

Dear Damian,

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

Comment on “Expanding seawater carbon dioxide and methane measuring capabilities with a Seaglider “ by Hauri et al.

Summary: The manuscript by Hauri et al. presents the development and comprehensive testing of a glider-based sensor package for measuring water column partial pressures of CO₂ and CH₄. This is a timely task because it directly addresses the need of increasing observations of the spatial and temporal variability of these major greenhouse gases.

Overall assessment: This contribution is significant for the field in that paves the way for large-scale surveys of greenhouse gases during process studies ranging from mid-water to the deep (1000 m) ocean, which in turn could be useful for the combined investigation of physical and chemical oceanographic variability. The manuscript is well written, figures and tables are mostly of adequate quality and the literature choice is appropriate. However, a major caveat I see in the study is the applicability of this approach given the different “readiness” level of the sensors for in-situ measurements of CO₂ and CH₄. While for the former the authors clearly show that the data obtained during both laboratory and at-sea conditions matches the desired accuracy and resolution (albeit uncertainties), for the latter both aspects rise questions on whether gliders are actually the right platform to be used (which is also an issue raised by Dr. Atamanchuk on her comment to this manuscript). The authors indicate that their pCH₄ measurements have an uncertainty of +/- 2 µatm, which is problematic because it would not allow fully distinguishing between under- and supersaturated conditions in the water column. Although I would expect this to be a more serious issue in open rather than in near-coastal settings, I would still expect that in the latter part of the seasonal variability (and possibly a large fraction of the water column) could not be adequately resolved. Recent intercomparison efforts from different groups (see Wilson et al., 2020; <https://doi.org/10.5194/bg17-5809-2020>) came to the conclusion that although there is no consensus on the threshold for “high-quality” seawater CH₄ measurements, the

achieved accuracy should be such that it allows tracking the ocean's response to increasing tropospheric CH₄ in time scales of 5 years (which translates into an analytical agreement of <1 % between independent observations). I would argue that, at this point, this should be the target for marine water column measurements. Furthermore, the slow response time of the CH₄ sensor poses a practical constraint for vertical profiling, at least if combined with the CO₂ sensor. The CH₄ sensor used in this study is therefore rather suitable for stand-alone (and potentially shorter) deployments in gliders, or long-term applications in, for instance, moored observatories where seasonal and longer time scales of variability are the target. I noticed that the description of the validation conducted by the authors is mostly centered around CO₂ and my impression is that, at least in parts of the manuscript, the work done for CH₄ was addressed somewhat qualitatively. Since the authors do have data to substantiate their thorough tests, I would kindly invite them to discuss the above mentioned caveats in their manuscript. In the following, I list general and specific comments not only to support my general assessment, but also in the hope that this is useful for the authors in view of a potential revision.

We thank both reviewers for their thoughtful comments on the methane glider. We revised the discussion paragraph on methane:

“While tank experiments showed promising results, short field tests of the CH₄ Seaglider in shallow water revealed low and patchy methane concentrations near the detection limit. The CH₄ Seaglider requires further testing in environments with strong $p\text{CH}_4$ 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 CH₄ observations, deploying a CH₄ glider can help identify the location of methane sources and guide the placement of in situ observations to conduct a more quantitative assessment of CH₄ fluxes and dynamics.”

General comments: Introduction:

- The manuscript is strongly biased towards CO₂ and this is reflected also in the introduction. If the platform is to be presented as relevant for both gases, the key processes for CH₄ cycling in the ocean should at least be briefly mentioned. This is important because for a reader not familiar with trace gases, it will be hard to grasp where CH₄ comes from in the ocean and why processes such as seepage and permafrost thawing (as mentioned by the authors) are so relevant.

We now start the introduction with:

“Understanding the distribution and dynamics of carbon dioxide (CO₂) and methane (CH₄) in the ocean is crucial for predicting and mitigating climate change and ocean acidification impacts.”

After a paragraph on CO₂ we added:

“CH₄ possesses a global warming potential over a 100-year period that is approximately 28 times greater than that of CO₂ (IPCC AR5; Myhre et al., 2013). Sediments along the seafloor at continental margins contain large amounts of CH₄, with about ten times as much carbon than the atmosphere (Kessler, 2014). CH₄ is biologically produced in anoxic sediments and the surface mixed layer or released from geological sources like hydrocarbon seeps and degrading methane hydrate deposits (Barnes and Goldberg; Du et al, Skarke 2014). This powerful greenhouse gas is emitted to the atmosphere through bubbling (ebullition) or diffusive gas transfer (Reeburgh, 2007; McGinnis et al., 2006), which is limited by rapid oxidation to CO₂ during transport through the water column (Leonte et al., 2017). CH₄ occurs generally at low levels (background concentrations) throughout oceans, unless close to a source. Positive feedback mechanisms, like warming induced CH₄ seepage from destabilizing hydrates and thawing subsea permafrost, may further accelerate ocean acidification and climate change (Garcia-Tigreros et al., 2021; Sparrow et al., 2018; Shakhova et al., 2010; Rees et al., 2022).

