Dear Paul Stoy et al.,

I have now received the third review for your revised manuscript on methane emissions from bison. While the reviewer is not completely convinced by the responses you have given in your previous response I am confident the currently remaining issues can be clarified, particularly in terms of explaining which errors you are showing based on how many data points and how these were derived. Similarly, I suggest to address the importance of specific biological parameters on CH4 emissions from ruminants (ie. diet quality not just quantity). Of course, there aren't many studies on bison available and this seems to be a scoping study to extend future measurements, I would like to point towards comparing your results to either "similar" species such as buffalo or cattle that are grazed even during colder periods in the year. You may further think about stressing more strongly that your system is a mix between a grazed and a feedlot system, given the supplementation with hay.
With this, I am looking forward to receiving your revised manuscript.

with kind regards

Lutz Merbold

Dear Dr. Merbold,
Thank you for reconsidering the manuscript. We made a number of additional changes to the manuscript which we believe improved the analysis and its presentation, and also added a short Conclusions section to the Discussion and a Land Acknowledgement to the Acknowledgements section. We added new references on ruminant nutrition and methane efflux and revised the Discussion section to explain our comparison with other literature estimates and to frame our results in a broader context. We thank you for your support of the manuscript and please do not hesitate to write if questions arise; we would be happy to further reconsider any analysis and were relieved that our efforts in comprehensively re-analyzing all values did not result in meaningful changes to our original study.

Sincerely,
Paul C. Stoy

Review of manuscript bg-2020-38-manuscript-version5

This is my second review of the manuscript and the third review round in total. I am not satisfied with the revisions and the response of the authors to my previous comments. The authors decided to rebut most of my major comments and their response contains, in my view, partly insufficient or inappropriate arguments. In the following I list the remaining issues and requested improvements.

We have endeavored to make the suggested changes and are grateful for the suggestions which improved the manuscript. We disagreed with some points from the previous review that we *have now added to the manuscript for completeness. One point regards the Conclusions section which we argued is a matter of preference. We added a succinct conclusions section that highlights the key findings of the manuscript and agree that it helps frame the analysis. We also now adopt the Referee's suggestion change the discussion of the equations while still ensuring that we credited the studies that inspired our approach. Another important point that was raised before is the issue source height. Source height would have been interesting to include if we had a basis for determining it, but we had no observational basis for doing so and felt that it would add unnecessary speculation into the analysis. This is one of the most cautious flux studies that we are aware of and we are trying to be very careful to not introduce unnecessary uncertainties. The third regards the z0 calculation. We felt that it would be inaccurate to change our approach, which was designed for the specifics of the study field and measurement period. We do note that the median z0 value that we arrive at with bison,*

MAJOR COMMENTS

The uncertainty of the main result, the average per-bison-emission, is not treated appropriately. In the new manuscript version there is even a new aspect on this issue. The authors give the results as "mean ± standard deviation" (e.g. abstract, line 249, ...). It is unclear what that means (standard deviation of what dataset?). Also a standard deviation is usually not a useful uncertainty measure. This needs to be clarified.

We summed variances calculated from multiple independent sources of uncertainty. We agree with the Referee that we did not communicate uncertainty correctly for two reasons. One is that we maintained both mean and median values after propagating uncertainty. We should have only kept the mean because large values are 'averaged out' to a degree. The non-normal distribution of methane flux observations that we show in Fig. 11a are sensible in our opinion because we expect intermittent large pulses of methane efflux. It is common to sum observations from non-normally-distributed eddy covariance observations to arrive at temporally-summed observations.

In response, we decided to retain standard deviation when presenting the average fluxes with and without bison because we felt that presenting observations in this way is actually more conservative. Had we used variance, for example, the average methane flux without bison would be −0.0009 ± 6.4e-5 µmol m^{-2} s^{-1} and had we used standard error of the mean this value would be −0.0009 ± 0.0002307 µmol m^{-2} s^{-1}. Needless to say we are more than happy to present values in this way if you prefer.

We re-ran all analyses (at a relatively large amount of computer time given our stochastic approach for estimating bison location uncertainty) and feel that our results are accurate but as the Referee notes should have been presented more clearly. We have revised the discussion of mean fluxes and their sums as a result to be even more conservative with our analysis as described in more detail below.

- In response to my previous comments, the authors have added some text statements about additional uncertainty sources, but at the same time, they apparently have reduced the uncertainty estimation, instead of increasing it.

This is correct. Uncertainty need not monotonically increase every time that it is re-evaluated.

- line 198: The 17% uncertainty "for longterm sums" adopted from the literature can hardly be used for the present extremely non-homogenous situation and a very limited measurement time of only daytime cases during only about 3 winter weeks (bison present and camera pics available).

