Anonymous Referee #1

Received and published: 12 August 2019

The paper by Webb et al presents CH4 and CO2 data from 101 farm ponds. Alongside these GHG measurements are an impressive array of variables of water chemistry, hydrological characteristics, and landscape attributes. The authors investigate these variables as drivers of the GHG emissions. The paper is well written and I enjoyed reading it. It is within the scope of BG, and presents novel data insomuch as the fact that more pond GHG data is needed (and this point was explicitly raised in the recent IPCC refinement). If small, artificial waterbodies can be designed to minimise CH4 emissions, and to act as CO2 sinks, then this could lead to them acting as natural climate solutions.

Methods and analysis are well explained with sufficient detail, and the results support the conclusions. Presentation is good, language is fluent, abstract is suitable. The work is mostly well referenced (I suggest two older references of farm pond emissions that the authors may have missed). I particularly enjoyed reading the succinct and to-the-point results section, which was enough to get the authors' points over without endlessly writing numbers out, as so many results sections do. The one thing I find lacking from the paper is a visual presentation of the underlying CO2 and CH4 data, and in my comments I suggest a way to address this. I think it is important that readers are offered an easy way to understand the variation in the GHG data across all 101 waterbodies. I suggest the paper is acceptable following minor revisions. Below are my detailed comments.

Response: We thank the reviewer for their positive review and their constructed comments and suggestions offered. Detailed responses to the comments are addressed in blue font below.

L29. "Small waterbodies have recently been recognised as substantial contributors to global carbon emissions from inland waters." This is true, and missing from somewhere in the introduction (and discussion) is a mention that the recent 2019 IPCC Refinement explicitly addresses the issue of CH4 emissions from artificial ponds. The Refinement can be found at the link below, and the relevant chapter is in vol. 4 (AFOLU), chapter 7 (Wetlands). The emission factor given for artificial ponds is 183 kg CH4/ha/yr, but there is currently not enough data to disaggregate pond emissions by climate zone. How does your data compare to this emission factor? https://www.ipccnggip.iges.or.jp/public/2019rf/index.html

Response: We appreciate the reviewer raising awareness of the latest IPCC estimate. The following sentence has now been added to the introduction:

"*The recent 2019 IPCC Refinement has assigned a CH_4 emission factor of 183 kg ha^{-1} yr^{-1} to constructed waterbodies, however data is greatly limited, both geographically and in number (n = 68), that climatic-zone emission factors cannot be estimated (IPCC, 2019).*" Line 61

We also now compare our average farm dam CH_4 emission with the IPCC estimate in the discussion:

"Average CH_4 fluxes from our farm reservoirs correspond to 417 kg CH_4 ha^{-1} yr^{-1}, which is greater than the current IPCC emission factor estimate of 183 kg CH_4 ha^{-1} yr^{-1} (IPCC, 2019). Considering the skewness of our CH_4 data, our median value of 184 kg CH_4 ha^{-1} yr^{-1} agrees with the emission factor of other artificial ponds." Line 368

L36. It's worth noting the recent paper by van Bergen et al who measured CH4 (including ebullition) and CO2 emissions, and C burial of an urban pond. Ideally we need studies that quantify GHG emissions and C burial, so the net balance can be calculated. van Bergen, T.J., Barros, N., Mendonça, R., Aben, R.C., Althuizen, I.H., Huszar, V., Lamers, L.P., Lürling, M., Roland, F. and Kosten, S.,

2019. Seasonal and diel variation in greenhouse gas emissions from an urban pond and its major drivers. Limnology and Oceanography.

Response: The van Bergen reference has now been added to the following sentences in the introduction.

"Artificial reservoirs have the potential to be potent sources of CO2 and CH4 (Downing et al., 2008; Holgerson and Raymond, 2016). This can be demonstrated by a carbon budget estimate from an urban pond where carbon emissions (both diffusive and ebullitive for CH_4) offset carbon burial by >1,000% (van Bergen et al., 2019)." Line 59

