

Dear Drs. Scranton and Neubauer,

We very much appreciate the time and effort you put into reviewing this manuscript. We have synthesized your comments and questions into a single response in the form of an author's comment on the discussion site, but we kept different reviews separate. We have also attached the revised manuscript as a supplement, where the changes inspired by Dr. Scranton's comments are marked with blue text and those inspired by Dr. Neubauer's are highlighted in yellow. Again, we appreciate the feedback and recognize that your efforts have resulted in a greatly improved manuscript.

Respectfully,

Carmella Vizza, Will West, Stuart Jones, Julia Hart, and Gary Lamberti

RESPONSE TO REVIEWER 1, DR. MARY SCRANTON:

Response to 2nd set of comments (SC2):

Dear Dr. Scranton,

Thank you again for taking the time to review our response to your comments. We appreciate the friendly tone of the conversation, and we acknowledge that many of the misunderstandings have resulted from differences between marine and freshwater fields. Having your perspective as a marine geochemist has helped us to hone our language and hopefully the manuscript will now be more easily understood across fields. Although we know that you disagree with our specific sea-level rise simulation, we appreciate your comment that the study "does contribute to an understanding of how increasing sulfate concentrations might change methane production." In the hope of keeping the second response succinct, we will briefly address some of the major points in your second and first set of comments in that sequence. We hope that our actions will speak louder than our words in the supplement that we have attached, which is a revised version of the manuscript with blue text for the changes that have been made in light of your comments and suggestions.

First of all, we agree that there may be differences in terminology between scientists studying marine and freshwater ecosystems, so we have tried to make our terminology clearer throughout the manuscript by describing our wetlands as fully inundated (e.g., **please see pg. 4, lines 6-12**), and we also separated water column parameters from sediment parameters (**please see Tables 1 and 2**) and added data on porewater acetate, nitrate, and sulfate concentrations. Regarding the pooling of samples, if we had wanted to assess the "electron tower" paradigm in freshwater wetlands, we agree that it would be not be appropriate to make composite sediments for microbial samples or even make homogenous 20-cm sediment slurries for methane production without doing more explicit surveys of how porewater chemistry and microbial communities vary by depth of the sediment. However, we believe that assessing the validity of this paradigm is beyond the scope of this paper. While we call on it to inform our results, the main purpose of the first portion of the paper was to assess how average methane production varies across freshwater and brackish wetlands and how microbial communities and normative chemical conditions may influence this process. Our goal was thus to characterize methane production over a large spatial extent, which inevitably involves some homogenizing and compositing of sediments for study feasibility.

Additionally, we explicitly present how much carbon we added in the organic matter simulation (**please see pg. 5, lines 14-17**) and appreciate you detecting this oversight. We also formalized our analyses of litter quality, discussed litter quality more in depth, and tried to be explicit about what Tiegs et al. 2013 previously measured in the Methods, Results, and Discussion (**please see pg. 5, lines 17-18, pg. 8, lines 28-29, p. 10, lines 20-21, and pg. 15, lines 2-9**). We also added more details in the Methods about our chemistry and statistical analyses (**see Sections 2.3.2 and 2.3.3 and pg. 8, lines 1-8**). Regarding your concern with the sea-level rise simulation, we now refer to it throughout the manuscript as a “biogeochemical sea-level rise simulation” (e.g., **please see pg. 1, line 14, pg. 3, lines 19-23, pg. 13, lines 20-25**). Lastly, we agree with you that hydrogenotrophic methanogenesis is potentially important and that we cannot rule it out in our ecosystems (**please see pg. 12, lines 14-18, and pg. 12, last paragraph**).

Again, we appreciate the time and effort you have taken to make this manuscript a better one. Below you will find your set of comments point-by-point to which we have already responded and posted online, but we have added in the exact changes that we have made in the **revised manuscript in bold** for your reference.

Respectfully,

Carmella Vizza, Will West, Stuart Jones, Julia Hart, and Gary Lamberti

Response to 1st set of comments (SC1):

Dear Dr. Scranton,

Thank you for taking the time to read and comment on our manuscript. We appreciate your observations and questions. Below is our response to your comments.

Comment: “The basic premise on which the authors base their study is that methane production depends on substrates produced during fermentation of organic matter and that methane production and sulfate reduction usually do not occur in the same sediment, presumably because of competition for these substrates.” **Response:** In this comment, you seem to imply that methane production and sulfate reduction do not occur in the same sediment. Methanogens and sulfate-reducing bacteria do directly compete for organic substrates such that when sulfate or other more energetically profitable alternative electron acceptors are available, some methane is still produced, but it is at lower rates (please see manuscript citations: Achtnich et al. 1995; Lovley and Klug 1983, 1986).

Comment: “The authors mimic sea level rise by adding brackish water to their samples (which ignores the possible importance of increased water logging and decreased oxygen in the sediments).” **Response:** These sediments came from freshwater wetland ponds (0.4 – 1.2 m total water depth) that were already waterlogged and therefore had decreased oxygen in the sediments. We have no reason to suspect that the physical act of sea-level rise would change the saturation status of the sediments. **Please see pg. 4, lines 2-8.**

Comment: “All incubations are done after vials are flushed with nitrogen although the natural sediments apparently all had some level of oxygen present in situ.” **Response:** In this comment and throughout your review, you seem to imply that purging vials with nitrogen gas is problematic. It is widespread practice in methane production bottle assays to purge with nitrogen gas in order to remove the oxygen

that was introduced by removing sediment from below the water's surface and exposing it to the air in the field and in the lab (Lofton et al. 2014; Sinke et al. 1992; West et al. 2012, 2015). Of course, it is possible that there are microsites where oxygen is present in the freshwater wetland sediments, but we believe that the same may be true in the bottle experiments. **Please see pg. 5, lines 23-27.**

Comment: "Values for nitrate (another potential substrate for carbon remineralizing organisms) and for acetate (the putative important substrate for methanogens) were measured but data were apparently quite variable and are not reported." **Response:** Summary porewater acetate levels are reported in Results section 3.1.1. You are correct in that we did not report actual values for nitrate, and we will add a short line in the results to indicate that nitrate was usually 2+ orders of magnitude lower than total sulfate levels. We measured acetate, nitrate, and sulfate fairly extensively in both the water column and the porewater as well as other parameters. We did not present all the data we collected because we did not want to overwhelm the reader, but these data are all available in the archived data. If reviewers and editors feel that these data would enhance the manuscript, we will add them in tabular or figure form to the manuscript or as supplemental information. **Please see Table 2.**

Comment: "No measurements were made of H₂, another potential substrate for methanogens and methanogens were not examined to see whether they were actually acetoclastic or hydrogenoclastic." **Response:** We did not measure H₂ production as we do not currently have an instrument available to measure this nor did we classify the type of methanogens. However, acetate was the predominant factor explaining methane production patterns in almost all experiments, which suggests that a large proportion of the methanogenesis is acetoclastic (e.g., pg. 8 lines 29-30). However, we do not deny the potential of hydrogenotrophic methanogenesis occurring in our ecosystems (please see last paragraph on page 10). **Please see revised manuscript, pg. 12, lines 14-18, and pg. 12, last paragraph.**

