Response to comments by Referee #3

Referee comments are provided in grey and author responses in *blue italic style text*.

The authors are to be commended for working toward affordable instrumentation for trace gas measurement. In addition, information provided for construction of a lowcost datalogger can be useful for other environmental measurements as well. The authors take a primarily empirical approach to the problem of extending the range of usability of a commercial sensor. The goal is to be able to measure low levels that would be of interest in natural ecosystems. Toward this goal they describe a variety of curve-fitting calibrations.

It seems possible that many readers will find the most value in knowing how best to apply these sensors, and what ultimate performance can be achieved. This reviewer suggests that revisions focus mostly on optimum calibration procedures and the perfomance metrics that can be so obtained. The following are some specific comments addressing individual statements in the manuscript.

We thank Referee #3 for all work and are glad that the efforts towards affordable trace gas instrumentation are seen as valuable. We will try to highlight and clarify our optimization approach further, including that we focused on flux chamber use of these sensors and that this use have different calibration requirements than e.g. use in open air to follow atmospheric mol fractions. In the flux chamber applications, accurate determination of relative changes in gas mol fractions is more important than accuracy in determining absolute mol fraction values. Because of the flux chamber focus, we also calibrated the sensors when positioned in a flux chamber, i.e. at field-like conditions, and we tried to evaluate also simplified calibration procedures as a way to present different acceptable solutions being optimal for field-conditions and various access to laboratory analyses. We are well aware that this is not optimal from the absolute "maximum accuracy and precision" perspective, where it would be desirable to keep a stricter environmental control during calibration. Hence, we need to clarify that we, as also noted by Referee 3, have a more empirical and field-use oriented approach with the aim to facilitate reliable flux chamber measurements by as many sensor users as possible (which is different from assessing the maximum performance of the sensor under stable lab conditions).

Line 59 typo (is are)

Will be fixed.

L 63 It would be helpful to know what mixing ratios were successfully measured

This range is provided in Table S2 that will be cited here in the revised manuscript.

L 77 why were 10 RH sensors used? Was this to provide some averaging?

We had 10 sensors measuring CO2, T and RH in the same chamber as the CH4 sensors to also evaluate our sensor network solution for the CO2/T/RH sensors (separate study). Because we could not link values from any specific CO2/T/RH sensor to each specific CH4 sensors (they all shared the same chamber volume) we decided to just average their values and use this average in the evaluation of the CH4 sensors. We will clarify this in the revised manuscript.

L 84 How does measured RH compare with the vapor pressure of water at the given temps?

We used absolute humidity in the unit of ppm water vapor, corresponding to vapor pressure in µatm. The study covered a humidity range from 9000 to 35 000 ppm water vapor. At temperatures below 20

°C the RH was usually 50 - 70 %, while at temperatures > 20 °C, RH ranged from 18 to 60 %. Our result indicates that the absolute humidity was the most important factor, followed by temperature, and therefore we think the absolute ranges in these variables matters more than the RH values for the sensor tests. We will try to clarify this in the revised manuscripts.

L 89 typo (form) , superfluous 'before'
Will be fixed.

L 92 some further discussion of time response is called for, especially if some data points are to be removed from the analysis on the basis of delayed responses

Gas concentrations in the chamber with the sensors were impacting the sensors immediately and recorded at their respective logging frequency. It took some time for this gas to pass of the tubing and flush the measurements cell in the reference instrument. Hence, the reference instrument responded with some delay. When the concentrations were relatively stable or changed at slow constant rates this could be corrected for by considering the time offset. However, when concentrations in the sensor chamber changed rapidly relative to the time off-set, e.g. right after CH4 was added to the chamber or when the chamber was ventilated to reduce gas concentrations, it was not possible to correct for the time off-set. The reference instrument measurement cell was simply large enough to be influenced by CH4 from the chamber over a certain time period (the measurement cell residence time), and if the concentration change in the chamber is more rapid, the data from the reference instrument and sensors become incomparable and need to be omitted to not bias the calibration. We tried to explain this in the previous sentence but will consider further clarifications when revising the manuscript.

