Geosci. Model Dev. Discuss., 7, C3–C6, 2014
www.geosci-model-dev-discuss.net/7/C3/2014/

Geoscientific Model Development Discussions
Open Access

Interactive comment on "Estimation of uncertainties due to data scarcity in model upscaling: a case study of methane emissions from rice paddies in China" *by* W. Zhang et al.

Anonymous Referee #1

Received and published: 17 January 2014

This study presents the methodology to upscale the uncertainty in model outputs associated with data sharing and spatial resolution taking a methane emission model for paddy rice applied to China as the example. Model upscaling and the quantification of model uncertainty due to upscaling themselves are commendable aims and of interest for broader modeling communities. However, after reading twice, the following concerns came out:

1. The concept of data sharing is in question. Fig. 1 itself is clear and I could roughly understand what the authors proceeded. However, what is the difference among the following three: (1) input at a site is shared by that at a neighboring site; (2) the identical value is used for multiple sites because only coarser-resolution data are available; (3)

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper

and there is less spatial variability of a variable of input. In my understanding, (1) and (2) are the same. If so, the simplest approach is to create inputs at difference grid sizes from the finer input and run the model at each scale. Otherwise the impacts of decreased variability of input (coarser-resolution input in general have less variability than finer-resolution input) on model output have never quantified. I don't understand why the authors go such a tricky way that aggregates model outputs using correlation between two cells.

2. Theoretically, a model should be calibrated at any scale before its application. As far as I read Haung et al. (1997) describing the modeling of CH4MOD, the structure of this model is very simple and is almost a multiplicative product of a few (five or so) variables. Why don't the authors directly calibrate empirical parameters of the model using coarser-resolution data? Also even for a complex process-based model (though the model is not for methane emission), the methodology to incorporate the information on subgrid-scale heterogeneity into a coarse-resolution model is feasible (e.g., Iizumi et al., 2013).

3. There is no mentioning of the importance of parameter values of a model as an important source of uncertainty of model output (e.g., Kennedy and O'Hagan, 2001; Makowski et al., 2002). This is especially true for empirical model like CH4MOD.

Specific comments: 4. P182L8. "five of the most sensitive model variables" is misleading. As long as I read Haung et al. (1997), this CH4MOD model has a few (five or so) variables as a whole.

5. P183L23-25. I suggest adding the number of harvesting of rice crop in a year and historical crop calendar here. These variables are important in estimating methane emission from rice paddy and have been changing with time (e.g., Tao et al., 2003).

6. P186L6. Does V_i(M_i,sigma_i) indicate a normal distribution?

7. P186L17-18. Could you elaborate how you can distinguish high correlation due to

similar geographical condition across sites and that due to the sharing of an identical data value associated with the limited availability of finer-resolution data?

8. P186L20. "share common data" would be more precise to say "share common data and modeled processes".

9. P187L21. This "DS matrix" is the first time to appear. It is more appropriate to write "data-sharing (DS) matrix" here if my understanding is correct.

10. P188L19. "semi-empirical" Please be consistent throughout the manuscript (P195L19 says "empirical").

11. P189L24. It is more readable to say "grand region scales" than "regional scales".

12. P190L1. The "CI" is the first time to appear as far as I read. Please replace this by "confidence interval (CI)" if my understanding is correct.

13. P190L25-28. I have no specific criticism to use a Gunma distribution; however, I wonder if a use of bootstrap method is more useful to avoid a fitting error of a Gamma distribution because a Gamma distribution fails to capture the frequency of low methane emission values (Fig. B1) and this study focuses on the uncertainty of mode output (you have a lot of model output data, indicating that a fitting is not necessarily essential).

14. P191L3. You can delete "1M=10ˆ6".

15. P193L3-5. I don't understand which result support this statement. Figure 2 doesn't work for this. This seems to be contradictory with the results presented in Table 4. The same comment can be applied to P193L14-15.

16. P194L26-P195L2. It would be nice if the author present this in a figure.

17. P194L22-23. However, both I_DS and I_R require finer-resolution data to calculate.

18. P197L1-2. I think the spatially interpolated temperature is an important source of

GMDD
7, C3–C6, 2014

uncertainty of model output. Why don't the authors include this into the analysis?

19. P197L5. It is well-documented that crop phenology including rice in China changed with time (e.g., Tao et al., 2006). Is this factor considered for the estimation of national-level methane emission? The estimated value appeared in P182L9 is tied to the reliability of input data. It is recommended the authors mention this point clearly in Abstract.

20. P198L7-9. Please elaborate how you select these papers from which sources.

21. Fig. 3. As far as I read, Fig. 3 is not cited in the main text.

Iizumi et al. (2013) An ensemble approach to the representation of subgrid-scale heterogeneity of crop phenology and yield in coarse-resolution large-area crop models. Journal of Agricultural Meteorology, 69, 243-254. Available at: https://www.jstage.jst.go.jp/article/agrmet/69/4/69_69.4.2/_article

Tao et al. (2006) Climate changes and trends in phenology and yields of field crops in China, 1981–2000. Agricultural and Forest Meteorology, 138, 82–92.

Kennedy and O'Hagan (2001) Bayesian calibration of computer models. J. Roy. Stat. Soc., Ser. B, 63, 425-464, doi: 10.1111/1467-9868.00294.

Makowski et al. (2002) Using a Bayesian approach to parameter estimation; comparison of the GLUE and MCMC methods. Agronomie, 22, 191-203.

Interactive comment on Geosci. Model Dev. Discuss., 7, 181, 2014.