Comment: "The data for the *dsrA* and *mcrA* genes are again not presented and in the results section appear to contract the conclusions made about their abundance. The authors also ignore the fact that numbers of genes do not directly relate either to number of cells (a cell can have more than one copy of a gene) or to gene activity." **Response:** The *dsrA* and *mcrA* summary data are presented in the Results section 3.1.3, and individual gene data per sample are also archived in the data set (see microbial tab). Although we acknowledge that the number of genes does not equate with number of cells or gene activity, qPCR of functional genes for particular guilds is a commonly used approach to estimate the abundance of a functional group. We also acknowledge in the manuscript that it is possible that some of the genes we detect are from dormant microbial communities (please see pg. 11, lines 8-12). The number of genes corresponding to abundance is always going to be a bit of an assumption, but in a broader survey of our wetlands, we have unpublished data from our system that *mcrA* abundance relative to *dsrA* abundance is correlated with methane production rates. Also, in *Letters in Applied Microbiology* (Volume 62, Issue 2, Pages 111-118), Morris et al. 2016 found that hydrogenotrophic methane production rates corresponded to *mcrA* abundance. Although observation of functional gene transcripts or even the actual enzymes catalyzing the reactions would clearly be preferred, logistics associated with our remote field sites precluded work with these more labile macromolecules. Despite challenges associated with interpreting DNA copy number of functional genes, we believe that our use of qPCR substantially adds to the literature as there are few studies that simultaneously measure methane production and microbial functional group abundances. For example, the studies listed on pg. 9, lines 22-26, have hypothesized that microbial community processes are behind methane patterns

along a salinity gradient, but none of these studies actually tested this hypothesis by measuring microbial communities. **Please see revised manuscript, pg. 7, lines 7-9 and pg. 11, lines 12-25.**

Comment: “Finally the authors use natural organic matter to enrich their incubations but do not indicate the amount of carbon added or the relative lability of that carbon. (Clearly a gram of sucrose and a gram of twigs would not be expected to stimulate microbial activity to the same extent).” **Response:** We added 3 g of live tissue per macrophyte species, which translated to 0.10-0.12 mol of C added to each incubation, and we will add these values to the methods of the manuscript. Dr. Scott Tiegs of Oakland University has measured the % C, % N, and % P of the four litter species we added: Maretail (45% C, 1.7% N, 0.17% P), Buckbean (44% C, 0.94% N, 0.15% P), Lily (45% C, 1.7% N, 0.17% P), and Horsetail (47% C, 2.5% N, 0.24% P), but the patterns we observed did not follow %C, %N, %P, C:N or C:P, which are considered standard indices of litter quality. We discuss the phosphorus in detail because Tiegs et al. 2013 found that litter decomposed more rapidly when P was higher (see pg. 11, lines 14-24). You are correct in that we did not specifically measure the lability of the carbon from each macrophyte species, but we are leveraging information from Tiegs et al. 2013, whose study had already assessed this by measuring decomposition rates of these litter types in Copper River Delta ponds. We will clarify this in the discussion. **Please see revised manuscript pg. 5, lines 14-18, pg. 8, lines 28-29, pg. 10, lines 21-22, and pg. 15, lines 2-9.**

Comment: “P2 Line 6: I do not like the term green house compensation point. It is not widely used and does not directly relate to what you are measuring (as you say nothing about carbon sequestration in this system).” **Response:** We believe that the term “greenhouse compensation point” accurately describes wetlands’ dual roles in carbon sequestration and emissions, and we believe that because northern wetlands are on the edge of this compensation point that it is so important to study how global change may alter the methane cycle. Greenhouse compensation point is a good way to set up the broader context of this study. **Please see pg. 2, lines 6-8.**

Comment: “You do not mention carbon fertilization again, and there is no indication in the text of whether any increases in carbon production in the CRD system would be due to warming or to carbon fertilization. I think you hurt yourself by trying to draw connections to too many issues. I would recommend drastic simplification of this section and sticking to the facts.” **Response:** Increased CO₂ levels lead to warming as well as CO₂ fertilization, both of which could affect the amount of substrate available. We are not precisely sure whether increased organic matter will result from warming, from longer growing seasons directly, or from CO₂ fertilization; therefore, it seems prudent to mention both potential mechanisms. Regardless, we will reevaluate our discussion of these points based on your recommendation and consider ways to focus our comments. **Please see pg. 2, lines 9-12 and pg. 3, lines 4-5.**

Comment: “P2 Line 16: redox conditions are only indirectly related to climate change. You seem to imply the mere presence of nitrate or sulfate changes redox conditions but this is not true, especially if oxygen is present. In fact there can be a lot of nitrate and sulfate in oxic surface sediments. Did you ever have sulfide in your samples? Does oxygen penetration vary with ecosystem? How far below the surface were the samples collected?” **Response:** True, redox conditions are a potential indirect effect of climate change. Sea-level rise could increase sulfate levels just as higher decomposition rates due to warming could deplete oxygen levels, both of which affect redox conditions. The sediment came from the top 20 cm, and although some oxygen is likely present in small amounts, it is also likely depleted quickly. For

example, the layer of water directly above the sediments often has low DO levels of around 1 mg/L, particularly in the evening, and it is not uncommon to have anoxic groundwater upwelling in these systems. We also observe that water column DO levels drop throughout the season as vegetation begins to senesce. As for hydrogen sulfide, we did not directly measure that, but sediment characteristics (black coloration and pungent odor) suggest the presence of sulfide. Usually sediments become anaerobic within the first few cm of freshwater ponds, which is another reason that we make the incubations anaerobic by purging with nitrogen gas. Of course oxygen matters, but the amount of nitrate in particular is orders of magnitude lower than agricultural and other human impacted systems. Lastly, if significant levels of oxygen were present in these sediments, it would kill the methanogens because they are extremely sensitive to O₂ (please see manuscript citation: Whalen 2005) and we would therefore see very little methane production in our experiments. **Please see pg. 2, line 18.**

Comment: “P2 Line 23: I don’t think an “abundant supply of organic matter can reduce competition for methanogens by increasing substrate availability”. Try instead “abundant supply of organic matter can increase substrate availability”” **Response:** We appreciate the wording suggestion and will make the change. **Please see pg. 2, lines 28-29.**

Comment: “P2 Line 28: replace “are likely results of” with “may be influenced by” “ **Response:** We appreciate the wording suggestion and will make the change. **Please see pg. 3, line 2.**

Comment: “P3 lines 7 and 10. Use same unit for sealevel rise (100 and 170 cm)” **Response:** We appreciate the wording suggestion and will make the change. **Please see pg. 3, lines 14-16.**

Comment: “P3 line 15: Numbers of methanogens not as important as whether or not the methanogens are active.” **Response:** Please see our response to your major comment above about qPCR and the feasibility of RNA work. **Please see revised manuscript, pg. 7, lines 7-9.**

Comment; “P4 line4: The range of physicochemical parameters in table 1 are actually pretty small for most measurements. Perhaps more important is whether the intertidal sediments are exposed to the atmosphere at low tide (tidal range?). How long are they submerged? What is the water content? Again it matters how far below the surface these sediments were collected. From the table it must be shallow since there was more O₂ in these sediments than the freshwater wetlands.” **Response:** The intertidal sediments were collected from the top 20 cm just as in the freshwater wetlands, and they are covered with freshwater during low tide and increasingly brackish water during high tide. So again these sediments are waterlogged with depleted oxygen levels. In Table 1, the DO data are actually from the water column, as we do not have oxygen data for the sediment. In the freshwater ponds where limnological profiles could be conducted, we reported DO levels from the bottom of the water column, but in intertidal marsh, DO levels came from the surface layer. This is detailed in the legend for Table 1. We will remove the water column DO data altogether since they seem more misleading than helpful. **Please see Table 1.**

