Dear editors,

The manuscript (No.: bg-2016-250) entitled “Uncertainties in the national inventory of methane emissions from rice cultivation: field measurements and modeling approaches” by Wen Zhang, Wenjuan Sun and Tingting Li has been revised according to the comments from the reviewers. We greatly appreciate the two anonymous reviewers for their helpful comments that led to improvement of this paper.

Detailed description of the revision is given in an appended **Response to the Reviewers** and **List of Changes Made in the MS.** Should you have further comments, please let me know. Thank you for your consideration.

Sincerely,

Wenjuan Sun

Institute of Botany
Chinese Academy of Sciences
Beijing 100093, China
Tel: +86 10 6283 6052
Fax: +86 10 8259 6146
E-mail: sunwj@ibcas.ac.cn

Response to Reviewer #1:

The manuscript provides a comprehensive analysis of the sources of uncertainty in the national inventory of methane (CH4) emissions from rice agriculture in China. Three approaches were used to estimate the inventory and the associated uncertainties (i.e. direct field measurements, two empirical regression models, and the process-based model, CH4MOD). Additionally, the sensitivity of the levels of uncertainty using each approach to various scenarios of data scarcity was assessed. The more complex, process-based model had the lowest total error compared to the two empirical models. All approaches had higher error when average values were used for input data compared to case-specific values, highlighting varying degrees of model instability to insufficiency of supporting data. Interestingly, even when no case-specific input data were used in the processed-based model CH4MOD, it still had lower total error than the least complex empirical model when all case-specific input data were used (i.e. organic matter input modified by water regime). This in-depth comparison of approaches, their associated errors, and the sensitivity of the errors to input data availability is a significant contribution to the scientific community. It examines very relevant issues and challenges that modelers are faced with when scaling up field-validated models to larger spatial scales. The manuscript nicely quantifies and discusses the trade-offs associated with using the different approaches. It also outlines a method for assessing various sources of uncertainty and distinguishing model structural uncertainty from the uncertainty in input data.

Re: We greatly appreciate the reviewer's comments on the scientific significance of the study.

-There is no mention of total estimated national CH4 emissions using each approach in the abstract. I actually I think the estimation of national CH4 emissions using the empirical models is missing from the whole paper. It seems like this is a major comparison to include in the paper and highlight in the abstract. Instead, the average CH4 emissions and 95% confidence intervals of the mean are reported. I think a comparison of the national CH4 emissions and their respective 95% confidence intervals for each approach and data-availability scenario is a very important application of this analysis and should be in the abstract. Similarly, I think it is important to highlight which case specific data (e.g. organic matter inputs, water regime, or soil properties) mattered the most in terms of its effect on uncertainty when it was omitted.

Re: Many thanks to the suggestion that "a comparison of the national CH_4 emissions and their respective 95% confidence intervals for each approach and data-availability scenario is a very important application of this analysis and should be in the abstract". In the revision, we have made the comparison of the national CH_4 emissions and their respective 95% confidence intervals for each approach and data availability scenarios. The results of the comparison were showed in Table 3 and the description of the results was also added in the main text (P13 lines 3-12 in the ‘clean revised manuscript’).

In Table 3, the estimated national CH_4 emissions ranged from 6.43 (3.79–9.77) Tg to 13.59 (1.45–38.98) Tg for the M-S0 scenario R1-S0 scenario, respectively. The 95% CIs of the national estimation differed more greatly among the approaches than those among the data availability scenarios of each approach. As an indicator of the trade-off between the complexity of the approach and data availability, the σ_d/σ_{b+v} ratio in Table 3 was 0.87 for M-S0, closer to 1 than those for the other approaches and scenarios, which also yielded the narrowest 95% CI in Table 3.

The factors affecting methane emission from rice paddies (e.g. organic matter inputs, water regime, or soil properties) had been incorporated into CH4MOD as input variables. The importance of those factors on uncertainty had been discussed in a previous study (Zhang et al., 2014). Stating briefly, the factor of high sensitivity will result in larger uncertainty when omitted, from water regime down to soil properties and organic matter inputs.

As suggested by the reviewer, we also add statement of the total estimated national CH_4 emissions in the revised abstract (P1 lines 28-29).

-Overall the paper is lacking in citations of current research articles. Most articles cited are >10 years old.

Re: The topic of the present study, uncertainties in the modelling approaches closely related to methane emissions from rice paddies and the relevant, had been dedicatedly discussed in few previous studies (Ogle, et al., 2010; van Bodegom et al, 2002a). In the present study, we compared performances of CH4MOD and two empirical methods that had been developed and utilized in early days (Neue et al., 1990; Khalil et al., 1991, 1993; Bachelet et al., 1995; Kern et al., 1995, 1997), and had to reach out to studies 10-20 years ago. We, however, didn’t omit relevant studies in recent years, e.g., the study of mitigating methane emission from rice cultivation by gene transcription (Su et al., Nature, 2015), the study of methanogenic community structure involving methane production (Singh et al., SBB, 2012), and national/global estimation of methane emissions from rice paddies and wetlands (Chen et al., GCB, 2013; Ren et al., Tellus B, 2011; Zhang et al., GCB, 2011). In the revision, we referenced major results of the recent studies concerning methane emission from rice paddies (Ito et al., 2012; Tian et al., 2016; Weller et al., 2016; Zhang et al., 2016; Dijkstra et al., 2012).

Missing description of model calibration of the two empirical models and CH4MOD.Thus, it’s unclear whether data used for model validation (i.e. comparison to measurement-based estimations of fluxes) and uncertainty analyses are independent from data used to calibrate the internal model parameters.

