the CO_2 Seaglider trial site, measured an average surface *p*CO_2 of 240.7 +/- 10.4 µatm during the time of the May 2022 mission, which compares remarkably well with the glider based measurements. A minimum of 140 µatm was measured in Sunny Cove in mid-April (3-day average) (Figure 12, Monacci et al., 2023), suggesting that the peak of the spring bloom happened three weeks prior to our glider mission."

L415: What could be done to increase data accuracy further? Some recommendations and the vision forward would be useful for the reader.
We added: -4H-Jena is reassessing their sensor calibration methodology and data post-processing algorithm to further improve the HydroC's data accuracy."

L417: Not only the pCO2 gap but also carbonate system dynamics.
Was added to text.

L452: Could this pCO2/pCH4 sensor solution easily be integrated into other gliders (e.g. Slocums)? Would data processing be similar? Could other glider users easily replicate this integration?
This was addressed in lines 429-437.

Figure 8: How was 1-meter binned calculated? It seems like RTC applied to downcast/upcast profiles using lab-derived time constants is insufficient. If you were to find an in situ time constant that would collapse the profiles on each other, that would be a really useful exercise.
We looked into this at first, similar to the methods used for thermal lag correction (Garau et al., 2011) but chose to use the response time correction based on the known sensor response time. This was done because, though the assumption of measuring the same water column on the downcast and upcast for shallow dives may be justifiable (we did not quantify this), we wanted our methods to translate to deep dives and missions. Additionally, using a published correction method with the provided t_63 of the sensor and pump, eases the burden of future glider users and data processors and was found to produce good results (see Tables 1-3).

We added the following to the Response Time Correction section: "HydroC data at the original resolution (2 s) and RTC resolution (8 s) were linearly interpolated onto the Seaglider timestamp and 1-meter binned data were calculated by first averaging 1 meter (+/- 0.5 m) upcast and downcast data independently, linearly interpolating over gaps, then averaging the interpolated 1-meter binned upcast and downcast together. "

L913: add legend to the plot
The different lines are described in the caption and do not need to be shown in a legend.

L930: add Legend to the plot
The different lines are described in the caption and do not need to be shown in a legend.

Figure 10: The glider dataset is indeed impressive; however, it is of no relevance to this paper. Therefore, please show only the period when Contros was on.

Since this paper also discusses scientific results and provides the first ever winter high resolution O2 and carbon measurements in Resurrection Bay, we believe that this figure and discussion of the full mission is valuable to the community studying this region.