Biogeosciences Discuss.,
https://doi.org/10.5194/bg-2017-281-AC3, 2017

Biogeosciences Discussions

Interactive comment on "Impact of land use and soil properties on soil methane flux response to biochar addition" *by* Weiwei Cong et al.

Weiwei Cong et al.

vivien8026@syau.edu.cn

Received and published: 15 November 2017

Dear Reviewer, Many thanks for the constructive comments and suggestions on our manuscript (bg-2017-281-SC1). We have carefully addressed all the issues raised by you in the comments. Please find our point-by-point responses under each of reviewer' s comments.

1.page 4, line 14. Be specific about the metric(s) used by Song et al, who used a response ratio. Typically such a measure is ln-transformed, but I gather they did not apply a transformation because they had fluxes <0 (a similar problem that limited your use of lnRR).

Response: Thank you for your suggestion. We will rewrite this sentence.

2.Page 5, line 13. Please be more specific in the main text – saying "not appropriate" is too vague. I think all you are saying is that you required 3 or more sampling dates for the study to be included in your dataset.

Response: Thank you for your suggestion. Yes, we are saying "Studies where gas sampling frequency were taken 3 times or more during the entire experiment were included." We will rewrite this sentence.

3.Page 5, line 7. You had the potential for a single paper to contribute many effect sizes (i.e., an average of 5 given that you had 268 "experimental treatments" and 50 articles; although the use of the term "experimental treatments" is misleading because an experiment has two treatments (biochar addition and a control)). You do not discuss how you dealt with the non-independence arising from multiple estimates from the same paper. Nor do you even discuss the nature of these multiple estimates (sometimes the same study is repeated over multiple time periods; or in multiple locations; or data from an experiment with crossed factors leading to multiple biochar-control comparisons). At a minimum, you need to explain the structure of the data to the reader. Ideally you would incorporate this hierarchical structure of the data into your analyses. I return to this point below, where it appears that multiple effects from the same paper might drive some of your conclusions (suggesting that you've "over-represented" a single study).

Response: You are correct. We check all the original data, if the data is from repeated studies, we will only use the estimate one time. Because we didn't want to lose the information about the crossed factors leading to multiple biochar-control comparasions, we will use the multilevel meta-analytic model to reanalyze the data in order to deal with nonindependence among data points originating from the same studies.

4.The Page 7, line 16. Specify that experimental duration was measured in days.

Response: Thank you for your suggestion. We will rewrite this sentence.

5.Page 8. I agree that many published papers have used Hedges' d. I do not agree

that it is a suitable metric of effect. This point has been most cogently made two decades ago by Osenberg et al. (1999; Ecology). One concern is that Hedges' d is standardized by the standard deviation (variation among replicates). As a result, two groups of studies could differ in d, not because of shifts in means (or modes) but because one group of studies yielded more precise values. Let's say, for example, that you had a group of lab studies and group of field studies. Each showed a shift in the CH4 flux from 5 (ambient) to 2 (biochar addition). However, the lab studies were less variable (s=0.1) than the field studies (s=1.0). In this case, the d for the field studies would be 30 and the d for the field studies would be 3. This 10-fold difference does not reflect a difference in expected CH4 flux, but rather a difference in how variable replicates are when placed into an incubator in the lab vs. plots in the field. One of the other problems that can arise with Hedges' d is that studies with small n, can give rise to unusually small s (just due to sampling error). If that's the case, then your estimate of d can be greatly inflated (just due to sampling error). Indeed, you report d's that are often >10, which is a huge standardized effect. I suspect you get many of these results because you are using st. dev.s based upon small n's (i.e., s's that are estimated very poorly). However, you don't report the n's so it's impossible to judge the magnitude of this problem. With respect to understanding how biochar will affect CH4 flux to the atmosphere, I think you are much more interested in the flux itself and not its experimental variability. Given my concern about Hedges' d, I encourage you to seek a more appropriate measure of effect – one that is well matched to the question you are asking. I think one viable option is a simple difference: CH4 flux in the biochar – CH4 flux in the control. This response can be directly scaled to the issue of question (effect on CH4 emissions or update), and can be applied to different land-use scenarios to estimate large-scale effects of biochar application that go beyond specific experimental results.

Response: Your are correct. We checked the original data, Hedges' d was influenced significantly by the lab precision. Therefore, we will use raw difference : CH4 flux in the biochar – CH4 flux in the control as the effect size and reanalyze the whole data. We

will separate the data into two groups in analysis: incubation studies, field studies , to deal with the variation among replicates in different scales.