Re: The approaches in the study had been used in previous studies (Bachelet et al., 1995; Kern et al., 1995, 1997; Zhang et al., 2011) to estimated methane emissions from rice paddies on regional, national and global scales. When analyzing the performances of the approaches in the present study, we validated them with data excluding those had been used for calibration to maintain the independence between the validation and calibration. We explicitly addressed the situation in the revised MS (P5 Lines 23-24).

It's unclear whether the direct measurements used in the analyses are cumulative CH4 emissions or daily CH4 fluxes from the same experimental plots. If it's the latter, then the errors are not independent, and this issue should be explicitly addressed in the paper. The issue of non-independence of errors was discussed, but it was unclear whether this was due to measurements taken in close proximity versus repeatedly from the same location.

Re: All the measurements of CH_4 emission in the present study are cumulative CH_4 emissions over the period from rice transplanting to harvesting. We explicitly stated it in the revised MS (P10 Lines 13-14). We discussed the non-independence of the measurements due to spatially close proximity in Section 4.1, when no temporal dependence of the daily measurements involved.

Additional comments, questions, and technical corrections:

P 1, Lines 26-29: Revise to account for the exception in which M-S3 performed better than R1-S0 (Table 2).

Re: We revised the sentence as "Comparisons revealed that the CH4MOD model may perform worse than the comparatively simple regression models when no sufficient input data for the model were available".

P 1, Line 33: Do you mean "between-grid variations", i.e. differences among grid cells?

Re: It is the within-grid variation calculated via the Monte-Carlo method. To make it clearer, we revised the sentence as "the within-grid variations, $\sigma_{T,i}$, were found to be 81.2%–95.5% to the grid cell means (F_i)."

P 2, Line 2: I think a slight rewording should be made, i.e. "Reducing the total uncertainty in the national methane inventory depends on a better understanding of both the complexity of the mechanisms of methane emission and the spatial correlations of the factors that influence methane emissions from rice paddies."

Re: Thanks for the suggestion. The sentence was revised.

P3, Line 16: Reference needed.

Re: The appropriate literature references were added.

P 6, Line 12: Provide detail on the parameters and assumptions for substrates produced from added organic matter and root exudates.

Re: We added sentences to briefly describe the substrate production from added organic matter and root exudates in the revised Supporting Information (Appendix B). The amount of the substrate derived from rice root exudate was simulated by a power function of the rice biomass, scaled by the parametric influence of the soil context and the rice cultivar. The substrate derived from the added organic matter was calculated by a first-order kinetic decomposition equation of the organic matter in soil, also scaled by the parametric influence of the soil context and the temperature. Details can be found in Huang et al (2004).

P 7, Line 9: Can you provide a reference or derivation of equation 4?

Re: We detailed the derivation of Equation 4 in the revision.

P 7, Line 10: Given that the focus of the manuscript is on uncertainty in national inventories, it seems that the methods section should be framed under national-level uncertainties as opposed to regional-level. It's my understanding that national inventories represent an aggregation of multiple regions. Thus, perhaps the section title here should state "national scales" as opposed to "regional scales", and translate this distinction into the text that follows.

Re: Thanks for the suggestion. We revised it throughout the section and other places in the MS.

P7, Lines 28-33: Please clarify each step of the process in which SAND data were obtained. What method of interpolation was used (e.g. ordinary kriging, inverse distance weighted)? What is meant by "missing spatial variation" in your dataset – how was this determined and quantified? Were some grid cells missing survey data all together?

Re: Soil properties have extremely high spatial variation and may vary largely from one place not far from another. We obtained the data from Institute of Soil Sciences, Chinese Academy of Sciences, as indicated in the MS. They collected more than 7000 soil profile measurements sampled during the period from 1980s to the present and linked them to the a soil database of 1:1,000,000 scale (Shi et al., 2004), and produced the gridded data of soil properties with geostatistical methods. We compared the spatial variation explained in the gridded datasets of soil properties against the variations in the profile measurements to analyze the 'missing spatial variation' (Bodegom et al., 2002b). The 'missing spatial variation' is the proportion of spatial variation of the soil properties (the sand content of the surface soil layer in the present study) that were not accounted for by the gridded datasets. We used the missing variation to build the PDF of SAND in Monte Carlo simulation by assuming normal distributions of the missing variation. We added the brief description of the soil property datasets in Appendix B.

P 9, Line 1: Please provide a reference or derivation of equation 5.

Re: Equation 5 is derived from Equation (C9) in the Supporting Information (Appendix C) and Equation 4 in the main text, when used in each grid cell. We added the description and derivation in revision (P9 Lines 27-30 and P10 Lines 1-10).

P 9, Line 9-10: You refer to the "three components of the estimation uncertainties" in equation 5. I assume you are referring to (1) (Fj x Br)2, (2) (Fj x CV)2, (3) _DJˆ2, which is analogous to the three terms in equation 6. Can you please provide a meaningful definition of what each of these components of uncertainty represent? Later in the discussion you explain that (Fj x Br)2 +(Fj x CV)2 represents model fallacy, while _DJˆ2 represents uncertainty due to input data. I think including this type of description in the methods section would be helpful to read leading into the results section.

Re: Thank for your suggestion. We added explanation of the terms in the revised MS in the method section. In Equation 5, $\sigma_{d,i}{}^2$ signifies the uncertainty caused by the error and availability of data, $(F_i \times r_b)^2$ represents the modelling bias, and $(F_i \times r_v)^2$ represents the rest parts of the model fallacy error apart from $(F_i \times r_b)^2$. We provided more details of the derivation and explanation of Equation 5 in the Supporting Information (Appendix C) in more general terms than the main text. The three components in Equation 5 do correspond to those in Equation 6 and the derivation from Equation 5 to Equation 6 was also provided in the Supporting Information (Section D).