Comment: “P4 line 26: Were no replicates run for sediments from a single site? How can you tell if observed variability is just typical of replicate samples? I would also think you MUST indicate how much macrophyte tissue you added (probably in terms of gC/g sediment or something like that) to even know if these treatments were similar since the lability of the carbon is likely not the same.” **Response:** We ran five control replicates without added substrate at each freshwater wetland (n = 5), which we then averaged to form the basis of the delta CH₄ production metric. Each macrophyte treatment was

replicated in 5 different ponds. In the increased organic matter simulation, we used pond as the replicate because we were more interested in capturing how wetlands differing in biogeochemistry along a glacial to oceanic gradient would respond to organic matter addition rather than how much variability there exists within a single wetland's response. **Please see pg. 5, lines 7-12.**

Comment: "P5 Line 6: You mean incubation temperature not ambient temperature?" **Response:** Yes, we do. We appreciate this good suggestion for a wording change. **Please see pg. 5, line 24.**

Comment: "P5 line 8: Purging with nitrogen will likely have a bigger effect than incubating at a few degrees cooler than the actual sediment. I would expect stimulation by this as you allow more anaerobic processes to occur. Were any blanks or controls run?" **Response:** Please see our response to your major comment above. **Please see pg. 5, lines 23-27.**

Comment: "P5 line 12: "flame ionization" not "flame ionizing" detector" **Response:** We will make this change. **Please see pg. 5, lines 31.**

Comment: "P5 line 20: You report that only sulfate was detectable in the water column. Did you present any water column data? Do they mean anything? What are detection limits for pore waters? Present nitrate and acetate data for sediments as well as sulfate." **Response:** We do have extensive water column data for these wetlands, some of which are in the archived data. If reviewers feel these data add to the manuscript, we can add them in tabular form or in supplemental information. Some (i.e., DOC and sulfate) are also already presented in Table 1, but acetate and nitrate were not detectable in the water column (we will add concentration detection limits to the manuscript which were about 10 and 5 μM , respectively). **Please see Table 1 for water column data and Table 2 for sediment porewater data.**

Comment: "P5 line 27: the method you describe is typically called "loss on ignition" and represents a loss of organic matter. Did you convert to organic C? Is this data reported anywhere in the paper? Was DOC measured the same way? Blanks? Detection limit?" **Response:** Yes, we took the sediment organic matter data (in % of OM) from loss on ignition and converted it to organic C (pg. 5, lines 25-29). We will add in the following citation about conversion to organic C: Thomas et al. 2005; *Aquat. Sci.* 67:424–433. The data were not reported in the paper as to not overwhelm the reader, especially since the amount of carbon did not affect the results, but the data are archived and available online. DOC was measured using a Shimadzu TOC-V (pg. 5, lines 18-19), and we did run blanks. All samples had detectible DOC, but our lowest standard was 1 mg/L and all samples registered above this level with the exception of five of the intertidal samples, which fell in between the blank and the lowest standard. If reviewers would like more detail on these analyses in the methods, we will add them to the manuscript. **Please see sections 2.3.2 and 2.3.3.**

Comment: "P6 line 2: I would use the word "converted" rather than "scaled" " **Response:** Thank you for the wording suggestion. **Please see pg. 6, lines 18 and 23.**

Comment: "P6 line 5: I couldn't tell what this composite sediment was used for. Were incubations for each site or was all sediment made into a composite? Why do this? It seems that then the averages for the genes refer to different samples than the physiochemical properties. Can you properly do statistics comparing the combined samples in one parameter to the individual samples in one parameter?"

Response: The composite sediment was a combination of the five sediment samples from different sites

for each freshwater wetland, but this was only done for the microbial analyses. Making a composite is highly practiced in soil microbial ecology for the purpose of controlling analytical costs while still capturing significant spatial heterogeneity. We made a composite for each freshwater wetland which were then compared against ten sediment samples from intertidal marsh. So yes, it is true that we could not directly link methane production of a single sample or its physicochemical properties to the functional group abundances, but overall we were able to characterize methanogen and sulfate-reducer abundances across ecosystems and relate these to average methane production rates. **Please see pg. 6, lines 30-31 and pg. 7, line 1.**

Comment: “P7 line 2: move “log transformed” to immediately follow “four factors”” **Response**” Methane production rates were log-transformed, not the factors. We cannot make this change because it would misrepresent our statistics. **Please see pg. 7, line 27-29 where we rearranged the wording for clarity.**

Comment: “P7 line 26ff: You use a lot of significant figures for something that is so variable. Maybe you are justified in two significant figures but not 3 or more. Again blanks and detection limits need to be mentioned as your averages are pretty close to zero (or at least almost include zero). I would like to see these data associated with specific samples” **Response:** We will reduce the number of significant figures and make the needed detection statements in the methods. We only reported parameters if they were detectable. The reason that these averages are close to zero in this particular line is that we converted actual concentrations to the total amount in the incubation bottle (i.e., $\mu\text{M} * \text{PW volume in incubation} * (\text{PW volume per ml of sediment}) * \text{sediment volume in bottle}$). **Please see sections 2.3.2, 2.3.3, 3.1.1., 3.2 and Tables 1 and 2.**

Comment: “P8 line 14: This paragraph seemed very odd. You seem to contradict yourself a lot. For intertidal samples, three had no *dsrA* but this was found in all freshwater AND the *dsrA* was independent of ecosystem type. This seems to contradict your hypothesis which you nevertheless cling to. Remember the presence of a gene doesn’t mean the organisms are doing anything at the moment and the number of genes does not necessarily indicate the number of cells of a particular organisms.” **Response:** Indeed, we did not detect *dsrA* in 3 of the 10 intertidal samples. Nevertheless, *dsrA* abundance tended to be higher in intertidal ecosystems, although not significantly so. In contrast, the *mcrA* presence and abundance did vary significantly by ecosystem, and methanogens are the group that we tend to use to explain our methane production results directly, while we use sulfate and sulfate-reducer abundance as supporting evidence (please see pg. 9, lines 14-16). **Please see revised manuscript pg. 9, line 22, and pg. 11, lines 18-25.**

Comment: “P8 line 27ff: I really don’t like your equating sealevel rise purely with sulfate concentration. It is ok to say that sea level rise will flood current intertidal areas, but the vegetation will change and the water will be permanently water logged rather than periodically exposed to air. You have no data on estuarine wetlands to compare to the freshwater wetlands (although others have done this comparison in other systems). I would guess the reason you see no effect here is that you are mimicking the process the wrong way. You are really looking at increased sulfate levels, not sealevel rise.” **Response:** In these systems, which are waterlogged throughout the year, we expect sulfate concentration to be a major biogeochemical change. Our simulation not only adds sulfate, but it also adds other nutrients and the microorganisms brought in with the saltwater. You are correct that over the long term vegetation and other characteristics of these ecosystems will also change. We will add comments at pg. 3., line 12, and pg. 4 line 20 to reflect our biogeochemical focus on the effects of sea-level rise. Our data on intertidal

ecosystems for comparison with freshwater ecosystems can be found in Results section 3.1.2 and Figure 3. **Please see revised manuscript, pg. 1, line 14, pg. 4, lines 6-8, pg. 13, lines 20-25.**