• The motivation of the study (beyond the notorious technological advances) is not clearly specified. A line of thought which – in my opinion- would help introducing that motivation would be to first mention the extent of the ocean contribution to the natural sources/sinks of these two gases, see e.g.: Global Carbon Budget (<https://doi.org/10.5194/essd-15-5301-2023>) The global Methane Budget 2000 – 2017 (<https://essd.copernicus.org/articles/12/1561/2020/>) After which the issue of undersampling would underpin why extending the usage of autonomous platforms to climate-relevant trace gases is urgent. Overall, I would recommend revising the structure of this section to show more clearly why this is great progress that needs to be further developed and implemented. While I appreciate that the focus of the manuscript is more technical, doing this would increase the impact of this work.

We now start the intro with: “Understanding the distribution and dynamics of carbon dioxide (CO₂) and methane (CH₄) in the ocean is crucial for predicting and mitigating climate change and ocean acidification impacts.”

We believe that this was also addressed in the preprint with the following lines on CO₂: “On top of this natural variability, the ocean has absorbed about one third of the CO₂ emitted by humans since the industrial revolution (Sabine et al., 2004; Gruber et al., 2019). In doing so, it has played an important role in mitigating climate change (Sabine and Tanhua, 2010).”

We also added a new paragraph on CH₄:

“CH₄ possesses a global warming potential over a 100-year period that is approximately 28 times greater than that of CO₂ (IPCC AR5; Myhre et al., 2013). Sediments along the seafloor at continental margins contain large amounts of CH₄, with about ten times as much carbon than the atmosphere (Kessler, 2014). CH₄ is biologically produced in anoxic sediments and the surface mixed layer or released from geological sources like hydrocarbon seeps and degrading methane hydrate deposits (Barnes and Goldberg; Du et al, Skarke 2014). This powerful greenhouse gas is emitted to the atmosphere through bubbling (ebullition) or diffusive gas transfer (Reeburgh, 2007; McGinnis et al., 2006), which is limited by rapid oxidation to CO₂ during transport through the water column (Leonte et al., 2017). CH₄ occurs generally at low levels (background concentrations) throughout oceans, unless close to a source. Positive feedback mechanisms, like warming induced CH₄ seepage from destabilizing hydrates and thawing subsea permafrost, may further accelerate ocean acidification and climate change (Garcia-Tigreros et al., 2021; Sparrow et al., 2018; Shakhova et al., 2010; Rees et al., 2022).”

• In line 57 and following, the authors mention how key parameters are under-sampled. Most of the variables measured by the platform presented in this new application including gases are actually classified as essential ocean variables of the Global Ocean Observing System (see <https://goosocean.org/what-we-do/framework/essentialocean-variables/>). I would suggest the authors to mention this framework in their manuscript in order to emphasize how scientific and technical developments in this direction do address a timely task at an international level.

To address this and the next comment we rewrote this paragraph: “To effectively observe and understand the complex processes and feedback mechanisms regulating Earth's systems, certain key parameters, known as essential ocean variables (EOVs), must be measured accurately. However, these variables are often vastly undersampled across time and space due to traditional sampling methods, which rely mainly on discrete water sample collections from dedicated research cruises, underway measurements from transiting vessels, or time series measurements from in situ sensors on fixed moorings. Although biogeochemical sensors deployed on autonomous platforms like moorings and Argo floats have become more prevalent in recent years, challenges such as high power requirements, sensor size, and data quality hinder their widespread use on underwater gliders. Autonomous, spatially resolved surface measurements of

pCO₂ and pH are collected using wave gliders and sail drones (Chavez et al., 2018; Nickford et al., 2022; Manley and Willcox, 2010). The state-of-the-art biogeochemical (BGC) Argo floats measure variables like pH, O₂, NO₃, chlorophyll-a, suspended particles, and downwelling irradiance in subsurface waters (Claustre et al., 2020). These floats can last for several years at low sampling resolutions, such as a 2000-meter depth profile every ten days, or they can be programmed for high-resolution and shallow sampling. They can even sample beneath seasonal sea ice (Briggs et al., 2018). Despite their capabilities, their trajectory is hard to control, and they are usually not recovered after their mission, which prevents sensor calibration and post-mission corrections.”

- With regards to undersampling, it would be important to mention that this does not only refer to spatial, but also temporal coverage. This would emphasize further the potential contribution of the approach presented in this manuscript, as the largest sampling deficits occur in winter.

We agree. We now write: “However, these variables are often vastly undersampled across time and space due to traditional sampling methods, which rely mainly on discrete water sample collections from dedicated research cruises, underway measurements from transiting vessels, or time series measurements from in situ sensors on fixed moorings.”