P 9, Lines 30-31: Explicitly state the water regimes.

Re: Revised.

P 10, Line 27: What "estimated CH4 flux" are you referring to? Are you referring to an example of a single flux? If so, I would start the sentence with: "For example, in one case the modeled CH4 flux was . . ., while the measured flux was . . ."

Re: Thank you. We revised the sentence (P11 Lines 28-30).

P 11, Lines 16-18: Specify which model the simulated fluxes are based on. Please clarify this in Fig. 6 and Table 3 as well.

Re: Thank you for the comments. We added information to specify the model used (P12 Lines 19-20) and the caption of Fig. 6.

P 12, Lines 13-14: Didn't the authors also apply the two regression models to the 10 x 10 km grids? A comparison to the other two approaches (direct measurements and process-based model) should be discussed here.

Re: The two regression models were not used to the 10 x 10 km grid in the BGD version of the MS. As suggested by the reviewer, we applied all the three models and data availability scenarios in the revision and list the results in the revised Table 3, focusing on the national CH_4 emission and the relevant uncertainties. While there are no measurements on grids, comparison of the estimation by modelling can only be carried among the models (Table 3 and P13, Lines 3-12) instead of against measurements.

P 12, Lines 26-29. Nice explanation!

Re: Thank you.

P 14, Lines 19-33. See comment above for P 7, Line 10. Reframe conclusions to include national estimates and uncertainties at the broadest level of discussion.

Re: We revised the conclusion and discussion section and added information of the national estimates in both the abstract and the conclusion.

Response to Reviewer #2:

I agree with referee #1 that the paper 'Uncertainties in the national inventory of methane emissions from rice cultivation: field measurements and modeling approaches' by Zhang et al. is an important and nice study regarding general uncertainties evolving during regional/national GHG emission inventories. I also agree with referee #1 that national estimates of CH4 emissions should be more emphasized. My main criticism relates to the presentation of the study. Material and Methods, Results and Discussion sections all need revisions in order to improve the reader's access to the main points of this study (see specific comments).

Re: We thank the reviewer for the comments and made revision to the MS to show the national estimations of CH_4 emission, and also the writing of the MS emphasizing the formulation and nomenclatures.

Specific comments:

P1 L25: Mention that regression models are taken from literature.

Re: Revised (P1 Line18 in the 'clean revised manuscript').

P1 L27-28: Use clear measures and give respective values instead of using the vague term 'model performance' only.

Re: The 'model performance' refers to how the model representing the variation in the observations, evaluated by the difference between the observations and the corresponding model outputs. Conceptually, the model performance here covers the 'parameter uncertainty' and 'model inadequacy' in Kennedy and O'Hagan (2001) and errors in observations, because we can't distinguish them with model validation, which was used to evaluate the 'model performance' in the present study. The measures to quantify the model performance here are two statistical parameters of the modelling residuals (difference between the observations and modelling outputs): bias (means of the residuals) and variance (statistical variance of the residuals) as showed in Equation 1 and Equation 2. We revised the equation and the main text to make the meaning of the terms clearer.

P1 L30: Absolute values of simulated methane fluxes are meaningless here since context (e.g., different irrigation, straw management, ...) is not clear yet.

Re: The modelling result here is the result of CH4MOD with available information of irrigation, straw management and soil properties of paddies in rice cultivation of China. We revised the sentence as 'As simulated by CH4MOD with data of irrigation, organic matter incorporation and soil properties of rice paddies, the modelling methane fluxes varied from 17.2 kg CH_4 ha^{-1} to 708.3 kg CH_4 ha^{-1}'

P4 L19-21: Statement is not very intuitive. Why should 'non-key' factors lead to significant errors? Factors leading to significant errors are implicitly named key.

Re: Here we intended to say that they were not 'non-key' at all. To avoid misunderstanding, we revised the sentence by replacing 'non-key' with 'other'.

P5 L21-22: Imprecise formulation, inaccuracies of models are manifold and should be defined more clearly based on common nomenclatures in literature, see for example nomenclature and definitions by (Kennedy and O'Hagan, 2001). Nomenclatures and definitions should be revised and standardized in many parts of the paper. Kennedy, M.C., O'Hagan, A., 2001. Bayesian calibration of computer models. J. R. Stat. Soc. Ser. B Stat. Methodol. 63, 425–464. doi:10.1111/1467-9868.00294

Re: We thank the reviewer for the comments and recommending the literature. Here, the model inaccuracy refers to the combination of 'model inadequacy' and 'parameter uncertainty' in Kennedy and O'Hagan (2001). In other places of the MS, 'model fallacy' means the same. We revised the MS to use the term 'model fallacy' throughout the MS and explain explicitly the mean of it. We also rewrite the equations of the MS.

P5: L34-36: Why were these two regression models chosen? It would be very interesting to see how IPCC emission factors, which also account for, e.g., different amounts of straw and different irrigation schemes would behave.

Re: One of the objectives of the study was to compare the performance of models with different complexity with different levels of data availability. We chose the two regression models because: 1) they had been used to estimate regional/national/global methane missions in many previous studies, and 2) they differed from each other and from CH4MOD explicitly in levels of complexity. There are many other models that developed and used widely in modeling methane emissions from rice paddies and wetlands *etc*. But we can't tell which one is more complicated in structure that the other. We briefly explained it in the revise MS (P6 Lines 13-15).