6.Weighting of effects. Traditional meta-analyses weight effect sizes by estimates of their precision. It's entirely unclear if you did this. I don't think you did, but you should have. Note that this is a different issue than dividing by s to obtain Hedges' d and it's a different issue than LOESS. Please note that Song et al. did not weight (in their analysis of response ratios) because they did not want to discard studies that failed to report variances. However, because you used Hedges' d, you could only use studies that reported variances and sample sizes. Thus, you have all the information needed to weight by precision (the inverse of the variance of the effect size).

Response: Thank you for your suggestion. We will use multilevel meta-analytic model instead of LOESS to reanalyze the data in order to deal with nonindependence among data points originating from the same studies. we will use raw difference : CH4 flux in the biochar – CH4 flux in the control as the effect size. Study will be used as random factors (estimating variance components).

7.You need to give more specific about the LOESS methods. E.g., what weighting function did you use? I assume the tri-cube? What method did you use to construct 95% envelopes?

Response: Thank you for your recommendation. We will use multilevel meta-analytic model instead of LOESS to reanalyze the data in order to deal with nonindependence among data points originating from the same studies.

8.I found the reference to Hedges' d in section 3.1 a bit confusing because you refer the reader to Figure 1, which does not present d, but instead gives a plot of the two means. Furthermore, you don't present any CI's on these estimates in the text. Perhaps that's because you give these data in Figure 2, but I'm uncertain if that's the case.

Response: Thank you for your suggestion. The Figure 1 showed CH4 flux from un-

treated control soils and the biochar treatment soils. The Figure 2 showed the Hedge's d and the estimates CI of Hedge's d. We will rewrite this section to make the description for Figure1 and 2 clearly.

9.Page 9, line 18. What does an "increase", "decrease" or "no change" mean? Do these categories refer to results that were "significant" in the original study or does it mean something else? Perhaps it's just a statement about the direction of the difference, without regard to significance, but then I'm surprised that 6 estimates would be exactly equal to zero.

Response: Thank you for your suggestion. We want to illustrate the difference in CH4 sink/source strength of each treatment, without regard to significance. Yes, there are 6 biochar treatments showed no difference in CH4 sink/source strength. We will rewrite this section to make the statement clearly.

10.Page 10, lines 19-21. Similarly, I'm confused by this statement, which suggests that some studies give evidence of demonstrable increases or decreases. However, you give no indication of which of the effect sizes are different from 0? Just saying that some estimates are >0 and some are <0 does not mean that positive and negative effects cancel out – the variation might just represent sampling error. [Of course, you are probably inferring these effects given the subsequent analyses, but the reader is not privy to these results.

Response: Thank you for your suggestion. Yes, just saying that some estimates are >0 and some are <0 can't conclude that positive and negative effects cancel out. We will rewrite this section.

11.In part, some of my concerns reflect a need to conduct a more formal evaluation of heterogeneity, as is done in classic meta-analyses with weights (with Q statistics). Given that you have estimates of s and n, I would expect you to do this (no matter what your metric of effect is).

Response: Thank you for your recommendation. Heterogeneity test for the biochar effect among different subgrouping categories in this study would be added in the manuscript.

12.Figure 2: I found it odd that you repeat the top set of comparisons (mean, upland, paddy soil) under "soil", "biochar" and "management": i.e., the top three panels are identical (except that the scale varies among the three columns). I suggest you move this result to its own figure since it does not involve the use of soil or biochar or management covariates.

Response: Thank you for your suggestion. We will move the repeated comparisons and redraw the Figure 2.

13.I found the discussion of the LOESS results unconvincing. That may be because I'm not an expert with LOESS methods, but the patterns that you present are quite perplexing. For example, there are multiple areas where the 95% CI shading does not include the 1:1 region, but there are either no data in this region (e.g., see Fig 3h, x-axis range from 13-22; Fig 3j from x=40-70) or the data that do exist seem to contradict the pattern. Clearly these patterns are greatly influenced by points "far" from the region, suggesting that the weighting function you use to fit the LOESS is having an influence that I can't clearly intuit – on one hand there are lots of local influences (the lines bounce around a lot), but on the other hand, there are "significant" effects even when the local information is very sparse.

Response: Yes, there are only a few data are in some areas. To make the trend are more convincible, we will use multilevel meta-analytic model instead of LOESS to re-analyze the data in order to deal with nonindependence among data points originating from the same studies.