L60. "Currently, only three studies have comprehensively assessed C fluxes from small agricultural reservoirs." What does "comprehensively" mean in this case? These three studies are slightly different – Ollivier et al did not measure ebullition whilst the other two studies did. Ollivier et al and Paneer Selvam et al were 'snapshot' studies whilst Grinham included some temporally repeated measurements (but didn't measure CO2). So are they all comprehensive really? I accept this is a minor point of language but it does matter. Additionally, there are two other papers that have measured farm ponds. Stadmark et al made repeated measurements of CH4 and CO2 emissions from agricultural ponds created to retain N: Stadmark, J. and Leonardson, L., 2005. Emissions of greenhouse gases from ponds constructed for nitrogen removal. Ecological Engineering, 25(5), pp.542-551. There is also data in an old and rather blandly titled paper from two farm ponds. Baker-Blocker, A., Donahue, T.M. and Mancy, K.H., 1977. Methane flux from wetlands areas. Tellus, 29(3), pp.245-250. L62. "Large fractions of CH4 being released." Fractions seems like an odd and unsuitable word. Change for "volumes", "amounts", "quantities", etc?

Response: We have removed "comprehensively" and replaced with "at regional scales" in the sentence which now reads:

"Currently, only three studies have assessed C fluxes from small agricultural reservoirs at regional scales and these support the notion that they are important landscape sources of GHGs (Panneer Selvam et al., 2014; Grinham et al., 2018a; Ollivier et al., 2019)." Line 64

Because here we are referring to studies with a high number of sites spanning a regional scale, we will not refer to the other two studies mentioned given they only measured a couple of sites.

L80. The study region occupies a large area, but seeing as temperatures are given it would also be good to give a value (or range) for annual precipitation. Reading on, I see the results says "precipitation ~60% less than the long-term climate average of 390 mm in Regina." Please give the value in the methods.

Response: The following sentence has been added to site description:

"Average annual precipitation in the area ranges from 354 to 432 mm." Line 89

L86. It says 101 ponds were sampled, but in table 1 some variables have N = 102. Where does 102 come from?

Response: We did sample 102 sites but lost GHG measurements from one. Because we are focusing of CO_2 and CH_4 samples in this study, we will refer to total number of sites as 101 and replace 102 in Table 1.

L113, L118. Floating chambers are not "incubations". This word should be altered to something like "deployments" or similar. L121. It says DO was measured in mg/l but in table 1 it is given as %. The methods text should be amended to % instead.

Response: "Incubations" have now been replaced with "deployments". DO units have also been amended to read % saturation in Methods text.

L149. Inflow is mentioned here. Do these systems have inflows? Is water pumped in for storage, or do they simply collect rainwater?

Response: With the water isotope mass balance method, inflow here refers to precipitation, snowmelt, and groundwater inputs. These farm reservoirs are designed collect most water than falls on the landscape due to being positioned in depressional area.

L183. "To avoid multicollinearity, correlation coefficients between pairs from Pearson linear correlation tests was used to guide covariate choice before model fitting." This is vague. Did you use a Pearson correlation coefficient of a certain value to decide when multicollinerity was present?

Response: Here if the correlation was significant then it was decided that multicollinerity was present. We have added that detail to the sentence, which now reads:

"To avoid multicollinearity, correlation coefficients and statistical significance (p <0.05) between pairs from Pearson linear correlation tests was used to guide covariate choice before model fitting (Table S1-3)." Line 199

L197. Something I desperately miss from the paper is a figure allowing the reader to visualise the raw CH4 and CO2 data and its distribution. I strongly advise the addition of a figure to show this. It could take numerous forms, such as a scatter plot of CH4 vs CO2 for all 101 ponds, or a box plot of GHGs (grouped by pond size, or pasture vs cropland), or even a bar plot showing individual concs for 101 ponds (large and unwieldy perhaps, but visually useful). Reading on I see figure.3 has a very small land-use graph, but I think a more obvious, up-front figure would be better.

Response: We have now added a figure (Figure 2) to illustrate the distribution of CO_2 and CH_4 concentrations across all sites.