P6-7 Formulas 1-4: Unclear why these measures have been used. Give proper descriptions, meanings and references to 'bias' and 'total error' and compare both to each other.

Re: 'bias' is the statistical mean of the modelling residuals. We admit that 'total error' is not a proper term. In the MS, it means the 'mean of squared errors' in model validation. In the revised MS, we used the term 'mean of squared errors' in the main text.

P7 L12: ' errors in the performance of the method': unclear formulation, use consistent nomenclature for different error/uncertainty sources

Re: revised as 'model fallacy'. Because in the MS, our emphasis was on the quantification of the uncertainty in the national inventory by modelling approaches, the rationale of the uncertainty was mainly provided in the Supporting Information (Appendix C and D).

P7 L 15: Give more information regarding your Monte Carlo simulation and PDFs since this is an important determinant of posterior uncertainty.

Re: To measure the uncertainties in model outputs due to insufficient data quality and availability, we applied Monte Carlo simulations to the CH4MOD model. Statistical characteristics were derived from the available datasets to develop probability distribution functions (PDFs) for each model input variable. The PDF of field irrigation were defined by the occurrence percentage of each irrigation pattern (Table B2). Table B1 shows the statistical parameters of the PDF (normal distribution) of organic matter incorporation in each province. The PDF of the soil sand percentage was also built as normal distribution with parametric information from the literature (Shi et al., 2004).

We performed Monte Carlo simulation in the way of randomly drawing values of the model input variables from their PDFs and then run the model (e.g. CH4MOD). This process iterated 1000 times and at the last step, the mean and 95% CI of the calculated methane fluxes were derived from the iterations (P8 Lines 3-6).

P8 L1 On what is this assumption (amount of stubble) based?
P8 L1-11: What is the difference between stubble and incorporated straw?

Re: Stubble is the part of rice stem that left after rice harvesting. Traditionally, both the rice grain and rice straw were harvested and stubble was left in field. The harvested straw may be taken away or left in field, but stubble was always left. The amount of stubble accounts for about 10% of the aboveground biomass of rice according to previously published literatures (Huang et al., 2004; Zhang et al., 2011). We noted the literature in the revised MS (P8 Line24).

P8 L15-16: Be more precise here and mention considered irrigation schemes and how the model handles them.

Re: The irrigation in rice cultivation in China was summarized into five patterns: 1) flooding-drainage-flooding-intermittent irrigation, 2) flooding-drainage-intermittent irrigation, 3) flooding-intermittent irrigation, 4) continuous flooding and 5) continuously intermittent irrigation (Gao and Li, 1992; Huang et al., 2004). Appendix B in the Supporting Information provides necessary information of the irrigation in China. Table B2 list the percentage of each water pattern in different regions of China. More information of how CH4MOD handles the irrigation may refer to the literature of the model development (Huang et al., 2004). We also add brief description of the irrigation in the revised MS (P8 Lines 32-34 and P9 Line 1).

P8 L28: Probably Appendix B is meant.

Re: Revised.

P8 L33: In section 2.4, the description of used formulas should be improved since the combination of model and model input uncertainty is a central point of this study. The derivations of formulas in the Appendix are unclear. Give consistent names and meanings to each symbol that is used. Parts in the Discussion sections refer to the meaning of formulas and measures and should be moved here.

Re: We revised the relevant part of the MS and the Appendix, emphasizing the consistence of the names and expressions.

P9 L27: Do you mean 'harvested-area-weighted' or 'cultivated-area-weighted'? Since cropping intensity (number of crops per year) varies, the weighted mean should be derived based on harvested area. In addition to area weighted means you should also consider seasonal means. A given amount of data may refer to different seasons, e.g., winter, spring, summer and autumn with strongly varying potentials of CH4 emissions. Most likely the seasonality distribution of observations does not correspond to the actual seasonality distribution of rice cultivation in China.

Re: Yes. The 'area-weighted' in the MS means 'harvested-area-weighted' and we revised the expression.

In China, the rice cultivation is different from north to south: single rice cultivation in north-eastern China, rice-upland crop rotation in eastern China and double rice cultivation in southern China. The 'harvested-area-weighted' analysis in the present study distinguished the harvested area of different water irrigation, because irrigation is the most important factor for methane emission. Seasonality also affects the methane emission but not as important as irrigation, according to both observational and modelling studies (Yan et al., 2005; Zhang et al., 2011). We agree with the reviewer that mismatch between the seasonality of the observations and the actual rice cultivation in China may bias the national estimation of CH4 emission via the statistical summation of the observations, and contributes to the uncertainty of the estimation.

P10 L18-21: Be more precise how measurements are dependent from each other. The potential dependency of measurements is not discussed in the Discussion section.

Re: The dependence of measurements here means the possible spatial correlation among them because of the common environmental conditions they may share. It is not the meaning that they were dependently obtained by sampling. The spatial aggregation of the measurements obtained at different places to produce national estimations may introduce biases if the spatial correlation among the measurements were not handled properly. We didn't make in-depth discussion about the spatial correlation because it is beyond the topic of the present study. In the revision, we revised the sentence to clear that it is about the spatial correlation. We also provide literature reference (Legendre, 1993; Dormann et al., 2007) for those interested in spatial correlation.

P10 L12-16: Standard Error (SE) and deviation are very common measures and do not need explanations/references. To my understanding, the presented SE refers to the variability of different observed mean fluxes from different field sites. How are measurement errors reflected? What do you mean with representative error?

