Reply on RC1

Rodrigo Vargas and Van Huong Le

Author comment on "The paradox of assessing greenhouse gases from soils for nature-based solutions" by Rodrigo Vargas and Van Huong Le, Biogeosciences Discuss., https://doi.org/10.5194/bg-2022-153-AC1, 2022

Response to reviewer

Comment: The authors present an interesting and novel evaluation of the bias inherent in sampling strategies at relevant timescales to capture estimates of GHG fluxes. As the authors point out, the ability to measure all three GHGs (CO_2, CH_4 and N_2O) simultaneously is now more common and has advanced understanding of the complex drivers of these important gases. Discrete, manual flux chamber sampling in which all three GHG fluxes from soils as tends to be the most common method, with good spatial but limited temporal representation. For convenience and cost effectiveness, discrete sampling strategies simultaneously measure all three GHGs, however, this strategy relies on the underlying assumption that each GHG responds similarly to biological and physical drivers at these same fixed temporal steps. Systems that automated the GHG flux sampling process are becoming more common but are still limited in application due to the costs associated with them, limiting spatial representation but providing high temporal sampling frequency. Automated, continuous measurement of all three GHG fluxes as high temporal frequency is better able to capture their temporal response to drivers that may not be co-occurring and offer a better understanding of the underlying drivers of each GHG flux as well as estimates of annual GHG budgets.

In this work, the authors aim to address how discrete manual flux sampling strategies in which all three GHGs are measured simultaneously at fixed temporal stratification (FTS) may violate the underlying assumption of co-occurring responses at temporal timesteps and bias the interpretation and understanding of each GHG. The authors utilize a dataset in which all three GHGs were sampled at hourly timesteps via an automated sampling system, for one year (Sept 2014-Sept 2015) in a temperate forest. By extracting subsets from this dataset at discrete timesteps, the authors create a series of examples of FTS at common sampling strategies (12, 24, 48 sample dates per year). The authors then utilize a novel technique, temporal univariate Latin hypercube sampling (tuLHs) to subsample the same annual dataset at the same temporal frequency (12, 24 and 48 annually). tuLHS optimizes the temporal selection of these subsets to reflect the same statistical properties and temporal patterns specific to each individual GHG reflective of the yearly GHG dataset. The authors argue that optimizing the sampling strategy for each GHG (tuLHS) is needed to avoid bias that may be inherent in FTS, particularly when the annual sample size is small (for example monthly, 12)

Response: We appreciate the detailed summary of this study by the reviewer.

Comment: The authors carefully show that measuring GHGs at common FTS biases estimates at annual timesteps, for this specific dataset, and that the tuLHS method produces a more representative reflection of yearly patterns of GHG fluxes providing a proof of concept for this novel method.

Response: We appreciate the supportive comment to recognize the novelty of our work.

Comment: This work is useful and informative and will provide a method (tuLHS) to aid researchers when developing a discrete manual sampling strategy for each GHGs. My concern is how easily this method is implemented broadly, either across years at the same site or how representative a tuLHS derived sampling strategy may be across similar ecosystems.

Response: We appreciate the supportive comment to recognize the novelty of our work.

We clarify that the main goal is to introduce the application of the tuLHS and show that the underlying assumption that each GHG responds similarly to biological and physical drivers may not be universal and should be tested. We provide a case study to introduce tuLHS as a proof of concept to show how the method works and to show that the general assumption that sampling at the same time all GHGs may not be appropriate to represent the probability distribution nor the temporal dependency of each GHG.

We propose that the application of the framework and the tuLHS can be applicable with any time series of GHGs, but we recognize that the specific results presented in this case study are specific from the statistical properties and temporal dependency of the time series analyzed. Predicting when to sample for a future year is extremely difficult but this method could be applied if a time series derived from a forecasting approach is available. If we assume that environmental variables would be similar for a future year, then the results generated from past information could provide insights about when to measure in the future. That said, forecasting sampling schemes is beyond the current scope of the present study.

In summary, we propose that the tuLHS approach is general and can be applied to different time series derived from multiple ecosystems. We will edit the discussion to address this comment and touch on the challenge of forecasting and potential applications of this approach for future sampling.