Comment: “P10,line2: The sentence beginning “our study also demonstrates...”seems to directly contradict your own data as you said before that *dsrA* did not correlate with ecosystem type” **Response:** Just because *dsrA* presence and abundance did not significantly correlate with ecosystem type does not mean the presence or abundance of sulfate-reducing bacteria does not affect methane production in intertidal ecosystems. Even though these bacteria are also present in freshwater ecosystems, they could be limited by sulfate availability. Also *dsrA* tended to be higher on average relative to freshwater ecosystems, just not significantly so. We will change the wording to “tended to have generally higher sulfate-reducer abundances when present” to address this concern. **Please see revised manuscript pg. 9, line 22, and pg. 11, lines 18-25.**

Comment: “P10 line 13ff and next paragraph: These two sections are wild speculation with absolutely no data behind them. There are a lot of other factors that might be important that you haven’t included.” **Response:** We provide potential hypotheses and explanations for our data that were grounded in the literature, which is not an uncommon practice in the discussion. Specifically, we discuss the acetoclastic pathway of methane production and how we might expect that to change in the future as well as possible reasons for why methane production rates varied by an order of magnitude throughout the season. We could omit these sections, but felt that it was important to say something about the results we observed. Regardless, we will reevaluate the nature and extent of our explanations. **Please see pg. 12, line 13, where we have extensively re-worked this paragraph. Also, we acknowledge that the last paragraph on pg. 12 involves hypothesizing about seasonal differences in CH₄ production because the data we collected fall short of explaining them, but the discussion is backed up by literature and we put forward a more formal hypothesis stating that future study is needed.**

Comment: “P11line14: This whole section is hurt by the fact that you don’t characterize the macrophytes at all in terms of potential lability. Did you add a constant amount of C or organic N? Or just dry mass? Or just a “chunk” of leaves? Again much of this section is speculation without more facts” **Response:** Please see our response above to your major comment about lability of macrophytes. We evaluated many different possibilities for explaining the results of the different macrophyte treatments including % C, % N and % P, but the quality measurement that seemed to matter the most was antimicrobial properties. None of these potential indicators of lability appeared to have an effect on methane production, which we will clarify in the Discussion pg 11, lines 14-30, and pg 12, lines 1-9. **Please see revised manuscript pg. 5, lines 14-18, pg. 8, lines 28-29, pg. 10, lines 21-22, and pg. 15, lines 2-9.**

Comment: “P12 line19: You have no idea whether there are hydrogen utilizing methanogens around or how much hydrogen there might be. You can reiterate what other people have said in other papers but I don’t see you connecting these to your system with any facts.” **Response:** We will try to improve the clarity of the manuscript, by acknowledging that methanogens in general utilize acetate, directly (acetoclastic) or indirectly (hydrogenotrophic). We will also add that CO₂ and H₂ are compounds that also come from fermentation of organic matter such as acetate. Hence, this is why acetate is important and was measured in this study. **Please see pg. 12, lines 14-18.**

Comment: “Modeling: I didn’t really understand the models. If you are making linear equations including various parameters and then constant factors for each, I would like to see more detail on how this worked. If you have a lot of variables, you can fit most data but figure 6 is completely mysterious and doesn’t convince me of the value of your model. I also didn’t understand the columns in tables 2 and 3. I have never heard of an AIC before, for example. Explain the statistics a bit to an audience that may include non-biologists.” **Response:** Akaike Information Criterion (AIC) is an increasingly utilized form of model selection that generates estimates of the model being the best representation given the data and the set of models being explored. It also penalizes for the number of parameters in the model. We will add more description of this in the methods. Basically, general linearized models are created, each of which is associated with an AIC value, which is then used to rank the models. We also corrected for small sample sizes (AIC_c). The model with the lowest AIC_c value is considered the best, and all remaining models are compared relative to the best approximating model using delta AIC_c (Δ_i). Models with a Δ_i less than or equal to 2 are considered to have substantial support, while models having a Δ_i greater than 7 have little support (Burnham and Anderson 2002). The relative strength of the models was then evaluated with Akaike weights (ω_i), which indicate the probability of a model being the best model, given the data and the set of candidate models (Burnham and Anderson 2002). **Please see revised manuscript pg. 8, lines 1-13.**

Respectfully,

Carmella Vizza, Will West, Stuart Jones, Julia Hart, and Gary Lamberti

RESPONSE TO REVIEWER 2, DR. SCOTT NEUBAUER

Response to 1st set of reviewer comments (R1):

Dr. Neubauer,

We greatly appreciate the extensive time and effort you put into reviewing our manuscript. Your comments were insightful and extremely thorough, and we believe that your suggestions have greatly improved the manuscript. Below please find our point-by-point responses to your suggestions and comments with the location of changes we made to the revised manuscript in bold. The attached supplement is the revised manuscript, which also has the changes your comments inspired highlighted in yellow.

Comment: “I was very interested in the topics covered by this manuscript and think that the measurements and experiments can provide some insight into questions about controls of CH₄ production and how global changes will affect CH₄ dynamics. However, I was a bit disappointed in the analysis and interpretation of the data, especially in how the authors tied their findings to existing knowledge. Similarly, I felt that the authors could have done a better job exploring the literature on how plants and organic matter inputs affect CH₄ production.” **Response:** We are pleased that you found our study interesting, and we thank you for pointing us towards several studies we had not found in our original literature search. We hope you find our discussion greatly expanded in light of the literature you recommended as well as a few other papers we came across in the process. **Please see pg., 11 lines 25-30, and pgs. 12-14.** Also, we hope that the expanded information on the AIC model selection as well as the fleshing out of some of our hypotheses in the discussion will aid the reader in understanding our data analysis and interpretation.

Specific comment 1: p. 1, line 13; p. 3, line 12; and throughout manuscript, “freshwater and intertidal wetlands”. As someone who studies tidal freshwater wetlands, I take issue with the way you are classifying wetlands as either freshwater *or* intertidal. Wetlands can be both! Indeed, there are over 19,000 ha of tidal *and* freshwater wetlands within the Copper River Delta (see Hall’s chapter in the 2009 “Tidal Freshwater Wetlands” book; Barendregt, Whigham, Baldwin (eds). Backhuys Publishers). A better way of characterizing your two groups of sites would be “brackish intertidal” and “non-tidal freshwater” (unless, of course, your freshwater sites were also intertidal). **Response:** Certainly, wetlands are difficult to classify and the terminology is something that we should clarify because it appears to be a source of confusion. In light of Dr. Scranton’s comment, we have added that the ‘freshwater wetlands’ are constantly inundated with fresh water so as to distinguish them from our ‘intertidal wetlands’, which are covered in freshwater at low tide and covered in increasingly brackish water at higher tides (rarely exposed to air at low tide). Although our freshwater wetlands currently do not receive any tidal influence, their surrounding sloughs sometimes do, which is why they are at risk of seawater intrusion. We have concluded that the best terms to call our wetlands in light of yours and Dr. Scranton’s comments are “tidal brackish wetlands” and “non-tidal freshwater wetlands.” **Please see pg. 1, line 13 in the Abstract and the terms were changed throughout the rest of the manuscript (we only highlighted the first instance in the revised manuscript so as not to distract from other changes) as well as pg. 4, lines 9-12.**