Re: Yes, we agree with you concerning the standard error. The measurement errors were not discussed separately in the present study. The reason of doing so was explained in Appendix C. The representative error in the present and other literatures (e.g., Van Bodegom et al., 2002a; Verburg et al., 2006) stands for the representativeness of the measurement obtained at a site to the area that enclose the site.

P10 L25: Present average values of overestimations for both models.

Re: Revised.

P10 L35: Why is ’total error’ and not ’bias’ interpreted as model performance? In order to underline this statement, more measures should be used, e.g., root mean squared error, R2, model efficiency.

Re: ‘bias’ is the average of the modelling residuals, accounting part of the errors. We use ‘mean of the squared errors’ to interpret model performance. ‘total error’ is not a proper expression and we replace it with ‘mean of the squared errors’ in the revision. There are other indexes, e.g., R^2 and RMSE, we use bias and ‘mean of the square errors’ in the MS owing to they are directly comparable to the errors from data availability.

P11 L22: I miss the discussion of these values. Are such uncertainties small or large compared to other studies?

Re: The within-grid estimation error ($\sigma_{T,i}$, calculated with Equation 5) is the error in each grid cell due to both the model fallacy and data scarcity when making estimation of a grid cell (10 × 10km). They are not shown in details because we emphasized the uncertainty in the national inventory, which was the spatial aggregation of the uncertainty in each grid cells. We didn’t compare the result of the ‘within-grid estimation error’ in the present study with other studies because no study had make estimation of the uncertainty in the way of the present study.

P11 L23-25: Discussion is missing.

Re: In the revision, we discussed the difference of the national methane emissions and the uncertainties estimated with different approaches and the data availability scenarios, as showed in the revised Table 3.

P11 L30 - P12 L14: This is rather introduction and representing of results than discussion.

Re: Thanks for this comment. We revised the MS by moving it to the introduction.

P12 L10: Temporal variations are not presented.

Re: Here in the sentence, we noted that there are temporal variations, annual, seasonal and even diurnal, in the methane emissions. But in the present study, we discussed the spatial variation and the estimation uncertainty in the national inventory of a specific year. Temporal variations of the methane emission were not discussed.

P12 L15-17: Unclear argumentation.

Re: Thank you for pointing it out. We revised the sentence as ‘This was partly due to the discrepancy in the spatial representativeness of the methane fluxes in field observations and model estimations’.

P12 L18-22: Unclear argumentation. Model performance was assessed with site-specific input and not with regional averages. The representation of experimental measurements for larger regions and associated uncertainties should be independent of models. Discussion of comparison between model and measurements at site scale could be moved to a separate subsection.

Re: This is what the ‘representative error’ means, which had been discussed in previous studies (Verburg et al., 2006; Van Bodegom et al., 2002a) and described in Appendix C of the MS. Model performance was assess with site-specific input. Here the ‘site’ means a small scale (e.g., a hectare or smaller) instead of a ‘point’, when the experimental sampling was taken at several ‘points’ called ‘duplicates’ at the experimental site. When we use the model for regional estimation, we make estimations for each grid cell (10 × 10 km in the present study). The mismatch of the scale supports the meaning of ‘representative error’.

P12 L31-33: Should be moved to the Results section. Use consistent nomenclature, i.e., the term ’model fallacy’ has not been used beforehand. Do not repeat formulas from the Material and Methods section in the Discussion.

Re: Thanks for the suggestion, we revised accordingly (P12 Lines 24-26).

P12 L35 - P13 L14: Much of this information belongs to the Material and Method Section and to the Discussion. Key results (e.g., ’56.6% of total uncertainty originates from the model’), which are also presented in the abstract should be first presented in the Results section and subsequently discussed. Appropriate discussion regarding the different uncertainty sources (model versus input) is missing. Argumentation regarding ’imprecision random noise and/or unknown factors’ is unclear.

Re: Section 4.1 discussed the different error sources to the uncertainties in the inventory. This paragraph around Fig. 8 was about the aggregation of $\sigma_{v,i}^2$. Material and Method Section described how the errors were quantified and aggregated, as showed in Fig. 2. We thank the reviewer for the revision suggestion and revised the MS accordingly.

P13 L27 - P14 L16: Remove this section from the Discussion. This is partly Material and Methods and seems to be an arbitrary example of model parameter uncertainty that has been neglected and thus is not much contributing to this study.

Re: Section 4.2 discuss how model improvement (e.g., parameterizing rice cultivar more specifically) affect the uncertainty analysis. We agree with the reviewer that the model parameter uncertainty wan not separately analyzed in present study. But because the parameter uncertainty contributed significantly to the model fallacy, it should be noted briefly in the discussion.

Fig. 5: Use identical axes for all plots.

Re: We guess you meant Fig. 5. We had at first used identical axes for Fig. 5-(a), Fig. 5-(b) and Fig. 5-(c). But it looked a little awkward, we, therefore, changed the y-axe of Fig. 5-(a) and kept the other axes identical.

List of changes made in revision

1. In response to the comments of Reviewer #1, comparison of the national CH_4 emissions and their respective 95% confidence intervals for each approach and data availability scenarios was made. The results of the comparison were showed in Table 3 and the description of the results was added in the main text.

2. By referencing to literatures, especially the paper of Kennedy and O'Hagan (2001) recommended by Reviewer #2, we changed the nomenclatures and formulation of equations in the main text and the Appendixes of the Supporting Information.

3. The results of recent studies concerning methane emission from rice paddies was addressed in the revision in accordance to the suggestion of Reviewer #1.