Comment: The authors acknowledge that the tuLHS method needs to be site specific, but a minimum of 1 year of automated continuous GHG fluxes (one without large data gaps) is needed to determine the optimal sampling strategy for each GHG using tuLHs. This also assumes that one year is representative of annual and interannual variation in each GHG flux patterns. Although this may be sufficient for CO2, CH4 and N2O are more variable at sub-daily to annual timesteps. A strategy developed in one year, may not be appropriate for the following year, especially if there are shifts in climate. It would seem that multiple years of site specific automated GHG measurements would be needed to determine if there are any wide variations in the optimal sampling strategy under different climate conditions.

Response: We appreciate the insightful discussion provided by the reviewer. These are important and interesting points that we can clarify in an edited discussion. Here we

provide responses on the main points that we would like to address in a revised version of the manuscript.

1- Our goal is not to prescribe a universal sampling time for each one of the GHGs, but to introduce the tuLHS approach and show that sampling all GHGs at the same time using discrete measurements may result in bias estimates. This is because a fixed temporal sampling is not able to capture the probability distribution nor the temporal dependency of each GHG when compared with automated measurements.

2- In theory the tuLHS can be used with any length of a time series as the method aims to optimize a sampling that represents the probability distribution and the temporal dependency. We present our case study with a 1-year time series as an example. That said, the longer the time series the better and if multiple years are available, then (arguably) the optimized sampling design could be more representative of the natural variability of the ecosystem.

3- The tuLHS could be applied to sub-daily time series but this was not the goal of the case study as we focused on daily time steps to simplify the example and present the case study. That said, this is possible, and the method could shed light about how to optimize measurements in a sub-daily time scale.

4- We agree that a strategy developed in one year may not be appropriate for the following year. We clarify that the goal of this study is to introduce the application of the tuLHS and show that the underlying assumption that each GHG responds similarly to biological and physical drivers may not be universal and should be tested. We previously discussed that forecasting a sampling design may be possible, but it is not the goal of the current study.

We postulate that there are several ways on how the tuLHS can be applied for future sampling designs. First, one could simply assume that the year tested is representative of the climate mean (probably not correct but a possibility) and therefore the strategy of one year can inform the strategy of the next year. Second, longer time series could be analyzed using the tuLHS and therefore the optimized sampling design may be more representative of interannual variability. Third, one could design a forecasting model and then use the tuLHS to inform the sampling design under that forecast scenario. We will consider these comments in a revised version of this manuscript.

Comment: Further, the tuLHS method may produce an optimal sampling strategy for each GHG, which logistically may be unreasonable to pursue given time, labor and cost constraints. To me this work highlights the need to either have an automated sampling system or co-locate automated and manual sampling strategies to truly capture temporal and spatial GHG fluxes from sites.

Response: We fully agree with this comment. First, co-location of automated and manual sampling strategies would be the best approach to capture the temporal and spatial variability of GHG fluxes from sites. Second, this study shows that the optimal sampling strategy is not to sample all GHG at the same time with a few discrete measurements. This is an important result because a few discrete measurements cannot reproduce the probability distribution nor the temporal dependency of the time series of GHG fluxes.

We recognize that adopting the sampling design recommended by the tuLHS will be logistically difficult but we show that a traditional fixed temporal sampling results in biased estimates. This has implications for quantification of the temporal variability and magnitude of GHG fluxes from soils. We will revise the discussion to make this point clear.

Comment: As a "proof of concept" in this site-specific case, the authors clearly show that FTS does produce bias in magnitudes and temporal patterns compared to tuLHS, which optimizes the sampling strategy, when compared to a one-year automated GHG flux dataset. More analysis, at multiple sites and conditions, is needed to ascertain the broad applicability of tuLHS.

Response: We thank the reviewer for the support of this study. We agree that this method should be tested in different ecosystems and will make it clear in a revised discussion.

Comment: I recommend minor revisions.

Response: We thank the reviewer for the support of this study.

Specific comments and questions:

- Was Sept 2014-Sept 2015 a typical climate (temperature and precipitation) year at the site? Can the authors provide insight on how deriving a sampling strategy from one year, particularly if it is not a normal climatic year, and utilizing that strategy in subsequent years may impact results?