Specific comment 2: p. 1, lines 15-16: Your data clearly show that rates of CH₄ production and porewater sulfate were each higher in the brackish sites. But, how did you determine that the high sulfate was the cause of the lower CH₄ production? You also reported differences in porewater nitrate and acetate (top of p. 8) between ecosystem types, and presumably there were differences in salinity as well. How did you conclude that sulfate was the driving factor? Sulfate wasn’t even the most important factor from your modeling exercise (p. 8, lines 11-12). **Response:** Table 3 shows that the top model (lowest AIC_c score) included all factors (ecosystem type, acetate, sulfate, and time period) and the second best model included ecosystem type, acetate and sulfate. Therefore, for explaining methane production rates in the non-tidal freshwater wetlands and tidal brackish wetlands, the most important factors are ecosystem type, acetate, and sulfate. First, nitrate availability was negligible in comparison to the total sulfate availability so we decided to remove that from the analysis (also based on your comment #16). Second, we concluded that acetate was not responsible for lower methane production rates in the brackish sites because acetate was actually higher than in the freshwater sites, and higher acetate availability should theoretically increase methane production. We believe that the variable ‘ecosystem type’ captures factors other than sulfate such as microbial communities and perhaps even salinity. It is difficult to disentangle the effects of sulfate from salinity since those two variables are highly correlated (**please see our expanded discussion of this on pg. 11, lines 18-30 and pg. 12, lines 1-12**). However, mechanistically from a redox perspective, we believe that it is most plausible that lower methane production rates in brackish sites resulted from sulfate-reducing bacteria (which tended to be higher when present in brackish versus freshwater sites) that outcompete methanogens (which were significantly lower in brackish versus freshwater sites). One could also argue that stress from salinity might also lower methanogenic activity, but if this were the case, this should be a more immediate effect that would have decreased methane production in the sea-level rise simulation. Conversely, the turnover of microbial communities in response to redox conditions may be a less immediate effect. You are correct in pointing out that our hypothesis, although supported by our data, does require some

assumptions and deduction on our part. We have changed the wording to “probably due to higher sulfate availability” on **pg. 1, lines 15-16**.

Specific comment 3: p. 1, line 19, “. . .increased organic matter generally enhanced CH₄ production rates.” This statement is too strong for your data set. You used four different organic matter amendments (= 4 plant species). The CH₄ production rates increased for only two of the species you tested (Fig. 5); the other two species had no effect. So, based on your data, I don’t see how you can justify saying that the amendments “generally enhanced CH₄ production.” If something happening 50% of the time means that it is a “general” occurrence, you could also say that the amendments generally had no effect on CH₄ production. **Response:** In calculating the difference in methane production between the paired treatment and control, this metric was positive in 15 out of 20 cases and so we believe that 75% of the cases is enough to say “generally.” However, we agree with your comment in light of the particular species and will change the wording to say “enhanced CH₄ production in 75% of the incubations, but this response depended on the macrophyte species added with half of the species treatments having no significant effect,” **please see pg. 1, lines 20-21**.

Specific comment 4: p. 2, line 4, “. . .21 times more effectively. . .” This value of the global warming potential for CH₄ over a 100-year time period is quite old and has been updated in each of the two IPCC reports that have been published after the Whalen paper you cited. The bigger point of this comment is that the global warming potential may not be the best way to compare CO₂ and CH₄ when one is talking about ecosystem processes, where C₃ gases are emitted or sequestered year after year. I discussed this in a paper that was published last year (Neubauer and Megonigal. 2015. Moving beyond global warming potentials to quantify the climatic role of ecosystems. *Ecosystems*. 18:1000-1013). **Response:** We appreciate your comment and acknowledge that using GWP may not be the best or most updated way of addressing why methane is an important greenhouse gas. We changed the wording to better reflect this, **please see pg. 2, lines 4-5**.

Specific comment 5: p. 2, lines 5-7, “Currently, wetlands at northern latitudes. . .” Other studies that have taken a longer temporal perspective have concluded that many northern wetlands have had a net cooling effect for the last 8,000-11,000 years (Frolking and Roulet 2007. *Global Change Biology*. 13:1079-1088). Because CH₄ is broken down in the atmosphere – and therefore the warming due to CH₄ emitted in any given year is transient – but the cooling due to C sequestration lasts “forever,” a wetland that is old enough can have a lifetime net cooling effect, even if its radiative balance over a shorter period implies net warming. So, a single wetland could have a warming or cooling effect, depending on what time scale you consider. **Response:** Thank you for your insightful comment. We agree with you that a wetland’s role in carbon sequestration and emissions is highly dependent on temporal perspective. **Please see pg. 2, lines 6-12** where we attempt to better reflect this point.

Specific comment 6: p. 3, line 28: Why is the word “wetlands” in quotes? Are you suggesting that your sites aren’t actually wetlands? Also, why are you comparing brackish intertidal marshes with (unvegetated?) freshwater ponds? My understanding is that you are trying to compare sites that differ in salinity and sulfate due to their effects on CH₄ production. Why not compare vegetated brackish marsh with vegetated freshwater marsh? Or, brackish ponds with freshwater ponds? **Response:** We were attempting to name a term that we would use to refer to them throughout the manuscript and therefore should have put both CRD and wetlands in quotes. For clarification, we are comparing vegetated freshwater non-tidal wetlands or “ponds” or with vegetated brackish tidal wetlands. We

consider the freshwater wetlands to be “pond-like” because they have more clearly delineated boundaries, whereas the brackish tidal wetlands are continuous. Unfortunately, there are no brackish water ponds that we know of in the Copper River Delta to which to compare our freshwater systems. We also wanted our sites to be comparable in depth. The freshwater sites with depths greater than 0.4 m all tend to be “pond-like” due to the clay-like sediments in the area that allow very little drainage, whereas the brackish sites tend to have a depth of up to 2 m during high tide and 1 m during low tide. In general, the sites we sampled in the brackish tidal wetlands tended to be a bit shallower than the average depth for logistical reasons (i.e., using a handheld bucket auger from a jet boat). We have attempted to make this clearer in this paragraph, **please see pg. 4, lines 5-12.**

Specific comment 7: p. 4, lines 10-11 and 17: I’m a bit confused by your sampling design. You collected a single sample from five sites along the salinity gradient. Elsewhere (e.g., p. 3, lines 16-17), you explained that you expect that the availability of sulfate will be an important driver of rates of CH₄ production. Given that, why would you combine all the sites along the salinity gradient into a single “brackish intertidal” value? **Response:** We combined all the tidal brackish sites into one value for the tables in the manuscript because this ecosystem was rather continuous and lacked clearly distinct boundaries like the non-tidal freshwater wetlands, **please see pg. 4, lines 15-18.** They are also combined in the figures as to contrast two distinctive ecosystem types. However, when it comes to data analysis through AIC model selection, each tidal brackish site is considered on its own such that widely varying sulfate levels are appropriately taken into account, **please see pg. 7, lines 27-29.**

Specific comment 8: You have quite a wide range of sulfate values (Table 1); was there a significant relationship between porewater sulfate and CH₄ production? This would be another way of getting at your hypothesis about the effects of sulfate on methanogenesis. **Response:** Yes, if we log-transform methane production rates from the brackish and freshwater wetland comparison and run correlation tests on both porewater sulfate and total sulfate (porewater + water column), the results are $r = -0.67$, $P = 0.0001$, and $r = -0.71$, $P = 0.00003$, respectively. Although we believe that doing separate correlations by factor are redundant with AIC model selection (although different philosophically), we are happy to include these correlations in the manuscript if reviewers and editors feel they would improve its clarity.