4. A briefly description of the substrate production from added organic matter and root exudates, as well as the data of soil properties were added in the revised Supporting Information (Appendix B) as suggested by both the two reviewers.

5. Overall use of consistent terminology throughout the text has been checked.

6. We changed the order of authors due to their contributions during MS revision.

Besides, there have been other changes in response to specific comments of the reviewers and all the details were marked in the revised manuscript with 'MS word tracking' below.

Uncertainties in the national inventory of methane emissions from rice cultivation: field measurements and modeling approaches

Wen Zhang[1], Wenjuan Sun[2*], Tingting Li[1]~~, Wenjuan Sun[2*]~~

1 LAPC, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China
2 LVEC, Institute of Botany, Chinese Academy of Sciences, Beijing, China
**Correspondence to*: sunwj@ibcas.ac.cn

Abstract. Uncertainties in national inventories originate from a variety of sources, including methodological failures, errors and insufficiency of supporting data. In this study, we analyzed these sources and their contribution to uncertainty in the national inventory of rice paddy methane emissions in China and compared the differences in the approaches used (e.g., direct measurements, simple regressions and more complicated models). For the 495 field measurements we collected from the scientific literature, the area-weighted 95% CI ranged from 13.7 to 1115.4 kg CH_4 ha^{-1}, and the histogram distribution of the measurements agreed well with parameterized gamma distributions. For the models, we compared the performance of methods of different complexity (i.e., the CH4MOD model, representing a complicated method, and two less complex statistical regression models taken from literatures) to evaluate the uncertainties associated with model performance as well as the quality and accessibility of the regional datasets. Comparisons revealed that the CH4MOD model may perform~~ed~~ ~~better~~ worse than the comparatively simple regression models ~~only~~ when no sufficient input data for the model were available~~, with the regression equations performing better otherwise~~. As simulated by CH4MOD with data of irrigation, organic matter incorporation and soil properties of rice paddies, the modelling methane fluxes varied from 17.2 kg CH_4 ha^{-1} to 708.3 kg CH_4 ha^{-1}, covering 63% of the range of the field measurements. When applying the modeling approach to the 10 km × 10 km gridded dataset of the model input variables, ~~within~~the within-grid variations, made via the Monte Carlo method, were found to ~~represent~~ be 81.2%–95.5% ~~of~~ to the ~~modeled~~ grid means~~ fluxes~~. ~~Moreover, u~~Up-scaling the grid estimates to the national inventory, the total methane emission from the rice paddies was 6.43 (3.79–9.77) Tg. ~~resulted in the models~~ The fallacy of CH4MOD contributed~~ing~~ 56.6% of the total uncertainty, with the remaining 43.4% being attributed to errors and the scarcity of the spatial datasets of the model inputs. Our analysis reveals the dilemma between model performance and data availability when using a modeling approach: a model with better performance may help in reducing uncertainty caused by model fallacy but increases the uncertainty caused by data scarcity, as greater levels of input are needed to improve performance. Reducing the total uncertainty in the national methane inventory depends on a better understanding of both the complexity of the mechanisms of methane emission and the spatial correlations of the factors that influence methane emissions from rice paddies~~Reducing the total uncertainty in the national methane inventory depends both on a better understanding of the complexity of the mechanisms of methane emission and the spatial correlations of the factors that influence methane emissions from rice paddies~~.

Keywords: Uncertainty, source and contribution, spatial variation, national inventory, methane emission

1 Introduction

Rice cultivation is a major source of anthropogenic methane and a prime target of greenhouse gas mitigation efforts (Tian et al., 2016; Smith et al., 2008). Globally, the methane emission from rice cultivation was about 18.3Tg CH_4 yr^{-1} under intermittent irrigation and 38.8Tg CH_4 yr^{-1} under continuous flooding in the 2000s (Zhang et al., 2016). Methane fluxes in rice paddies varied extensively with environmental and agronomic factors. Certain factors, such as rice biomass (Bachelet and Neue, 1993), organic matter input (Kern et al., 1995), water management (Khosa et al., 2011;Mishra et al., 1997), paddy soil properties (Yao et al., 1999;Gaunt et al., 1997), climate ~~(~~ Sass et al., 1991) and rice varieties (Su et al., 2015; Ding et al., 1999), have ~~previously~~ been recognized as having significant impacts on methane emissions. Other factors, such as atmospheric CO_2 and ozone contents (Dijkstra, et al., 2012; Bhatia et al., 2011;Inubushi et al., 2011), N fertilizer application (Banger et al., 2012;Xie et al., 2010a) and active soil organic C (Zhan et al., 2011), and even the field management of the previous crop that rotating with rice (Weller et al., 2016) are also receiving increasing attention. Because so many factors affect the production, oxidation and emission of methane from rice cultivation, the observed methane fluxes varied extensively both spatially and temporally.

Numerous methods have been applied for estimating national and global inventories of rice paddy methane emissions, including meta-analysis of direct measurements, process models and empirically based statistical models. However, the range of ~~regional~~national/global source estimates remains large (Cao et al., 1996;Sass et al., 1999;Chen et al., 2013). The major factors that are known to regulate rice paddy methane emissions include agricultural management practices (Khosa et al., 2011;Sanchis et al., 2012;Sass et al., 1992;Bodelier and Laanbroek, 2006) and environmental conditions, such as climate and soil properties (Conrad et al., 2007;Inubushi et al., 2011;Sass et al., 1991). Currently, techniques for calculating methane emissions differ substantially and ~~are far from perfect~~usually in the way of scenario simulations (Ito et al., 2012; Van Bodegom et al., 2002a, b; Verburg et al., 2006), ~~mainly because of issues involving~~without integrated consideration of methodological fallacy and data insufficiency.

