L33: ‘greater CH_4 activity’ – either specify the location or use something general e.g. ‘field mission’.

L40: pCO_2 and/or fCO_2

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.

L76: Here, at the start, some context as to why you mention pH sensors would be great.

L111: Zero correction is good, but span correction if also needed in post-deployment treatment.

L116: Indicate rpms of the pump used.

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

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.

L168: Please specify the wavelengths of the ecopuck channels.

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).

L190: on rocks??? Maybe a sketch diagram will help picture the flow-through setup.

L197: It’s not clear what you mean here.

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.

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

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

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.

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.

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.

L359: Check Equation 1, must be % diff = (delta/pCO_2 disc) *100%

L366: Sensor-pCO_2 or discrete pCO_2?

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?

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.

L415: What could be done to increase data accuracy further? Some recommendations and the vision forward would be useful for the reader.

L417: Not only the pCO_2 gap but also carbonate system dynamics.

L452: Could this pCO_2/pCH_4 sensor solution easily be integrated into other gliders (e.g. Slocums)? Would data processing be similar? Could other glider users easily replicate this integration?

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.

L913: add legend to the plot

L930: add Legend to the plot

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.