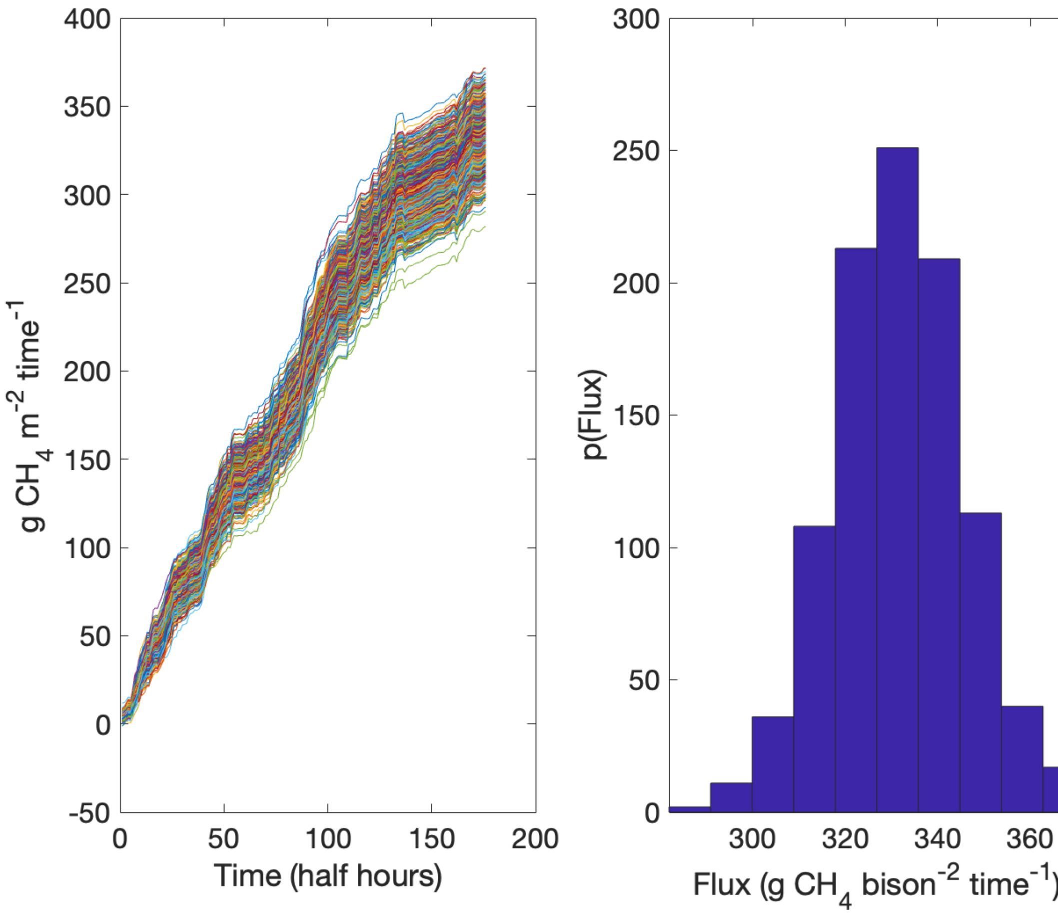
Random eddy covariance uncertainty is usually assumed to be on the order of 10-15% for carbon dioxide fluxes following for example Goulden et al., 1996, doi: 10.1111/j.1365-2486.1996.tb00070.x. The uncertainty of flux sums is smaller than individual measurements because random error is averaged out. We had used 17% as the most conservative value for flux sums that was supported by the literature, and we agree that 'long-term' is subjective and removed it. In the revised analysis we use the 41% value from Deventer et al. (2019) to estimate uncertainty due to eddy covariance observations and find that they now make a slightly larger contribution to total uncertainty (25%) but not enough to change our conclusions. To be honest we are more concerned with potential bias uncertainties in all eddy covariance observations than random uncertainties and members of our team have been very active in addressing the eddy covariance energy imbalance challenge and other potential sources of bias uncertainty, which may be important in all flux studies but difficult to quantify.

- It is also crucial to declare in the manuscript, how many half-hourly per-bison-emission values were available after all quality filters.

After all filters were applied, 158 half-hourly observations with bison in the flux footprint remained when applying the Hsieh et al. (2000) footprint model and 146 observations were available when applying the Kljun et al. (2015) footprint model. In both instances this is over 100 more measurements than the groundbreaking work of Galbraith et al. (1998) who measured five penned bison with seven replicates. These are also the first field observations made of a species of critical importance to the culture of multiple Indigenous Tribes of North America and their duration was a consequence of respect for the private landowner who allowed us to make these measurements. We now report these values in the manuscript.

- Figure 11A shows that the individual per-bison-emission data have a strongly skewed distribution. Therefore the random-like error cannot be well estimated according to Gaussian statistics rules. However the difference between arithmetic mean and median is an indicator of a large uncertainty.