By extrapolating field measurements obtained from experiments, methane emissions from the 30 million hectares or so of land under rice cultivation in China were estimated to range from 21.6 Tg CH_4 yr^{-1} to 30 Tg CH_4 yr^{-1} (Matthews et al., 1991;Taylor et al., 1991), much larger than the result of a recent study (Zhang et al., 2016). The extrapolation of methane emission rates from site measurements to larger regions is unlikely to yield reliable results because of the tremendous spatial heterogeneity in environmental conditions and agronomic activities (Ogle et al., 2010). Other studies have described the relationships between methane emissions and rice NPP (net primary productivity) (Bachelet and Neue, 1993) and organic matter inputs (Bachelet et al., 1995). Ambient temperature and the use of nitrogen (N) fertilizer have also been identified as determinants of methane emissions (Kern et al., 1995;Bachelet et al., 1995). Until the significant reduction in methane emissions caused by mid-season drainage was confirmed (Sass and Fisher, 1997;Yagi et al., 1997;Li et al., 2002;Yan et al., 2005), all previous regional and national estimates (obtained using extrapolation or regression equations) were derived from continuously flooded rice fields. More ~~recently,~~ factors such as the rice cultivar involved (Watanabe et al., 1995;Butterbach-Bahl et al., 1997;Ding et al., 1999;Inubushi et al., 2011), soil properties (Sass et al., 1994;Yao et al., 1999) and atmospheric CO_2 (Dijkstra et al, 2012; Xie et al., 2010b) and ozone ~~concentrations~~ (Bhatia et al., 2011) concentrations have also been incorporated into models designed to estimate methane emissions from rice paddies. Complex interactions among these factors have spurred model development (Cao et al., 1995;Li, 2000;Matthews et al., 2001;Huang et al., 1998;Van Bodegom et al., 2001;Huang et al., 2004). To delineate variations in methane emissions and to reduce uncertainties, the impacts of these factors on the production, oxidation and emission of methane were mathematically incorporated into the models. Models with ~~a greater number of~~more factors involved are able to reduce uncertainties in estimating methane emissions, but the estimates generated by these models still differ significantly across multiple spatial and temporal scales (Butenhoff et al., 2009;Ren et al., 2011;Chen et al., 2013).

Reduction of the uncertainty in estimated methane emissions requires the development of an effective and reliable model that incorporates various paddy environments and agronomic activities. However, our understanding of the complex biogeochemical processes that occur in paddy soils is poor. When estimating methane emissions from rice agriculture, only factors that are thought to be key determinants of methane emissions have been incorporated into the models. Excluding other~~“non-key”~~ factors introduces errors into the model output (Equations C2 and C3 in the Supporting

Information). Improving our knowledge of methane processes in the future will increase the number of factors that are integrated into models and potentially delineate details related to spatial/temporal variations.

Uncertainties in regional estimates of methane emissions from rice paddies stem not only from ~~deficiencies~~ fallacy in the applied models but also from errors and inadequate data, which we discussed in a previous study (Zhang et al., 2014; Appendix ~~C~~ D in the Supporting Information). A model with more factors generally performs better than a model with fewer factors but requires a larger amount of data to facilitate model ~~performance~~application. A model with good performance (less fallacy) can still result in large uncertainties when the available input data (e.g., soil properties, rice irrigation, types and amount of organic matter) are insufficient (Zhang et al., 2014; Ito et al., 2012).

In the present study, we analyzed the uncertainties in experimental measurements of methane fluxes in different rice paddies. We also evaluated the performance of different methods involving a diversity of input variables and the influence of data availability on the performance of these methods. Finally, the uncertainty in the national emissions inventory as a consequence of variable model performance and according to the quality and availability of input data were discussed.

2 Materials and methods

2.1 Field measurements of methane emissions from rice paddies in China

The observational data used in this study (Table 1) consisted of field methane fluxes measured at 33 sites (Fig. 1). We obtained these measurements from the published literature concerning all crop rotations with rice cultivation in China (double rice, winter wheat and rice rotation, single rice crop cultivation, and so forth) (Wei, 2012). A total of 495 measurements were taken at the 33 sites, after excluding those had been use for the model calibration (Neue et al., 1990; Kern et al., 1997; Huang et al, 2004). The amount of organic matter added to the rice paddies ranged from 0 t C ha^{-1} to 15.3 t C ha^{-1} and included animal manure, green manure, crop straw, biogas residuals and their various components. The applied water regimes consisted of continuous flooding, single mid-season drainage and multi-drainage irrigations.

Model performance was assessed by comparing the model estimates with the measurements. To drive the models, data pertaining to rice yields, soil properties and crop phenologies were collected from the relevant literature (Appendix B in the Supporting Information).

2.2 Performance of the methods used to estimate methane emissions

The uncertainties produced by the models derive from ~~inaccuracies in the models themselves~~model fallacy (Kennedy and O'Hagan, 2001, Appendix C in the Supporting Information) as well as from the quality and availability of data (Fig. 2). Model performance was assessed by comparing model outputs with the direct measurements (left part in Fig. 2). Errors in the input data of the model can be propagated in the obtained estimates (right side of Fig. 2, Appendix D in the Supporting Information).