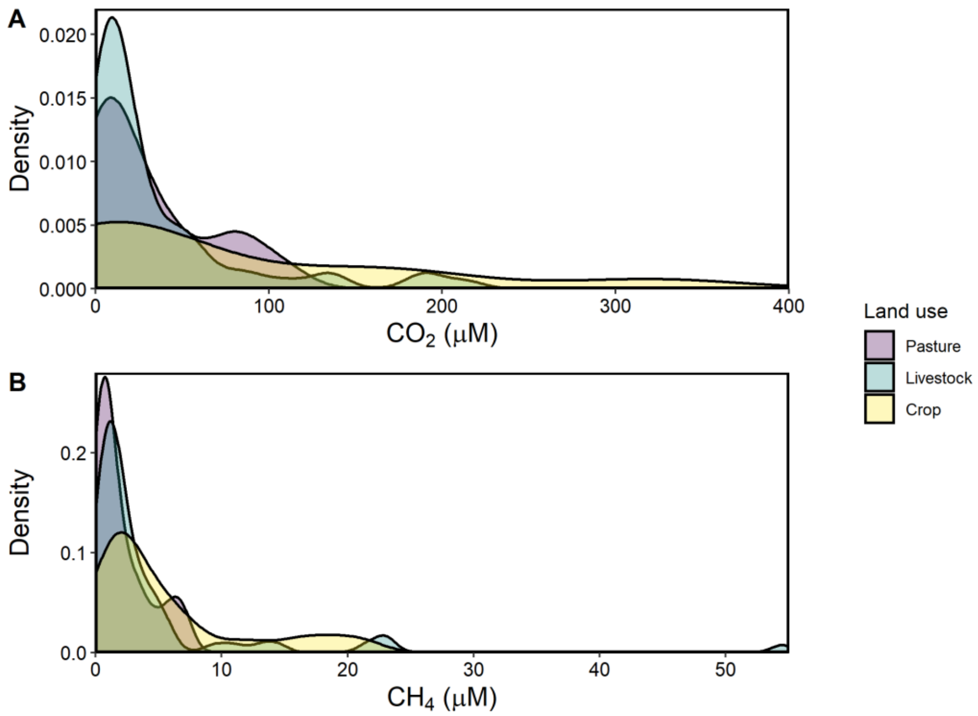

Figure 2: Kernel density estimates of CO_2 and CH_4 concentrations measured in 101 farm reservoirs grouped by land use.

Fig 2 and fig 3. In part this relates to my point above. Wouldn't these figures be improved by adding the underlying data points on to these figures as a scatter? That way the reader can see the model, and the raw data. It would help the reader visually determine the robustness of the models.

Response: While we understand where the reviewer is coming from regarding underlying data points, we chose to avoid adding these here as adding raw data to partial effects plots of GAMs does not provide a meaningful way to represent model fit. These figures illustrate the partial effects transformed on the response scale and the fitted relationship between each covariate and the response is affected by all covariates in the model. Instead, we have now provided diagnostic plots in the supplementary material (Figs. S2 and S3) to allow readers to visually assess the robustness of each model. One of these plots shows the observed versus predicted values of our CO_2 and CH_4 concentrations with the model.

