We thank the expert reviewer for the very fruitful comments and remarks, which helped us to improve the manuscript.

All the remarks been addressed below, and changes in the manuscript have been done accordingly using word track changes. Some grammatical improvements have also been done. The instruments have been named more consistently over the manuscript as SubOcean and HydroC HP.

We hope that the manuscript is now acceptable for publication in Geoscientific Instrumentation, Methods and Data Systems.

Reviewer(s)' Comments to Author:

This paper seeks to demonstrate the reliability of a membrane-based system to continuously measure methane in aquatic environments other than oceans. The authors place their research in stratified lake Kivu, in which water can have very high methane concentrations, and span several orders of magnitude. It is shown how a manipulation of the sensor system can accommodate this large dynamic concentration range, making the presented sensor system a potentially interesting and useful addition to the biogeochemists' toolbox. Such an in-situ sensor system, usable in methane-rich freshwater environments, is of course interesting and relevant, and some of the results are indeed very encouraging, but at the same time the manuscript suffers from some conceptual and structural weaknesses that in my opinion undermine its suitability for publication in the present form.

General points:
Introduction: The authors choose with Lake Kivu an interesting, but also particular case. To me, it is unclear, why the manuscript introduction focusses on lake Kivu (given the choice of journal) instead of focusing on the need for continuous methane profiles in aquatic systems to answer a large array of very interesting and pressing questions, including e.g., transport, production/consumption etc.. While this would require substantial changes to the introduction, it would make a much more useful and strong manuscript. The method presented here is different from a previous paper (Grilli 2018) mostly in the altered measurement mode that allows for much higher concentration to be quantified. That this extended range is still not extensive enough to cover the range in CH4 concentrations of lake Kivu is unfortunate but not critical. A review of the concentration range in freshwater systems would help the reader to understand the relevance of the presented instrument modifications.

We thank the reviewer for those remarks. The introduction has been adjusted as suggested. The importance of better constraint the CH4 emissions to the atmosphere and the processes behind those emissions in aquatic environments is now discussed. We did not reduce the information on Lake Kivu since, even if the paper is centered on the measurement technique, we found important to provide the contest to the reader, and to highlight the importance of the dissolved gas monitoring at this particular site. Safety reasons are driving this monitoring of the evolution of the lake, particularly now that power plants are exploiting the gas reservoir. The reviewer also asked a review of concentration ranges in freshwater systems. We are unfortunately not in the position of providing an extensive review, but we report two examples of meromictic lakes where the upper measurement range of the Sub-Ocean probe is compatible with the concentration of bottom waters. We hope that the reviewer will now consider this introduction better adapted to the focus of this work.

Methods: Even though much of the general aspects of the sub-ocean instruments have recently been described (Grilli 2018), the general setup must be presented here in detail. Further, it is unclear to me how the authors arrive at a systematic error of 10% for the HydroC. I am missing information for the use of the instrument (e.g. lowering velocity, frequency of measurement), which should be moved away from the results section.

The information about the lowering speed, frequency of the acquisition and response time for the SubOcean instrument have now been added in the Method section.

"The embedded spectrometer is continuously measuring the gas composition at 10 Hz, while the response time of the sensor during the campaign, expressed as τ_{90}, was ~10 sec. At a lowering speed of ~6 m min^{-1}, this corresponds to a vertical resolution of 1 m".

The 10% for the HydroC was specified by the manufacturer though a personal communication.

Further information on the extraction system has now been added in the manuscript.

"A stainless-steel membrane block (MB) was equipped with two 10 μm thick polydimethylsiloxane (PDMS) membranes of 56 mm diameter mounted face-to-face. The membranes were mounted on porous bronze frits of 3 mm of thickness (Poral, grade 20), providing mechanical strength for the membrane under high pressure differences. A schematic of the membrane block can be found in the supplementary information of Grilli et al., 2018."

Results: The authors often mix "discussion" and "results" elements which leads to confusion and unnecessary repetition (up to the point where they mention that the "discussion of a certain value can be found in the results section, L269). Any discussion around the underlying reasons for the measurement uncertainties surely do not belong in the results section. Many of the aspects on lake Kivu methane concentrations are unnecessary, and the comparisons of past CH4 measurements in Kivu through in time and space should be better linked to each other (making clear when technical and when ecological reasons drive differences), and synthesized, of course, in the discussion section. To me, it seems odd to compare the average results measured in the surface waters with other values that, depending on situation (e.g. season and mixing), span 3 orders of magnitude (L 236).

Indeed the comparison on surface measurements is not easy to do due to seasonality and mixing. The reviewer is right saying that the span of existing data is large, but we noticed that 1) our measurement is compatible with previous ones; 2) that high values (as high as our) were previously measured on this time of the year and reported by Pasche et al 2011; 3) The CDT data from few months prior the campaign highlighted a mixing event occurring at the beginning of March which would explain this large amount of CH4 at the surface. Even if we do not know what cause this mixing we could clearly observe it from the data reported in Figure 7. From the O2 profiles one can notice that from mid-December the lake started to stratify at 25 m (bringing water enriched in dissolved gas up to this depth), while at the beginning of March the oxic layer increased down to 35 m depth, highlighting a mixing of this upper part which would cause an increase in dissolved gas at the surface.

We now improved the structure, making a Results and Discussion section followed by a conclusion section.