Specific comment 9: p. 4, line 20: Does this sulfate concentration indicate the sulfate concentration in the water that was added when making the slurry or does it indicate the sulfate concentration in the slurry itself? Can you also report the salinity of the water added for the slurries (or the final salinity of the slurry, whatever is consistent with the sulfate concentration)? **Response:** We have added another table to distinguish between water and sediment chemistry data, with the water column concentrations presented in Table 1 and the porewater concentrations presented in Table 2. The salinity concentrations of the water column are reported in Table 1, but unfortunately our salinity probe is only designed for measurements of the water column in the field so we were not able to also measure the salinity of the porewater (i.e., we could only extract a few mL of porewater per sediment sample). **Please see Table 2**, where we have presented both the porewater sulfate data as well as the total amount of sulfate available in the slurry (PW + WC).

Specific comment 10: p. 4, “Increased organic matter simulation” section: How much organic matter did you add to each bottle? Did you characterize the organic matter (e.g., C, N, P contents? lignin content)? Did you use aboveground or belowground tissues? Were the tissues first cut to a standard size (e.g., passed through a grinder) before added to the bottles? Details such as those should be added to this

section. **Response:** This is a good suggestion. **Please see pg. 5, lines 14-18** where we have updated this information in the manuscript.

Specific comment 11: p. 4, line 30: The genus is *Menyanthes*, not *Menanythes*. The same genus name is misspelled in some of the figure legends. **Response:** We greatly appreciate you catching this error. It has been changed throughout the manuscript. **Please see pg. 5, line 13, and Figure Captions 5 and 6.**

Specific comment 12: p. 5, lines 5 and 9, and elsewhere in the manuscript. Generally, you include a space between a number and its units (e.g., “60 mL” on line 5) but other times you don’t (e.g., “250mL” on line 5). I also remember seeing some places where you said that your experiment lasted for “14d” (instead of “14 d”). When editing, check throughout the manuscript to see that you include a space between a number and its units. **Response:** We agree with you on consistency and will check all instances. In general, we did not include a space when a unit was used an adjective, as in describing a 250-mL serum bottle, which according to the style of *Biogeosciences* should be listed as a “250mL serum bottle.” Nouns such as “250 mL of sediment” should indeed contain a space as you suggest. In light of the confusion, we have gone back and hyphenated all the adjectives and highlighted these instances for reference. We hope that the editor will clarify whether he prefers the lack of space or the hyphenated adjectives in light of these instances.

Specific comment 13: p. 5, line 12: It is a flame ionization detector, not a flame ionizing detector. **Response:** Thank you. We have made this change, **please see pg. 5, line 31.**

Specific comment 14: In the context of comparing between treatments within your study, it is fine to report your CH₄ production rates as μmol per bottle per time. However, this really limits your ability to compare your results with those of others. Note, for example, that you would have gotten different rates (on a per bottle basis) if you had used a different volume of sediment, even if everything else was identical. At a minimum, you should tell the reader the weight of sediment in each bottle (e.g., “The 60 ml of added sediment was equivalent to 80-90 g of dry sediment.”). It would be even better if you reported your rates as μmol CH₄ per gram sediment per time. **Response: Please see pg. 5, line 22,** where we have added in the average amount of wet sediment. We agree with you that we would have gotten different methane production rates if we had added widely variable amounts of sediments, but the coefficient of variation in the wet sediment masses was approximately 3% and therefore did not affect the major trends when we reran all the analyses with CH₄ production rates as nmol CH₄ per gram of dry sediment per day. Originally, we reported values on a per bottle basis because we were treating each bottle as its own “wetland microcosm.” This is also the reason we ran analyses with the total amount of each anion per incubation converted from porewater concentrations. Nonetheless, we agree with your argument that it is important to be able to make this comparable to other studies. Therefore, we have added the line in the methods clarifying how much wet sediment was added (dry sediment mass is available in the archived data) and have completely re-done all figures, tables, and analyses with CH₄ production rates being reported and analyzed as nmol CH₄ per gram of dry sediment per day. **Please see pg. 6, line 3, Tables 3-4 and Figures 3-6.**

Specific comment 15: p. 6, lines 1-2, “Porewater concentrations were scaled. . .” I don’t understand what this means. Are you saying that you multiplied the anion concentrations by the volume of porewater in order to determine the total amount of each anion in the bottle? Based on your nutrient/anion results (bottom of p. 7 to top of p. 8), I think this is what you did. But, why did you do this? As with the CH₄ production rates (previous comment), reporting things on a per-bottle basis makes

your numbers completely dependent on the amount of sediment (and its water content) that you ran through the centrifuge. It seems more straightforward to report your nutrients using molar units (e.g., mM) because those numbers are independent of the volume of sediment that was processed and can be easily compared with other studies. **Response:** We have clarified how and why we converted porewater concentrations to the total amount per g of dry sediment (we removed bottle rates based on your comment 14), **please see pg. 6, lines 21-25**. Essentially, we converted the porewater concentrations to the total amount of each anion per bottle because the volume of porewater extracted from the sediments varied widely among samples (~47% CV). In light of your suggestion of being able to make comparisons between studies, we have now reported both the amount of each anion in the porewater per g of dry sediment as well as the traditional porewater concentrations **in Table 2**.

Specific comment 16: p. 7, line 4, “total sulfate and nitrate” Is this fourth factor the sum of sulfate and nitrate? If so, why did you add them together? I recognize that both sulfate and nitrate are electron acceptors. Adding them together makes the implicit assumption that one mole of nitrate is “the same” as one mole of sulfate. However, the thermodynamics (i.e., energy yield) and stoichiometry (i.e., moles of sulfate or nitrate reduced per mol of carbon oxidized) differ for sulfate reduction and nitrate reduction. So, in terms of competing with methanogens for electron donors, one mole of nitrate is not equivalent to one mole of sulfate. **Response:** Thank you for this insightful comment. Originally, we combined sulfate and nitrate as an overall representation of the presence of alternative electron acceptors and in order to minimize the number of parameters in model selection. Although we agree with you that the thermodynamics and stoichiometry varies per reaction, we thought that trying to account for this would be negligible in light of the fact that nitrate was such a small component of this factor (~ 4.5%) in comparison to total sulfate (~ 95.5 %). In light of how negligible nitrate is and because of the stoichiometric concerns you voice here, we have decided to remove nitrate from analyses and instead just focus on total sulfate availability and therefore redid all analyses to account for this. **Please see pg. 7, line 31 and pg. 8, line 1**.

Specific comment 17: p. 7, line 7: What is delta i? **Response:** **Please see pg. 8, lines 1-13** where we have updated the manuscript to reflect more details about AIC model selection.