- Since the authors aim to show the advantages of glider-based, large-scale measurements of CO₂ and CH₄, it would be good to provide the reader examples of the current approaches and their limitations (added as citations). Some suggestions (which include further examples therein) are as follows: Behncke et al (2024): A detectable change in the air-sea CO₂ flux estimate from sailboat measurements, <https://www.nature.com/articles/s41598-024-53159-0> Resplandy et al (2024): A Synthesis of Global Coastal Ocean Greenhouse Gas Fluxes, <https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2023GB007803>

We believe that the section about wavegliders, saildrones, and BGC Argofloats addresses this comment.

Methods: • Some statements in this section (mostly regarding CH₄) are vague or require additional explanation in order for the readers to be able to fully understand the methodological approach of the study (and ensure reproducibility).

We revised large sections of the methods- for example we added: “ Discrete CH₄ were converted from the concentration of dissolved gas in water (mol L⁻¹) to partial pressure (µatm) using the solubility coefficient following Sarmiento and Gruber (2006).”

- As part of the validation discrete samples for both gases were collected and analysed. While it is clear that the glider sensors provide CH₄ values as partial pressures, the discrete samples were measured as dissolved gas concentration in an aqueous matrix. This means that partial pressures needed to be calculated. I would suggest the authors to clarify this aspect in the manuscript.

Rephrase 2.7.3 pCH₄ post-processing

“SG HydroC pCH₄ data were response time corrected using a τ_{63} of 43 minutes (Dølven et al., 2022; Figure 4c). Discrete CH₄ were collected during the tank experiment (Figure 4c) and analyzed at John Kessler’s laboratory at the University of Rochester. Discrete CH₄ were converted from the concentration of dissolved gas in water (mol L⁻¹) to partial pressure (μ atm) using the solubility coefficient following Sarmiento and Gruber (Chapter 3, Table 3.2.2).”

- In lines 301–302 the authors mention that the response times of their gas sensors do not have a temperature dependency. I think this is an important advance, which should be shown with data in the manuscript.

There are results from laboratory experiments that describe the (compared to PDMS) very low $t_{63}(T_{\text{water}})$ dependency for TOUGH membranes. But as T_{water} is just one parameter influencing t_{63} it is more valuable to highlight that the HydroC CO₂ sensor provides the opportunity to determine t_{63} during the deployment.

We rewrote: “For this study, the SG HydroC CO₂ sensors were deployed with the new robust TOUGH membrane, which uses Teflon AF2400 as the active separation layer and which has a low temperature dependence on the permeability coefficient (Pinnau and Toy, 1996). The response time with the HydroC CO₂ TOUGH membrane is very stable but can be affected by the speed of water exchange across the membrane (e.g. pump speed, tube length, etc.). Field verification of the response time is recommended to ensure the highest quality post-processed data product.”

- I suggest the authors to make sure they write this section consistently in past tense. **We rewrote the methods section and used past tense wherever possible.**

Results and discussion:

- Besides withstanding high pressure and probably the need of extended battery power, I could not not see why differences in the performance of the deep water glider are to be expected (or it is at least not described in the manuscript). Even if there would be any changes, the tests shown on Figure 6 depicts a deployment that was even shallower than the first one, and therefore does not necessarily substantiates the author’s argument.

Section 3.1 aims to convince the reader that the glider was able to fly well because of our careful trim and ballast efforts (to account for the large amount of mass and length added). The fact that we could even confirm stable flight in a shallow dive shows that the glider quickly reached a stable equilibrium. Critically, this means that the flow met by the sensors was undisturbed and data quality maximized.

We rephrased: “Example flight profiles with the POM and Titanium integrated sensors are shown in Figures 5 and 6, respectively.”

• Section 3.2 (comparison with underway measurements): Neither during the methods, nor the results and discussion sections there is mention of how underway systems were used to cross-check the glider CO₂ and CH₄ sensors.

We rephrased: “The quality of the CO₂ Seaglider data was thoroughly tested with discrete measurements during a tank experiment and glider missions.”

• Section 3.3.1 (tank experiments): the description provided here is rather qualitative. I recommend the authors to present the corresponding results on this validation for CH₄, as they nicely did for CO₂.

We wrote: “The SG HydroC CH₄ was evaluated during the tank experiment described in section 2.4 (Figure 4c). Relative differences (Eq. 1) between discrete $p\text{CH}_4$ (average of triplicate samples) and $p\text{CH}_4^{\text{RTC}}$ were 6.3 to 14.6 % (Table 1). During the experiment there was a decrease in salinity from 30.95 to 29.88 where $p\text{CO}_2$ correspondingly decreased by 80 μatm . The corresponding $p\text{CH}_4^{\text{RTC}}$ signal decreased by 25.4 μatm from 32.3 to 6.9 μatm . Although the triplicate discrete $p\text{CH}_4$ water samples were slightly lower than the sensor measured $p\text{CH}_4$ values, they also reflected this step change.”