Discussion: basically missing. Much of the discussion (actually found in the results section) oddly focusses on Kivu-specific observations (e.g. temporal mixing dynamics) instead of methodological aspects. The actual discussion section is mostly a repetition of introduction elements alongside some of the uncertainty numbers mentioned in the results. The latter were derived based on the assumption of homogenous methane concentration across the lake (L263), although the distance from the sediments is known to have large influence on local methane concentrations (del Sontro 2018). While this observation is mostly true for epilimnetic waters, it is unclear how if the hydrodynamic features of Lake Kivu allows for omission of this important control of the spatial distribution of CH4 in lakes. Further, as a reader or potential user of such an instrument, I'd be highly interested in performace metrics (e.g. opposed to other techniques and instruments), or reasons for particular performance-related issues (e.g. regarding the carrier gas flow), and these aspects therefore require much more elaboration.

We thanks the reviewers for this remark and input. The reviewer is right saying that concentration of dissolved gas are not constant over the whole surface of the lake, and that could strongly depends on the depth (i.e. distance of the sediments to the surface). With the SubOcean instrument we recently conduct two field campaigns, one in the Black Sea and another one at Lake Aiguebelette (France) that both highlighted this behavior. We therefore added in the discussion the following part:

"This fast response sensor could be used to better investigate the fluxes of CH4 (or other greenhouse gases) from lakes, oceans, rivers and other water reservoirs. In this campaign only a specific location at 5 km from the coast with 410 m of water depth was investigated. The amount of CH4 at the surface could strongly depends on the water depth i.e. on the distance of the sediment to the surface. A fast sensor could allow to follow the spatial distribution of dissolved gas at the surface for different depths, as well as its variability over the seasons. This would help to better constraint the greenhouse gas emissions in the face of global change (DelSontro et al., 2018)."

The Discussion have now been merged with the Results section as suggested by both reviewers. We hope that the actual structure make the manuscript clearer.

Specific points:

Figure 5: panel TDGP, point 70-90m, how can there be a nonlinearity in the interpolation? What value is interpolated?

This has now been corrected. Thanks for spotting this out.

Figure 6: I understand the quantification of σ for precision and repeatability. However, it seems overly simplified to express the error quantified at one concentration in percent, and then extrapolate the relative error across all orders of magnitude (making very small absolute errors in small concentration).

This remark is indeed correct, but it is challenging to report the accuracy of the sensors in the upper layers where the lake is not stratified. We therefore decided to report the actual variability of the measurements at depths from 0 to 80 m, and to fix the error bars to the value estimated at 80 m at lower depth (since only one profile cover the 150 meters, and all the others saturated within the first 100 m). We added this discussion in the text and adapter the Figure accordingly.

« The uncertainty represented by the grey lines in **Erreur ! Source du renvoi introuvable.** represents the measured variability over the eights vertical profiles from 0 to 80 m, and was fixed to ± 22% at larger depths »

L90: It would be interesting to know how long the instrument could run with "between 2 and 40 bar" of carrier gas.

Different aspects must be considered for providing the maximum long-term deployment of the instrument. The carrier gas flow may be varied between 0 and 6 ml/min STP. The gas that is analyzed is stored inside the instrument. Since no miniaturization efforts were made for this prototype, the internal free volume is relatively large, estimated to 30 L, which allows to store a continuous flow of 6 ml/min of gas for 24h, before the pressure reaches 1.3 bar (absolute pressure) inside the instrument that would make the embedded vacuum pump no longer capable to provide the 30 mbar inside the measurement cell. The 40 bar reported in the manuscript represent a large excess than what is actually required, even with large carrier gas flows (with 6 ml/min STP of carrier gas 10 bar in 1L tank will already be enough for the 24h measurements). But for precaution, the tank is normally filled with a larger amount of carrier gas than what is required (eg. in case of small leaks in the high-pressure part of the line).

At low carrier gas flow, the amount of dry gas is very small, the mixture is mainly composed by water vapor which is trapped with silica gel before reaching the pump. So in the ideal case where all the water vapor is trapped and that the only gas reaching the pump is the dissolved dry gas (0.065 ml/min STP) then the autonomy of the instrument will be 90 days. This discussion has now been added on the manuscript:

"For an operation where the instrument is powered through an electromechanical cable the autonomy will be limited by the storage of the dry gas inside the instrument housing. For fast response measurements at maximum carrier gas flow of 6 ml min^{-1} this will correspond to 24 h autonomy, whereas without the use of carrier gas the autonomy will stretch to 90 days since most of the gas flow will be composed of water vapor that is trapped before the vacuum pump by the silica gel dryer (however, the long-term deployment may be limited by the capability of the silica gel)."

L185: general information about complementary sampling campaign and their published results should be shortened

Sorry, but we found difficult to make it shorter.

L214 N2 calculations from other measured gases is used in a figure and should therefore be part of the method section.

Changed accordingly

L221 Orange lines in what figure?

Figure 6. The information has been added.

L232 I don't understand this piece of information

We tried to make it clearer. The data at higher depths (>150 m) measured with the Contros HydroC sensor were about 12% lower with respect to the measurements by EAWAG and UFZ. We think that this is possibly due to the calibration of the sensor by the manufacturer but we do not have further information which could confirm this hypothesis.

Technical corrections

Sloppiness with the references does not increase the joy of reading this manuscript.

We are sorry to hear that the reviewer found missing references. But without further details on this remark we are not able to improve the manuscript accordingly.

I am no native English speaker myself, but many terms and expressions seem awkward (e.g., L86 "at the price of", L118 "electromechanical cable", L176, L191 "than", L236 "higher edge", L261 "seeing a background").

Those expressions have been changed/corrected (a part for "electromechanical cable and "higher edge" that we find correctly used).

Typos
L246 "O2 completely vanished"
L249 reasons
L244 e.g.

L433 replace "retrieved"

Typos errors have been corrected as suggested.