Specific comment 18: GLMs: I am approaching this manuscript as someone who is interested in the questions you are addressing but is getting lost when trying to understand and interpret the GLMs. Admittedly, this is because I have not ever received formal (or informal) training in GLMs. I am not expecting the revised manuscript or your “Response to reviewers” document to provide a tutorial in how GLMs work, but I hope that you will be able to make some modifications to the manuscript text that will make it easier for someone who isn’t familiar with GLMs (such as myself) to follow the analyses that you did. **Response:** **Please see pg. 8, lines 1-13**, where we have updated the manuscript to reflect more details about AIC model selection. **Specific comment 18a:** In a GLM, what is a parameter estimate (e.g., as mentioned on p. 7, line 7)? I am familiar with multiple linear regression analyses, where each explanatory variable has its own parameter (or, “slope”). But, multiple linear regressions use continuous explanatory variables. In contrast, you have some nominal variables (e.g., ecosystem type, time period, macrophyte species). I can’t begin to translate from my experience with multiple linear regressions to guess how you would come up with a parameter estimate for a nominal variable, or what such a parameter would even mean. **Response:** For categorical variables, a GLM essentially adjusts the intercept. For example, in considering ecosystem type, both freshwater and brackish wetlands would have their own intercept. **Specific comment 18b:** How do you estimate the importance of different

variables? And, what does “relative importance” mean? I first thought that relative importance would be where everything is expressed as a fraction of the importance of the most important variable (so, the relative importance of the most important variable would be 1 and everything else would have a lower relative importance). But, that must not be the case since none of the variables for the GLM from the organic matter addition experiment have a relative importance of 1 (p. 9, lines 10-11). **Response:** Great question, and we have included more details about this in the manuscript. **Please see pg. 8, lines 11-13.** Also, we have included the updated Akaike weights after model averaging in **Tables 3 and 4** so that the reader can mentally calculate the relative importance in their head. For example, in Table 3, there were three models included in model averaging with a $\Delta_i > 4$, and sulfate is a factor included in 2 of the 3 models whose $\omega_{i(MA)}$ were 0.61 and 0.26, so its relative importance is 0.87. **Specific comment 18c:** Given Figure 6, it is apparent that you can use GLMs to generate predictive equations. Would it be worthwhile to include your best predictive equations in the manuscript for the reader to see? **Response:** We will be glad to report this should the editor feel this would be helpful. Although since these models are fairly specific to this study and these ecosystems, we would not necessarily recommend that others try to extrapolate methane production rates from them. **Specific comment 18d:** I have no idea how to interpret the “Akaike weights” numbers in Tables 2 and 3. **Response:** **Please see pg. 8, lines 3-8.** **Specific comment 18e:** What is a null model? **Response:** A null model includes an intercept only and we have updated this in the methods for clarification, **please see pg. 8, line 9, and Tables 3 and 4.**

Specific comment 19: p. 7, Statistical Analyses: What are you using as your level of statistical significance – 0.01? 0.05? 0.10? **Response:** Great question, and we have updated our methods to reflect this. **Please pg. 8, lines 16-17.**

Specific comment 20: p. 7-8, Water column and porewater chemistry results: A student of mine did an experiment where he measured porewater concentrations of 130 μM sulfate, 5 μM nitrate, and 4 μM acetate. How do those numbers compare with yours? I don’t actually want you to make that comparison but I want to make the point (again) that it is impossible for the reader to make that kind of comparison. I only know that you processed “~50 ml” of sediment but I have no idea of the water content of that sediment. Therefore, I cannot convert your values of the stock of sulfate or nitrate or acetate in a ~50 ml chunk of soil to a concentration. **Response:** Again, you are absolutely right that it is important for others to be able to compare their ecosystems to ours in terms of chemistry, and therefore porewater concentrations have been added to **Table 2.**

Specific comment 21: p. 8, line 9 and Tables 2 & 3, “total sulfate/nitrate” Earlier, I thought that you calculated sulfate + nitrate. Did you actually use the ratio of these anions in your analyses? What is the rationale for doing that? **Response:** No, we did not use a ratio, but rather we originally added sulfate and nitrate together. However, we have removed nitrate from analyses based on your comment 16.

Specific comment 22: p. 9, lines 18-20 and p. 13, lines 4-5: I am unclear how you concluded that it is important to consider the interactions of multiple global change mechanisms. I totally agree with that idea, but do not see how your results demonstrate the importance of studying interactions. After all, you only looked at individual factors, not at interactions. If you had done an experiment where you manipulated salinity and organic matter availability individually and together and had found that the interacting factors gave results that were unexpected based on single factor experiments, then you would have support for the idea that it is important to consider the interactions of multiple global change mechanisms. **Response:** Good point, as our study did not directly assess the interactions of

simulated sea-level rise and increased organic matter availability. However, in assessing the differences in methane production between freshwater and brackish wetlands as well as in our sea-level rise simulation, we included both redox conditions (i.e., total sulfate availability) and a component of organic matter (i.e., acetate availability) as factors in our models. Both redox conditions and acetate availability were important factors for both modeling efforts, but interestingly the influence of each factor for the freshwater/brackish comparison was in the opposite direction for the organic matter simulation. This suggests that environmental context matters, i.e., higher sulfate levels generally lead to lower methane production, but methanogens in areas with more alternative electron acceptors might be more likely to respond to increased organic matter availability. This is why we believe that other studies should not only look into these factors in different ecosystems, but also examine how they interact. **Please see pg. 11, 3-6.**

Specific comment 23: p. 9, line 22-26, “Many studies. . .” Weston et al. (2014; *Biogeochemistry*. 120:163- 189) is an example of a study that measured CH₄ fluxes along a salinity gradient *and* measured rates of methanogenesis. Neubauer et al. (2013; *Biogeosciences*. 10:8171- 8183) reported CH₄ fluxes and CH₄ production for a wetland that experienced >3 years of experimental saltwater intrusion. **Response:** We thank you for drawing our attention to these sources. **Please see pg. 11, lines 12-16.**

Specific comment 24: p. 9, line 27, “directly linked lower CH₄ production to higher sulfate and nitrate concentrations” This is odd phrasing since you saw higher CH₄ production where you had higher nitrate; your sentence suggests the opposite pattern. I think the confusion here is related to the way you did the GLMs with sulfate+nitrate (or perhaps sulfate/nitrate) as a model factor. You chemically analyzed these anions separately, but then combined them in some way for the statistical analyses. As noted in earlier comments, I do not understand how/why you combined these anions for statistical analysis. **Response:** Our results did show lower methane production when sulfate availability was high (**please see Table 3**), and we removed nitrate based on your comment 15. Methane production and sulfate availability were inversely related (please see our response to your comment 8). We attempted to clarify the wording, **please see pg. 11, lines 14-16.**

Specific comment 25: In your Statistical Analysis section (p. 7), you say that you used “general linear models.” In the legends for Tables 2 and 3, you say that you used “general linearized models.” A quick Google search for the exact phrase “general linearized models” did not reveal any statistics-related results in the first two pages of search results (I didn’t look any deeper than that). A search for the same term (without the quotes) suggested Wikipedia pages for “general linear model” and “generalized linear model” as the top two search results. These are not the same thing. So, this comment is a long-winded request that you clarify whether you used “general linear models,” “generalized linear models,” or “general linearized models” (whatever they are) and to make sure that you use the correct terminology throughout your manuscript. **Response:** We appreciate you catching this, it is most common to just refer to them as GLMs, but the analyses used were actually named “generalized linear models” and **pg. 7, line 28**, has been changed accordingly as well as in **Tables 3 and 4.**

Specific comment 26: p. 10, lines 11-15: You cannot directly compare your CH₄ production rates with those reported by Hines et al. (2008). Most importantly, Hines used 50 ml of slurry with 1 part soil to 3 parts total slurry volume; in each of your bottles, you used 120 ml of slurry that was 1 part wet sediment to 1 part water. All else being equal, we would expect higher CH₄ production in your study simply

because you had more sediment in your bottles. Expressing the rates per gram of soil/sediment, as I suggested earlier, would go a long way toward making your results comparable with those from other studies. [Helpfully, Hines et al. reported the typical weight of dry soil per milliliter of their slurry so you can get a rough idea of what their rates would be if expressed per gram of soil. Your data repository file lists sediment weights (wet or dry?) for each bottle.] **Response:** Originally, we had wanted to present Hines' data in bottle rates so that they could be comparable to the rates we originally report in the manuscript, but we agree with your point that rates should be compared based on g of sediment. **Please see pg. 12, lines 18-25**, where we have updated these calculations and extensively re-worked this paragraph.