L 92 typo (therefor)

Will be fixed.

L 108 it is unclear why knowing RL is considered challenging, as 1% resistors are routinely available at low cost.

Referee #3 is correct that most of the variability between sensors are likely to regard the active sensor surface characteristics and that variability regarding RL is less likely. We just did not want to exclude this possibility, but given this comment, we will consider removing RL from this sentence.

L 123 this reviewer is not in a good position to assess the calibration approaches in detail, as they are mainly empirical and specific to these particular sensors (which this reviewer has not personally used)

Noted. Many thanks for good and important comments overall.

L 168 a few more details of the datalogger would be of interest to readers here. What is input voltage range and resolution (e.g. number of significant bits?). What other parameters would someone wanting to use this device in the field want to know? (see also below)

The Arduino based datalogger has an input voltage range of 7-12 V and a resolution of 10 bits. There is good documentation regarding logger board specifications on internet (e.g. https://store.arduino.cc/arduino-uno-rev3 and https://learn.adafruit.com/adafruit-data-logger-shield). We will check and make sure information about such links are available in the supplement.

L 176 typo (influence)

Will be fixed.

L 180 self heating is a very interesting issue. How much power is dissipated at the sensor surface, and would it be expected to produce heating that is significant relative to the uncertainty with which sensor temperature is constrained/influenced by environmental conditions?

The heating power consumption of the CH4 sensor is approximately 280 mW. This was not enough to notably change the temperature relative to other factors in the flux chambers tested so far. Not even under the experimental conditions when we had 20 sensors in a chamber with 7 L gas volume did we see any considerable heating effects.

L 220 It would be most helpful if the conclusions stated here were expressed more quantitatively, expressed perhaps in terms of accuracy, reproducibility, and long-term stability.

We considered this carefully and in principle agree. However, given that several different calibration models gave acceptable results and may be optimal for different conditions and different users, we would have to repeat rather large amounts of information to provide this information properly. In turn, this would lead to a rather long conclusions section. To try to resolve this situation we refer readers to Table 2 for quantitative information regarding different calibration strategies.

For long-term stability, we will add such information to the manuscript. Please see the response to the last comment below.

L 234 Good that code is provided for the datalogger!

Thanks.

Supplement
The datalogger may be of interest to many who plan on building their own field instrumentation. A more detailed circuit diagram, perhaps accompanied by a clear and more explicit image of the physical setup, would be helpful to those not experienced with Arduino.

We tried to provide a clear wiring diagram and illustrative images. We will consider how we can improve the visual descriptions of the system further in the revised supplement.

Fig S1 shows responses over ranges of several hundred ppm. It is suggested to also present data on an expanded scale at the lower end of the usable range.

We will add a similar figure covering the lower end of the range in the revised supplement.

Fig. S2 what are the units associated with RMSE in this figure. It gives the impression that acceptable results can not be obtained without using 8 or 9 reference samples.

The RMSE units are ppm, which will be clarified in the revised manuscript. The impression of Referee #3 is correct. A minimum of 8-9 reference samples are needed for acceptable results. In the text we choose to highlight that 20 reference samples would be preferable to reach even lower RMSE levels, but we will consider to relax this conclusion to 10 reference samples given Fig. S2 as suggested by Reviewer #3.

General: what are the authors' observations with regard to aging and long-term stability of these sensors?

We could not observe any tendencies of ageing during our studies so far, but we have added information about this based on other studies. van den Bossche et al., (2017; cited in manuscript) found no drift over 31 days. Eugster et al (2019) studied results from a very similar type of sensor used for outdoor measurements over 7 years and concluded that the drift was in the order of 4–6 ppb/yr and the variability drifted by –0.24%/yr. We will add this information in the revised study to address the drift question.

Reference:
Eugster, W., Laundre, J., Eugster, J., and Kling, G. W.: Long-term reliability of the Figaro TGS 2600 solid-state methane sensor under low Arctic conditions at Toolik lake, Alaska, Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2019-402, in review, 2019.