We also added $p\text{CH}_4$ to Table 1.

Triplicate Date Time (UTC)	$p\text{CO}_2^{\text{disc}} \pm \text{uc}$ (μatm)	$p\text{CO}_{2,\text{sn422}}^{\text{RTC}} - p\text{CO}_2^{\text{disc}}$	$p\text{CO}_{2,\text{sn0718}}^{\text{R}}_{\text{TC}} - p\text{CO}_2^{\text{disc}}$	$p\text{CH}_4^{\text{disc}} \pm \text{u}$ (μatm)	$p\text{CH}_4^{\text{RTC}} - \text{CH}_4^{\text{disc}}$
5/2/2022 2 3:25	298.7 \pm 10.2	-0.9 (0.3%)	-	-	-
5/2/2022 2 7:32	227.1 \pm 7.8	4.3 (1.9%)	2.4 (1.1%)	-	-
5/2/2022 2 11:27	223.3 \pm 7.7	0.7 (0.3%)	-2.6 (1.2%)	-	-
5/2/2022 2 15:30	227.8 \pm 7.9	-1.1 (0.5%)	-3.3 (1.5%)	-	-
5/2/2022 2 00:11	-	-	-	25.4 \pm 2.1	4.0 (14.6%)
5/2/2022 2 12:06	-	-	-	7.3 \pm 1.3	0.5 (6.3%)

Specific comments:

I. 21–22: “The key parameters to observing and understanding (...)”. Here I recommend revising the syntax.

We will leave as is since it is just a question of writing style.

I. 31: (...) provides (...). I would write “provided” instead.

Done.

I. 38–41: Spatial variability appears twice. I would check here to avoid redundancy.

We rephrased to: ”The oceanic reservoir of carbon dioxide (CO₂) is large and of critical importance to Earth's climate, biogeochemical cycles, and the health of marine ecosystems. Within the ocean, CO₂ levels (measured as the partial pressure of CO₂, pCO₂ and/or fugacity of CO₂) are spatially and temporally

variable as they are influenced by a myriad of highly dynamic physical, chemical, and biological processes.”

I. 67: The abbreviation “BGC” should be inserted here.

Done.

I. 68: This is unclear. Of which variables? In the sentence above only pCO₂ and pH are mentioned.

We rephrased to: “The state-of-the-art biogeochemical (BGC) Argo floats measure variables like pH, O₂, NO₃, chlorophyll-a,”

I. 91: Here it should be indicated what is the measurement standard (and/or reference) that substantiates this statement (e.g. accuracy, detection limits, etc.). Is this referred to Newton et al (2015) as included in the caption of figure 7?

We deleted this sentence because we couldn’t find supporting literature. Thanks for pointing this out.

I. 128: “POM”. This abbreviation should be spelled in full upon first usage.

Done.

I. 154: “headspace” instead of “head space”

Done.

I. 163: “percentage” instead of “%”

Done.

I. 223–224: Here the specifics of the CRM should be indicated (e.g. exact denomination, literature if any, etc).

We rephrased to: “This method requires Certified Reference Material (CRM, Batch #198 from A. Dickson’s Certified Reference Materials Laboratory) to create a three-point calibration line.

I. 238: I would suggest “processing” as more appropriate than “manipulation” here.

Agreed.

I. 258,259: “CRM” instead of “Certified Reference Material” as the abbreviation was already defined

Done.

I. 292: “required parameter”. This is rather vague. Here it should be stated which parameter was missing to be able to carry out the post-processing.

We rephrased to:”The $p\text{CO}_2$ data from February 2023 were not post-processed because a required parameter (p_NDIR) was not relayed in real-time...”

I. 450–451: Considering the comparatively low accuracy and long response time, I would have to disagree, unless it is clearly stated for which types of studies this is the case (see also overall assessment above).

We rephrased to: “CO₂ gliders are perfectly suited to contribute data for understanding relevant inorganic carbon processes in coastal shelf and boundary regions where mesoscale or sub-mesoscales dominate.”

The flaws of CH₄ gliders are now discuss in the paragraph just above:

“While tank experiments showed promising results, short field tests of the CH₄ Seaglider in shallow water revealed low and patchy methane concentrations near the detection limit. The CH₄ Seaglider requires further testing in environments with strong $p\text{CH}_4$ 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 CH₄ observations, deploying a CH₄ glider can help identify the location of methane sources and guide the placement of in situ observations to conduct a more quantitative assessment of CH₄ fluxes and dynamics.”

I. 492: There should be a link in the revised version.

The link will be provided in the revised version.

Kind regards, Damian L. Arévalo-Martínez