Dear Editor

Thank you for your helpful work on the manuscript. In response to the comments from the reviewers, we added a short paragraph to address the applicability and the implications for the indices in general meaning which was suggested by the Referee #3. We also added a few short words to the caption of Fig.2 to explain the meaning of the five symbols in Fig.2(b). Some sentences in the abstract were revised as suggested by the Referee#1. We also added the Author Contribution sector to list the contribution of each co-authors. All modifications made to the manuscript were highlighted by a colored background.

The Referre#1 raised several questions in doubt to the rationale of the method in the study. We made point-to-point responses to those questions as possible as we can at present and the itemized responses were listed to the end of this letter.

In addition, according to the co-authors' contribution, we changed the order of the co-authors. This was also highlighted in the manuscript.

Best Regards

T. T. Li
LAPC, Institute of Atmospheric Physics, CAS
No.40 Huayanli, Chaoyang District, Beijing, China 100029
TEL: +86-10-62012067
E-mail: litingting@mail.iap.ac.cn

Referee#1:

1. While both reviewers pointed out the same concern that reasons other than the data sharing (e.g., the similar geographical condition in temperature etc.) would lead to higher correlation across two grid cells and this could seriously affect the results and conclusions, as far as I read, it is rather hard to say that the authors did a good job to satisfy reviewers' doubts. I am still in doubt about whether the aggregation method proposed here truly deals with the data scarcity as the authors are discussing.

R: In the MS and the responses to the comments of the reviewers during open discussion, we explained the rationale of the methods and the difference between the correlations due to similar geographical conditions and that due to data sharing. We'd like to illustrate it here by a simple example. Suppose we have two cases, Case-I and Case-II. In Case-I, we have two "separate and identical" datasets for two cells (cell *i* and cell *j*). Here "identical" means each of the model input variables has the same values in the two datasets. In Case-II, we have only one dataset for some (not all) of the model input variables for the two cells and all the values are exactly the same as that in Case-I. In both cases, the model yields identical outputs in the two cells and they are "closely correlated" owing to "similar geographical conditions". When aggregating the model outputs in the two cells, we give C_{ij}=0 for Case-I because the two cells have "separate" datasets but in Case-II, the C_{ij} has a value greater than 0 because the two cells share data for some of the model input variables.

2. Ironically, this study suffers from the data scarcity while trying to quantify the uncertainties in national-level methane emission associated with the data scarcity. This in part relates the concern mentioned above. A concern came to me after reading the revised manuscript is that the experimental design needs improvement to be more straightforward. For instance, while the authors conducted the Monte Carlo simulation at a county level, this is less informative when aggregating because the information on the spatial heterogeneities of some input variables within a county is lacking. And this lead to an unnecessary doubt regarding the dependency of the confidence interval of national-level estimate to the scale of simulation. As the central purpose of this study is to propose an aggregation method, it would be more straightforward to start analysis from a spatial resolution where the information on the spatial heterogeneities of all input variables within a calculation unit is available.

R: As mentioned by the reviewer, the central purpose of this study is to propose an aggregation method of the model estimation when suffering from data scarcity. At a given scale, county level for example, the spatial heterogeneities of some input variables within a county can be analyzed when we have sufficient data, like the case for soil properties. But we don't have enough data for the spatial heterogeneities in other factors and the Monte Carlo simulation in a county had to be carried out by sharing the data and PDFs of those factors. We are aware that broader regions usually have larger spatial heterogeneity than smaller regions and sharing PDFs made for broader regions (a province for example) might overestimate the spatial heterogeneity in a smaller region (a county). But exploring details of the spatial heterogeneity needs more spatially available data and we return back to the problem of data scarcity. We thank the suggestion of the reviewer that "it would be more straightforward to start analysis from a spatial resolution where the information on the spatial heterogeneities of all input variables within a calculation unit is available". In the present study, however, we can't follow the suggested way of analysis because we don't have enough data to do so. Fig. 2 illustrates the spatial availability of the data in the present study.

3. "Improving the data availability of the model input variables is expected to reduce the uncertainties significantly… (L29-30)". The statement is one of main conclusion of this study (as far as I read), but conflict with the results. This comment is applicable to L264-266. I agree this in general terms, but what you present in Table 4 is that even a full set of spatially fine inputs are unavailable, the simulations at a finer level would help to reduce the uncertainties in national-level estimate. And there is no reasoning why the simulation at a finer scale showed the smaller uncertainty despite of the lacking of spatially-dense information on some input variables and relatively comparable model sensitivity across all input variables at a site level (see Table B1).

R: The simulations at a finer level would help to reduce the uncertainties in national-level estimate because some, though not ALL, of the model input variables have data at a finer resolution as showed in Fig. 2. We are not to expect a further smaller uncertainty if we made the estimation at a scale finer than the finest resolution of the available data showed in Fig. 2.

4. In my view, many statements in the manuscript are worded too strongly or without caution. The followings are the examples (but not all):

4.1 "Data scarcity is a substantial cause of the uncertainties in estimating the methane emissions on regional scales (L14-15)". It may be true, but you cannot say with absolute certainty as the relative importance of data scarcity against other sources of uncertainty is practically impossible

to quantify due to the lack of a full set of spatially-dense inputs.

R: We revised the sentence into "Data scarcity is one of the substantial causes of the uncertainties in estimating the methane emissions on regional scales".

4.2 "from site scale up to regional/national scale (L17)". I don't think you present a result at a site scale. The finest scale presented is a county scale. Table B1 shows to some degree the sensitivity of the model at a site scale, but this has little relevance to the data scarcity.

R: The phrase "site scale" is not proper here and we revised the sentence into "from county/provincial scale up to national scale". Table B1 is the analysis of the model sensitivity to support the application of the equation (4) and the discussion of how the model sensitivity influences the spatial aggregation of the model outputs (Line 338-346).

4.3 The statement such as L25-28 is likely unfaithful if no caution related to the limitations of the method proposed here is accompanied by.

R: We addressed the limitations of the methods in the MS between Line 287-309. In the abstract (Line 25-28), only the main results of the study were presented.

Referee#3

The authors addressed an important issue for upscaling the process-based model form site scale to regional one. They proposed two indices to quantify the degree to which data scarcity (availability) affects the uncertainties in upscaled estimations at various scale as well as the redundancy (cost) for implementing the heterogeneity at finer scale, that is, I_ds and I_R. These indices are useful for the applications. But, the applicability and the implications for these indices are not clear. More informative explanation about them should be added in the section of Results and Discussion. It will help readers to apply these indices to their specific study. I recommend this article should be accepted after minor revisions.

R: We appreciate the suggestion by the reviewer and add a short paragraph at the end of the Result and Discussion (Line 347-365) to address the applicability and the implications for the indices in general meaning.

Specific comment:

Figure legend for Fig.2 is not clear. Specifically, more explanations about Fig.2(b) should be added indicating what each symbol (SP, RC, OM, WR and VI) shows.

R: We add short explanations to each symbol of Fig.2 (b) in the figure caption. More information of the symbols was provided in Appendix A.