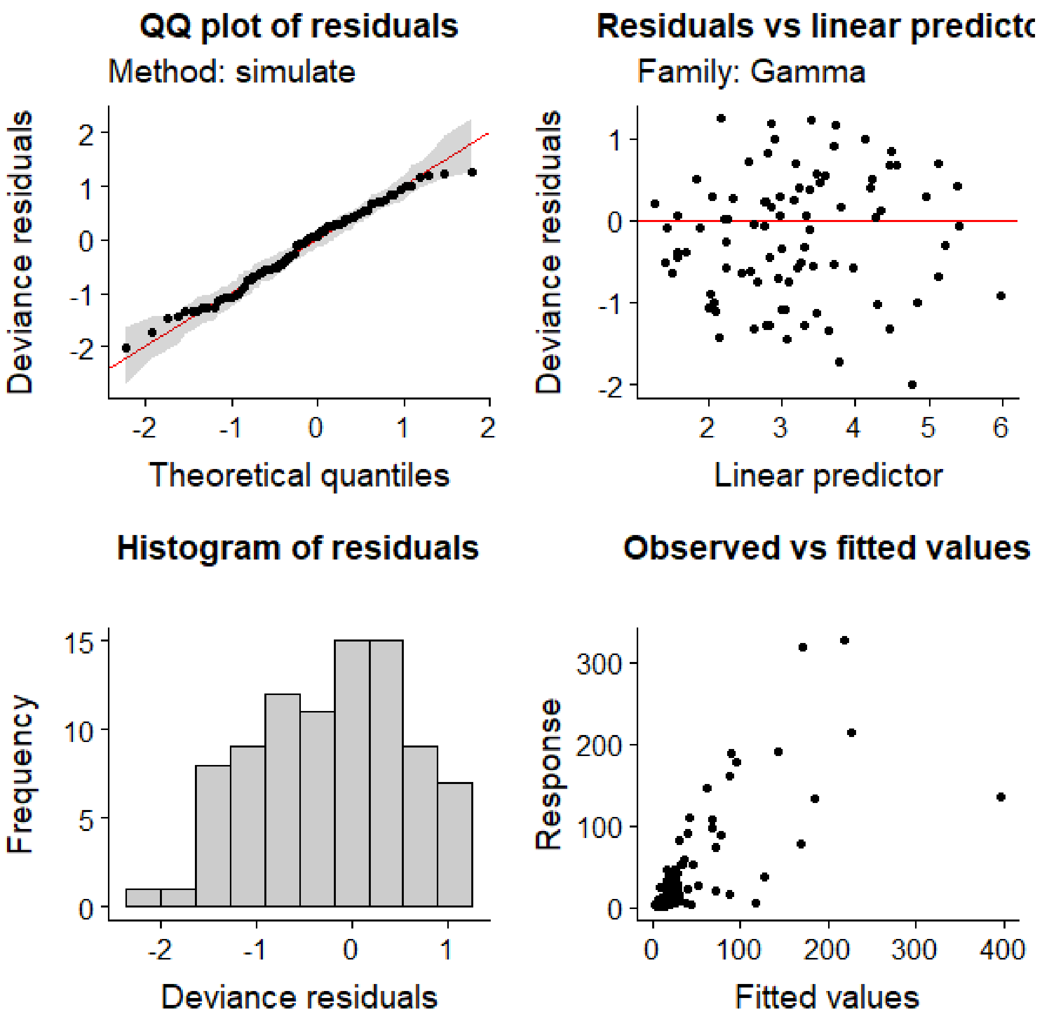

Fig. S2: R output of diagnostic plots for carbon dioxide model

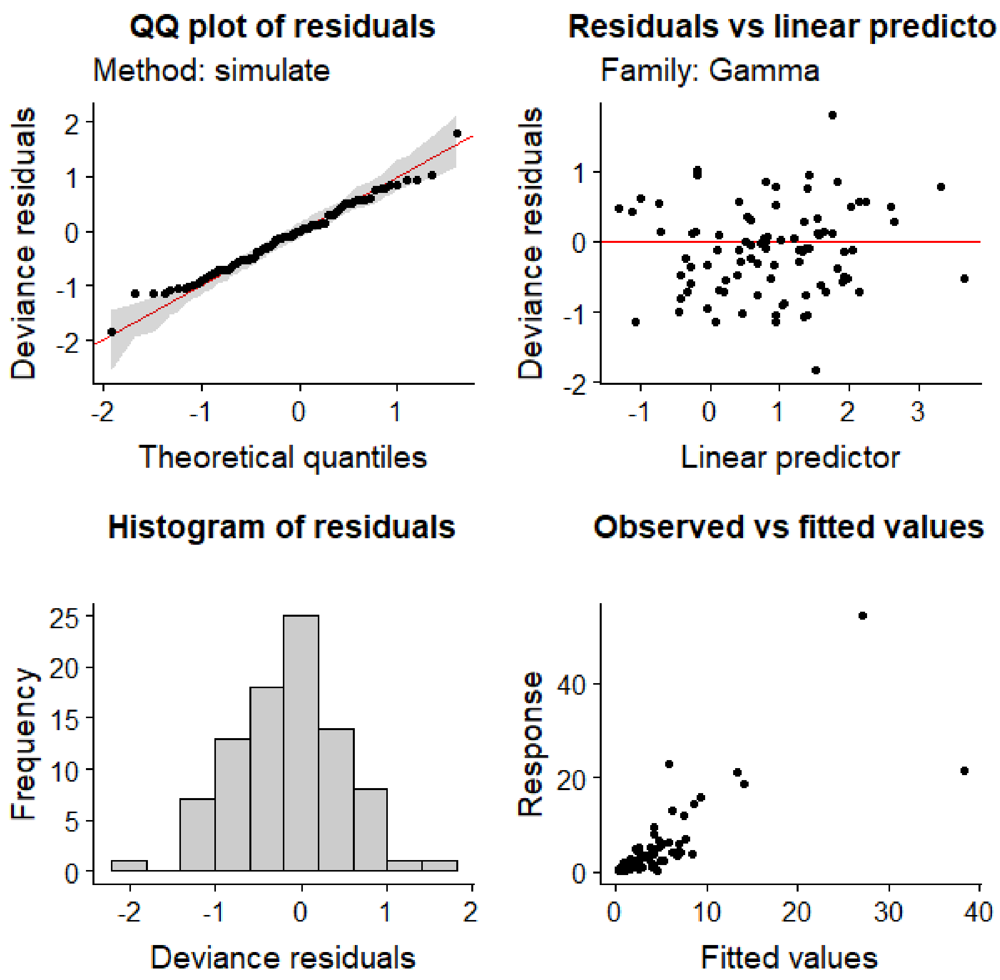

Fig. S3: R output of diagnostic plots for methane model

L210. "CO2concentrations displayed a positive response with. . .NOx" Whilst the upper 95% credible interval continues to increase, the black line presumably suggests that CO2 decreases at the highest NOx levels. Is there a mechanism that can explain this? Figure 3 has a land use graph, but figure two doesn't. Even if there is no difference in CO2 between land use a figure would still be interesting to see, and there is room for an extra panel at the bottom right anyway. For the land use panel in figure 3, the categories are pasture, livestock and cropland. However, line 87 in the methods only mentions pasture (n = 80) and cropland (n = 21). Where do these livestock ponds come from?

Response: We try to avoid over interpreting the tail-end of model responses where credible intervals increase, as this is often governed by less data at those extreme ends. The addition of supplementary figures S4 and S5 shows the distributions and correlations between covariate pairs to demonstrate this.

Fig. S4: Scatterplot matrices of covariate data used in the CO_2 model showing distribution and correlation pairs

Fig S5. Scatterplot matrices of covariate data used in the CH_4 model showing distribution and correlation pairs.

A plot for CO_2 land use model results, Figure 3I, has now been included. We have also corrected the definition of land use types in methods which now mentions livestock:

"We sampled 101 farm reservoirs between July and August 2017, ranging in surface area from 158 – 13,900 m^2 (Table 1), including basins in pasture (n = 18), pastures with livestock (n = 62) and cropland (n = 21) sites." Line 93