We are unaware of flux observations that follow a normal distribution and the sum of these values will approach a normal distribution due to the Central Limit Theorem. For example, the figure below demonstrates 1000 realizations of the cumulative sum of observed per-bison *methane flux with a random 41% uncertainty about each observation, which the upper end of the uncertainty established by Deventer et al. (2019).*



The cumulative sums, right, are normally distributed. We would agree that our observations are not strictly speaking independent as the central limit theorem requires, but we feel that it is a reasonable approximation.

Performing this analysis inspired us to reinterpret the uncertainty analysis using approaches that are perhaps a bit more established in flux science. We calculated the mean percent uncertainty for the spatial uncertainty analysis, the flux footprint analysis (without spatial uncertainty), and the 41% conservative uncertainty estimate of flux sums following Deventer et al. (2019). We used these percent uncertainties to calculate a variance for each observation, then drew 1000 samples from each of these distributions to create different realizations of the sum of fluxes as demonstrated above. We then calculated 95% confidence intervals about the flux sums and present these in the revised manuscript rather than standard deviations or variances. The percent uncertainty that we now calculate for each term differs slightly from previous uncertainty estimates as a result, but results do not change our conclusions.

B) I consider the discussion of the per-bison-emission results in comparison to the literature as still insufficient. In the response to my previous comment 6, the authors state that "Methane flux is related to the animal in question, its body mass, diet, metabolic state, pregnancy / weaning status, and more." But for the literature comparison they just selected"...results that are similar to ours". This is a clearly non-scientific approach. The authors should not just select literature per-animal emission values that are similar to the present study without considering/stating the relevant factors (body mass, diet, etc.) in the referenced studies.

We are first and foremost interested in ensuring that our observations are reasonable with respect to other ruminant systems: a simple logic check. We re-worded the text to be more clear on this point. This logic check follows the work of Galbraith et al. (1998) who also noted that bison methane flux are to a first order similar to cattle (at least when fed alfalfa). There are no existing measurements of methane flux from bison in a natural setting and we had to resort to comparisons with non-native introduced cattle grazing systems as a consequence.

There are hundreds of studies on cattle and for good reason; they are critical to the global methane cycle but are also favored by European management systems, which in North America is an introduced agroecosystem. It is of course of interest to synthesize such studies, but this exceeds the scope of the present analysis because it is a presentation of new results and not a review article. Studying a grazing system that is of cultural interest to Indigenous People is of scientific value and we wish to understand how these grazing systems compare to conventional grazing systems that tend to dominate the published literature.

In response, we added information to the Discussion section that notes the unique pasture & feedlot characteristics of the study field including new references from the literature on methane efflux.

Since the authors compare their results to feedlot studies, it is also not clear whether the present experiment is considered as representative/comparable to a grazed pasture system or rather to a feedlot system. This should be clarified.

The present study shares elements to both grazed pasture and feedlot systems; the animals are fed supplemental hay on a pasture in winter. We made extensive edits to the Discussion section and note that our study system shares elements to a grazed pasture and feedlot as noted.

MINOR AND LANGUAGE COMMENTS

C) Response to prev. comment 3:

Since the authors introduce their z0 determination in detail in Section 2.4 and because the z0 values are important inputs for the footprint models, it is surly warranted that the authors present some corresponding results in the text (e.g. range of obtained z0 values with/without animals in the footprint).

We now present median z0 values for measurements with and without bison.

D) Response to prev. comment 9:
The rebuttal indicate that the authors principally put the use of a 'Conclusions' section into question, because their arguments could be applied to most scientific papers. I do not agree but I leave it to the editor to decide this issue.

We feel that the end of the Discussion section adequately summarizes results and we agree that this is a personal preference, but we added a Conclusions section for completeness. We agree that it helps make the manuscript come “full circle”.

E) line 22: "...similar to eddy covariance measurements of methane efflux from a cattle feedlot during winter". It needs to be clarified here, that the mentioned feedlot results were not obtained in this study but that you mean "previously reported eddy covariance measurements ..."

This is correct. Our results are compared against other studies that arrive at similar values to gain insight into why the methane flux measurements may be similar. We re-worded the passage.

F) line 163: define the meaning of Fij in Eq. 4

F_{ij} is the contribution of the flux of scalar X from grid cell i,j. We changed the presentation of the equations per the Referee’s suggestions in the revised manuscript, which improved the discussion of the equations because it avoids the use of ‘phi_{source}’, which we did not define adequately in the previous manuscript.

G) line 321: It needs to be clarified whether the 36% is a attribution fraction or a relative uncertainty.

Thank you for noting this, it is the contribution to total uncertainty.

H) Figure 10 and 11A: The use of the term "probability" is strongly misleading here. The shown data rather are observed frequencies of occurrence.

The figures show kernel density estimates, which are non-parametric estimates of probability distributions. They are probability distributions. This is not intended to be misleading and we re-worded the passage to help ensure that it is clear for the readers.