Response: We have discussed the challenge of forecasting and we have emphasized that this is not the goal of this study. Our goal is to introduce the tuLHS approach and show that sampling all GHGs at the same time using discrete measurements may result in bias estimates.

We agree that a strategy developed in one year may not be appropriate for the following year. We previously discussed that forecasting a sampling design may be possible but it is not the goal of the current study. There are several ways on how the tuLHS can be applied for future sampling designs. First, one could simply assume that the year tested is representative of the climate mean (probably not correct but a possibility) and therefore the strategy of one year can inform the strategy of the next year. Second, longer time series could be analyzed using the tuLHS and therefore the optimized sampling design may be more representative of interannual variability. Third, one could design a forecasting model and then use the tuLHS to inform the sampling design under that forecast scenario. We will consider these comments in a revised version of this manuscript.

- Do the authors think there was any influences in results due to missing automated GHG flux data, which appears predominately in the winter- early spring? It seems curious for N_2O to have tuLHS select predominately in the fall/winter period as representative of annual N_2O flux temporal and statistical characteristics.

Response: Missing data due to quality assurance/quality control, electrical power or mechanical failure are common in automated measurements. The tuLHS will optimize the sampling approach based on the statistical properties of the time series, and our assumption is that the data presented is representative of the statistical properties of the year of measurements. We will discuss this assumption in a revised version of the manuscript.

In the case of N2O, its variogram shows that the temporal variability is constant, that is, there is no temporal dependence. The results were selected within optimized days located in winter/fall, but because the statistical properties of this time series, the selection of these days may not vary much if those alternative days are representative of the statistical properties of N2O fluxes.

We clarify that if there are substantial missing data, then the statistical properties of the time series will change and consequently the results of tuLHS approach. We reiterate that the main goal of this study is to introduce the tuLHS approach and show that sampling all GHGs at the same time using discrete measurements may result in bias estimates.

- Lines309-311: the results show that tuLHs provided closer estimates of cumulative sums and uncertainty ranges than FTS. Were these estimates significantly better?

Response: Figure 5 show the result of how the different sampling designs influence the cumulative sum and uncertainty ranges. We did not perform a formal test for significant statistical differences, but in all cases the annual sum and uncertainty ranges derived using the tuLHS are closer to those from automated measurements (Figure 5). We highlight that sampling N2O using FTS results in the largest bias (150%) in cumulative sums.

- Overall, since the means for FTS and tuLHs were not statistically different, if a researcher's goal is only to estimate an annual GHG flux, is FTS, particularly at bi-weekly time steps, a sufficient strategy?

Response: The means from FTS and tuLHS were not statistically different but that does not imply that cumulative sums nor uncertainty is similar (see Figure 5). Our results show that the cumulative sums and uncertainty derived from FTS are biased for all GHGs (Figure 5). The tuLHS approach consistently provided closer estimates for cumulative sums and uncertainty ranges than FTS for all GHG fluxes. We will revise the wording in the manuscript to emphasize the results of Figure 5.

Technical corrections/comments:

- In the graphs the authors use Time (days) from 1-365. I assume that is DOY and 1 is Jan 1. The data collected by the automated chambers is Sept 2014-Sept 2015 and I just want to clarify that day 1 is not Sept 2014 and the year follows that timeline.

Response: The dates are from September 12, 2014 to September 11, 2015. We use the DOY and day 1 is January 1. We will clarify this in a revised version.

- Figure 2: The blue line is very difficult to see. Perhaps make the open black circles smaller, thicken the horizontal lines for better clarity.

Response: We will edit Figure 2 following the comments from the reviewer.

- Figure A1, A2, A7 and A8: These figures are too small to read when printed.

Response: We will edit these figures to make them larger and easier to read.

- What program did the authors use to apply the tuLHs to their automated dataset and can they provide that code alongside their already referenced dataset?

Response: We used the program language: R and the integrated development environment: RStudio and the code will be available in a GitHub repository. We will add a link to the repository in the revised version of this manuscript.