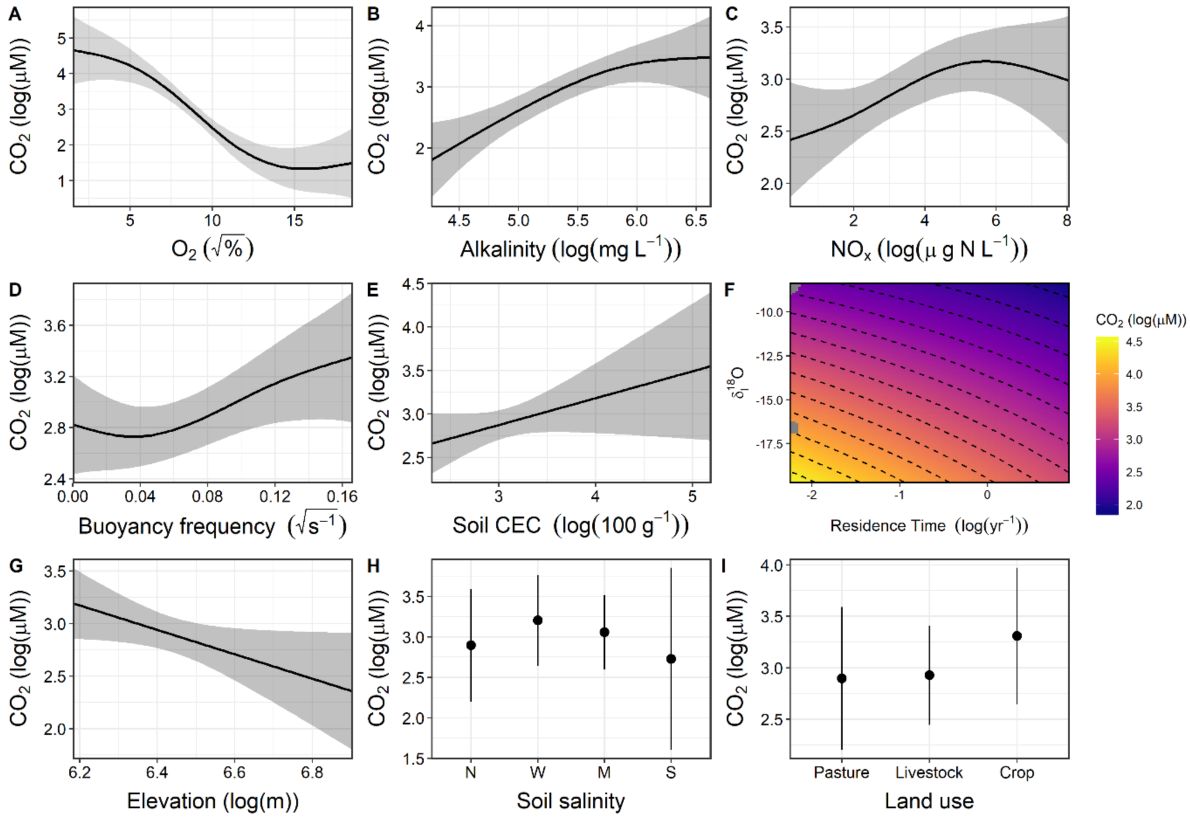

Figure 3: Response patterns farm reservoir CO_2 concentrations with abiotic, biotic, hydromorphological, and landscape variables based on GAMs. CO_2 was best estimated by a combination of a) DO saturation, b) alkalinity, c) NOx, d) buoyancy frequency, e) interaction between δ_I and WRT, f) soil CEC, g) and elevation, with soil salinity (h) and land use (I) not significant. Model deviance explained was 66.5%. The response patterns shown are the partial effect splines from the GAM (solid line) and shaded area indicated 95% credible intervals. See Table S4 and Figure S2 for summary of model statistics and model fit with observed data.

L224. "Our comprehensive spatial analysis revealed wide variations among CO2 and CH4 concentrations between farm reservoirs" As per my previous comment, there's currently no easy way to assess this until the raw data is more visible in a figure.

Response: A new figure (Figure 2), has been provided as suggested previously and is now referenced in that text.

L227. "CH4 was most correlated by internal abiotic and biotic mechanisms" Should this not be "most correlated with"?

Response: Corrected.

L282. "Additionally, smaller waterbodies with shorter WRT can support higher rates of internal CO2 production due higher rates of allochthonous DOC mineralisation" Needs amending to read "due to"

Response: Corrected.

L285. "This mechanism is also suggested by the observation that higher reservoir CO2concentrations are predicted in high CEC soils Alkaline high CEC soils retain more calcium ions within clay particles which releases carbonates and bicarbonates into soil porewater" It seems like something has gone awry in the writing here, and this should be two sentences or some words need removing.

Response: Yes, this sentence should be separated into two. This has now been corrected.

L331. "The effect potential effect of sulfate" The first "effect" needs deleting

Response: Corrected.

L336. "In contrast to the external drivers found for CO2, local land use had a significant effect on CH4 concentrations in farm reservoirs (Fig.3I), with significantly higher CH4 levels in cropland waterbodies than those in pasture. This finding contrasts with those from Australian farm reservoirs where diffusive CH4 fluxes were 250% higher in reservoirs with livestock compared to crops," I find this section of the discussion interesting. As the authors write, the intensive agricultural practices associated with cropland could be expected to result in elevated CH4 concentrations. Conversely, pasture/livestock emissions would depend on the system (intensive or extensive), livestock, etc. Intensive grassland systems could easily result in high emissions, whilst low-level grazing might result in emissions being less than those from cropland. So cropland > grassland and grassland < cropland are both explicable it seems to me.

Response: We agree that for all land use types, the intensity of agricultural production likely governs the effect on methane in the reservoirs, perhaps more so than land use type itself. Although assessing the intensity of each land use is beyond the scope of this research, we have expanded this section of the discussion with mention to livestock intensity:

"Our finding contrasts with those from Australian farm reservoirs where diffusive CH_4 fluxes were 250% higher in reservoirs with livestock compared to crops, although the mechanisms responsible for observed differences were inconclusive (Ollivier et al., 2019). This difference could be the result of the intensity of agricultural production, where farm reservoirs supporting high intensity grazing may also experience high CH_4 production as demonstrated by a couple of high CH_4 concentrations observed in our livestock pasture reservoirs (Fig. 2). In this case it's likely that CH_4 levels are more influenced by nutrient loading from the landscape which stimulates eutrophication (Huttunen et al., 2003), as suggested by the biotic variables in our model (Fig. 4). The intensity of agricultural production under different land use types should be an area of further exploration for external controls on farm reservoir GHG production." Line 357

Figure 4 and fig. 5. The study by Grinham et al of Australian ponds is referenced in the text but doesn't seem to be included in these figures. Is there any reason their data was left out?