Specific comment 27: p. 10, last paragraph: Ultimately, this paragraph is unsatisfying. After seeing the huge June vs. August difference in rates of CH₄ production, I was really hoping that you would be able to provide some strong insight into the cause(s) of that difference. I guess you are limited by data availability. Still, I wonder if others working in similar systems have reported order-of-magnitude changes in rates over such a short time period. I don't know anything about your system except what is in the manuscript but I'm wondering if the pattern could be related to the timing of soil thaw in the early growing season or perhaps the phenology of plant growth. Finally, I'll note that the measured concentration of acetate reflects the balance between rates of acetate production and acetate consumption. So, if higher acetate production in August was balanced by higher acetoclastic methanogenesis, you would see high rates of CH₄ production without correspondingly high acetate concentrations. **Response:** Yes, you are correct that we are limited by data availability, but wish we had been able to explain these intriguing seasonal differences. However, in the hope of making this paragraph more satisfying, we have fleshed out our hypothesis about why these seasonal differences might occur and it indeed pertains to plant phenology! Also, you are correct that porewater acetate availability is a balance between acetate production and consumption. **Please see pg. 13, lines 8-18.**

Specific comment 28: p. 5, line 20 and p. 11, line 8: Change detectible to detectable. **Response:** Done, **please see pg. 6, line 10 and pg. 14, line 17.**

Specific comment 29: p. 12, lines 5-8: The possible role of antimicrobial compounds is an interesting hypothesis and you presented some information to support it. However, I did not see where you tested for the effects of litter quality (e.g., C:N:P, percent lignin, lipid content) on CH₄ production rates. Without having run those analyses (either the chemical analyses or the statistical analyses), why are you discounting the possible influence of those factors that have previously been shown to be important? **Response:** We have formalized our analyses of the Tiegs et al. (2013) stoichiometry for the macrophytes used in this study and can therefore rule out C:N:P. **Please see revised manuscript pg. 5, lines 14-18, pg. 8, lines 28-29, pg. 10, lines 21-22, and pg. 15, lines 2-9.** Although we cannot rule out percent lignin or lipid content, it is our hypothesis that antimicrobial compounds were responsible for why methanogens responded differently to macrophyte species based on the literature we found about their antimicrobial properties. Even if we had measured percent lignin or lipid content, it would be difficult to conclusively rule out that some other quality measure contributed to the differences we observed without extensively characterized the quality of these macrophyte species. Again, we acknowledge that antimicrobial compounds hypothesis is untested, but this study does provide some anecdotal evidence suggesting that it might be worthy of further examination. **Please see pg. 15, line 24.**

Specific comment 30: p. 12, line 13, “Fewer studies have examined. . .” There have been studies looking at CH₄ emissions when wetland plants are grown in an elevated CO₂ environment. Although there are important differences between CH₄ emissions and CH₄ production, the elevated CO₂ studies generally find that CH₄ emissions increase with elevated CO₂, with this increase often being attributed to higher plant production (see, for example, Vann and Megonigal 2003. *Biogeochemistry*. 63:117-134).

Response: We appreciate you pointing us toward some interesting studies. **Please find our discussion updated on pg. 15, lines 29-31, and pg. 16, lines 1-6.**

Specific comment 31: Table 1: The legend says that you made 4-10 spot measurements per site. Given that, why aren't there any standard deviations or other estimates of error/variability for pond depth, temperature, pH, and salinity? **Response:** In light of the fact that these variables did not vary all that much, we originally did not report the error in order to avoid the table from becoming too cluttered. However, we are happy to report them, and please find standard deviations for all these parameters in **Table 1.**

Specific comments 32 and 33: Data repository: I took a look at the data that you made available at the knd.informatics.org site and had a question about your CH₄ production rates on the “All CH₄ data” Excel worksheet. In column R, you reported “areal CH₄ production (umol/m²/d).” How did you determine areal rates? What does an areal rate even mean in the context of a soil slurry in a bottle? Is the “area” the same as the cross-sectional area of the bottle? If so, that seems meaningless to me since a cylindrical bottle is going to have the same cross-sectional area whether the bottle is 1/4 full, 1/2 full, or 3/4 full but, C₁₀ presumably, would have different rates of CH₄ production due to the different amounts of soil in the bottle. Given that you didn't report the areal rates in your manuscript, this whole thing would be a question of curiosity. . . except that you used these areal rates to calculate the per-bottle rates (column S of the Excel file) that *are* reported in your manuscript. So, I need to know more about these areal rates before I can judge the validity of the per-bottle rates. I do not understand the formula that you used to go from areal rates (umol/m²/d) to per-bottle rates (umol/d): per-bottle rate = areal rate * 0.2 * sediment volume in liters. In order for the units to work out, the 0.2 factor must have units of m²/L. Those are odd units. Where does 0.2 m²/L come from and what does that conversion factor represent? **Response:** We apologize for any confusion caused by not removing the areal rates from the data repository, and we have updated those data included in the repository accordingly, **please see <https://knb.ecoinformatics.org/#view/doi:10.5063/F1028PF8>**. We have plans to use areal CH₄ production rates as a parameter for another manuscript, one that involves a process-based model with a methane budget for each pond. In order to use the methane production data as an input we needed to come up with a conversion factor of sediment volume to area, which is what the 0.2 m² per L represents. That conversion factor has nothing to do with this study though; it is just that we simply back-calculated the bottle rates in this particular spreadsheet from the areal rates. Just to be clear about our process: To get bottle rates, the GC concentrations are in ppm that are regressed against incubation duration in days (since we have measurements from multiple time points), which we then convert to μmol L⁻¹ d⁻¹ which we then multiply by headspace volume. To determine areal rates, we take the GC concentrations in ppm that are regressed against incubation duration in days, which we then convert to μmol L⁻¹ d⁻¹ which we then multiply by headspace volume and divide by sediment volume, and lastly we multiply by the areal conversion factor (1 L/0.2 m²). This is why in the spreadsheet online we took the areal rates, divided by that areal conversion factor (or multiplied the values by 0.2 m²/L), and then multiplied that value by sediment volume in mL and divided that by 1000 to convert to L.

Again, this was only a back calculation since we did not have the raw $\mu\text{mol L}^{-1} \text{d}^{-1}$ data in this particular spreadsheet.

Respectfully,

Carmella Vizza, Will West, Stuart Jones, Julia Hart, and Gary Lamberti