14.In addition, I found it hard to see the data in Figure 3. Points overlap and often the CI's are also hard to discern.

Response: Thank you for your suggestion. We will redraw the figure 3.

15.Figure 3 also suggests that some studies are having a disproportionate effect. For example, in Fig 3, it appears that there is a single study (with soil C=18, soil N = 1.6, soil C:N=11, soil pH=5.9, etc.) that contributes many points, most of which have very negative d's (i.e., 0 to -40). This one study is solely responsible for the "dip" in the LOESS fits seen around these values. There are at least two problems suggested by this: 1) I'm guessing from these exceptionally large values of d, that this study probably had few replicates and a very small estimate of variability (i.e., s(pooled) in the denominator of d was exceptionally small). With few replicates, s is estimated poorly and by chance you can get very large d's. 2) If you had weighted by 1/variance, I'm guessing that the influence of these points would have been reduced. 3) But even if you did weight by 1/variance, the influence of this one study would still be disproportionate because you have not addressed issues of non-independence.

Response: Your are correct. We checked the original data, Hedges' d was influenced significantly by the lab precision. Therefore, we will use raw difference : CH4 flux in the biochar – CH4 flux in the control as the effect size and reanalyze the whole data. We will separate the data into two groups in analysis: incubation studies, field studies , to deal with the variation among replicates in different scales.

16.Given the above concerns, I remain unconvinced of the role of any of the covariates shown in Figure 3. I further suspect that many of the "effects" seen in Figure 2, exist only because of the disproportionate weight given to a few papers that each yielded many estimates of effect.

Response: Your are correct. We checked the original data, Hedges' d was influenced significantly by the lab precision. Therefore, we will use raw difference : CH4 flux in the biochar – CH4 flux in the control as the effect size and reanalyze the whole data. We will separate the data into two groups in analysis: incubation studies, field studies , to deal with the variation among replicates in different scales. Covariates will be used

as fixed factors in the multilevel meta-analytic model instead of LOESS to reanalyze its influence in methane emission difference.

17.Please also note that it's very difficult to discern patterns in Figure 3 (with scales on d that go from -40 to +20) when the average effects you document in Figure 2, appear (at best) to be on the order of 2.

Response: We checked the original data, Hedges' d was influenced significantly by the lab precision. Therefore, we will use raw difference : CH4 flux in the biochar – CH4 flux in the control as the effect size and reanalyze the whole data. We will separate the data into two groups in analysis: incubation studies, field studies , to deal with the variation among replicates in different scales.

18.I appreciated having access to the supplement; however, it also raised very concerns. For example, if you look at the first figure under the column "Hedges' d" (by the way, Larry Hedges is the name, so I think the possessive should be after the "s": Hedges' d, not Hedge's d), the distribution of d seems to shift as you move from the top panels to the bottom panel. All panels appear to share the same scale, but the bottom panel has a mode around 0, while the panels above have a model closer to 10. That can't be correct – they should have the same distribution. A similar problem exists in the next figure (compare the distribution d for the clay covariate with the other variables), and other figures....

Response: Thank you for your correction. We checked the manuscript and corrected the expression of Hedges' d. The panel in the diagonal is the variables distribution. For example, the Hedges' d distribution showed in the column 1 and row 1 in supplement figure S1, the soil organic C content distribution showed in the column2 and row 2. Therefore, the distribution of the diagonal panel is not the same.

19.I'm not sure of the journal's policy, but I feel very strongly that the data should be provided as an appendix. That appendix should give the means, st. devs, and n's for the +biochar and ambient treatments, as well as the value of the covariates. Please do

not just give the effect size, as that precludes other investigators from reanalyzing the data using another metric.

Response: Thank you for your suggestion. We will provide the data as appendix including the means, st. devs, and n's for the +biochar and ambient treatments and the category for the soil, biochar and management variables.

20.Despite your reference to "linear additive models", I saw no place in which you fit models using multiple covariates simultaneously. Instead, all of the results, look at covariates one at a time (except, in part, for the decomposition of paddy vs. upland sites).

Response: Thank you for your comment and suggestion. We will use Multilevel meta-analytic model instead of LOESS to reanalyze the data. We will have covariates in the figure 3 as fixed factors and the study as random factors to reanalyze the soil , biochar and management properties influence on methane emission difference.

Interactive comment on Biogeosciences Discuss., https://doi.org/10.5194/bg-2017-281, 2017.