Response: The Grinham et al., 2018 study is included in Figure 4 under "Artificial ponds, Australia" for the CH_4 fluxes. We now realise this reference is not included in the supplemental table referred to in the figure caption. Reference details to this study is now included in Table S6.

L365. "Here, CH4 fluxes were converted to CO2-efluxes using the sustained-flux global warming potential over 100 years" I am not familiar with this metric, and suggest a few lines are included in the methods as to what it is and how it is calculated.

Response: We have added details to how the CH_4 fluxes were converted to CO_2-equivalent fluxes in the methods:

"For comparing CO_2-equivalent fluxes, CH_4 fluxes were converted using the 100-year sustained-flux global warming potential (SGWP, Neubauer and Megonigal, 2015). This metric offers a more attainable measure of ecosystem climatic forcing, assuming gas flux persists over time instead of

occurring as a single pulse as quantified using traditional global warming potentials (GWP, Myhre et al., 2013). Here, a SGWP multiplier of 45 was applied to all CH_4 fluxes in the literature comparison, which is slightly higher than the traditional GWP of 32 over a 100-year time frame (Myhre et al., 2013).” Line 130

Section 4.4. What (if any) vegetation colonises these pools? Is there no role for encouraging certain plant species that might promote C uptake? For instance, Moore & Hunt say: “The carbon sequestration assessment of constructed stormwater wetlands and ponds suggests that emergent vegetation is a significant source to the soil carbon pool (compared to allochthonous sources) and a critical component of carbon sequestration in these systems.” Moore, T.L. and Hunt, W.F., 2012. Ecosystem service provision by stormwater wetlands and ponds–A means for evaluation?. Water research, 46(20), pp.6811-6823.

Response: We agree that vegetation likely plays an important role in sequestering carbon in sediments and have added the following paragraph to the discussion in section 4.4:

“Studies have also shown the importance of emergent vegetation plant species in sequestering carbon in sediments. Emergent vegetation was found to contribute significantly to the soil carbon pool of stormwater ponds compared to allochthonous sources (Moore and Hunt, 2012). However, in our CH_4 model, the significant effect of sediment C:N ratios suggested that an autochthonous organic matter source from either phytoplankton or submerged macrophytes supports greater CH_4 production in farm reservoirs. The ability of farm reservoirs to have a negative climate forcing will rely on the balance between GHG fluxes and sediment carbon accumulation. The effect different plant species and other aquatic primary producers have on both these processes needs to be evaluated in future studies as the current design of farm dams within the study area minimises growth of emergent vegetation through steep sides and slopes.” Line 426

L392. “The flux of N2O was constrained in our earlier study (Webb et al., 2019), which found a small CO2-e sink (-89 to -3 mg CO2m-2d-1) for the majority of these farm reservoirs despite high N concentrations.” Something of a diversion here, but doesn’t this depend on how the data are interpreted though? In your earlier study the median N2O flux was negative, but the mean was positive (with 33% of ponds emitting N2O), whilst in this study (figs 4 and 5) you present mean CH4 and CO2. There’s probably a debate to be had concerning what average is most appropriate to use, but note the IPCC Refinement used a mean value calculated from log-transformed values.

Response: We thank the reviewer for their insight but have respectfully retained our original presentation. As noted above, we presented both median and mean in the Webb et al. 2019 publication because we wanted to make a clear point that most small agricultural reservoir was, unexpectedly, not a major source of N_2O. This result is not highly dependent on the form of the summary statistic (weak sink, weak source; neither are large). Similarly, in this paper, we focus on the mechanisms predicting variation in the C-based GHG fluxes rather than the absolute values. Thus, while we agree that the ‘optics’ of the presentation (interpretation by readers) of median and mean are slightly different, we feel that this is a ‘side issue’ better left for the IPCC committees to debate.

End of Referee #1 response file