Many techniques are available for calculating estimates of rice paddy methane emissions, such as extrapolation of measured emission rates (Khalil et al., 1991;Khalil et al., 1993), statistical regression equations (Bachelet et al., 1995;Kern et al., 1995;Kern et al., 1997) and the application of models of varying complexity (Cao et al., 1995;Matthews et al., 2001;Van Bodegom et al., 2001;Huang et al., 1998;Li, 2000). Here we chose two regression models (Neue et al., 1990; Kern et al., 1997) and CH4MOD (Huang et al., 2004) in that they differed explicitly in levels of structural complexity. We compared the performance of these methods under different levels of data availability (Table 1) using experimental field measurements as a point of reference (Fig. 1). In Table 1, R1 represents a simple regression equation in which the carbon (C) input is the sole predictor (Neue et al., 1990). Regression equation R2 is slightly more complicated in that it uses organic C and fertilizer N application as inputs (Kern et al., 1997). We assumed two data availability scenarios for R2. In R2－S0, both the C and N inputs are available; in R2－S1, only the C input is available (Table 1).

The third approach consists of a semi-empirical model, CH4MOD. This model was developed to simulate methane emissions from rice paddies under diverse environmental conditions and various agricultural practices (Huang et al., 1998;Huang et al., 2004). The input variables of the model include the climate, soil conditions, water management type, organic matter application and crop rotations. The model consists of two modules: the derivation of methanogenic substrates from added organic matter and rice root exudates and the production and emission of methane. Rice biomass is a key variable used to calculate the root exudates and the fraction of the methane emitted by rice plants and bubbles. The daily changes in the soil redox potential (Eh) were calculated according to various water manipulations conducted in the rice paddies (Xie et al., 2010b). The influences of other environmental factors, such as soil temperature and texture, on the decomposition of organic matter and the production of methane were expressed as specific coefficient functions (Huang et al., 1998). The input variables of the CH4MOD model (Appendix B in the Supporting Information) include the daily air temperature, soil sand percentage (*SAND*), organic matter amendment (*OM*), rice grain yield (*GY*), water management pattern (W_{ptn}) and rice cultivar index (*VI*).

Four model input scenarios (Table 1) were scheduled to evaluate the performance of CH4MOD under different levels of data availability. In M－S0, all of the model variables were assigned specific values. In M－S1, the application of organic matter was assigned the average value for all experiments, thus assuming a situation where no detailed information on organic matter application was available. In M－S2, detailed information on the water regime and soil properties was assumed to be unavailable. In M－S3, detailed information on all three major factors (organic matter application, soil properties and water regime) was assumed to be unavailable.

The estimation residuals (Δy, Equation 1), relative bias (r_b~~B_R~~, Equation 2) and coefficient of variations (r_v~~CV~~, Equation 3) were thus evaluated as follows:

$$\Delta y_k = \hat{y}_k - y_k, \qquad i = 1,2,\ldots,n \tag{1}$$

$$r_b = \frac{E(\Delta y)}{E(y)} \times 100\% \tag{2}$$

$$r_v = \frac{\sqrt{E((\Delta y)^2) - (E(\Delta y))^2}}{E(y)} \times 100\% \tag{3}$$

where *y* represents the measured methane fluxes; $\hat{y}$ is the estimate of *y*; and ~~*N*~~ *n* is the total number of measurements. *E*( ) indicates the statistical mean.~~, and *F* = *E*(*y*) is the mean of the measured methane fluxes (y_i)~~. The ~~total error~~ mean of the squared errors (MSE) of the estimation is calculated as follows:

$$\begin{aligned} MSE &= E((\Delta y)^2) = (E((\Delta y)^2) - (E(\Delta y))^2) + (E(\Delta y))^2 \\ &= (F \times r_v)^2 + (F \times r_b)^2 \end{aligned} \tag{4}$$

Where $F = E(y)$ represents the mean of the measured methane fluxes (y_k)

2.3 Uncertainties in estimating rice paddy methane emissions on ~~regional~~national scales: data error and availability

In addition to ~~errors in the performance of the method~~model fallacys, the difficulties in estimating ~~regional~~national rice paddy methane emissions also stem from errors in, and limited availability of, input data. To measure the uncertainties in model outputs due to insufficient data quality and availability, we applied Monte Carlo simulations (Penman, 2000) to the CH4MOD model. Statistical characteristics were derived from the available datasets to develop probability distribution functions (PDFs) for each model input variable (Table B1, B2 of Appendix B in Supporting Information). We performed Monte Carlo simulation in the way of randomly drawing values of the model input variables from their PDFs and then run the model. This process iterated 1000 times and at the last step, the mean and 95% CI of the calculated methane fluxes were derived from the iterations~~At the last step of the Monte Carlo simulation, the mean and the 95% CI of the methane flux were derived from the iterations of the CH4 MOD simulation (Zhang et al., 2014)~~.

The factors involved in the uncertainty analysis included organic matter application, soil properties and water regimes; these variables (*OM*, *SAND*, and W_{ptn}) were parameterized as input variables in the CH4MOD model (Huang et al., 2006;Zhang et al., 2011). The other two model input variables were the rice grain yield and daily ambient air temperature. These two variables were not used in the uncertainty analysis because sufficient relevant data were available, which were characterized by less error compared with the other variables (Zhang et al., 2014).

The *SAND* data were obtained from a 10 km × 10 km grid dataset interpolated from soil survey data (Oberthür et al., 1999;Shi et al., 2004;Liu et al., 2006). It is possible that approximately half (Van Bodegom et al., 2002b) of the immense spatial variation in soil properties can be lost after spatial interpolation (Goovaerts, 2001); as a result, the missing spatial variation was ~~interpolated from~~attributed to the PDF of the gridded *SAND* data (Appendix B in Supporting Information).

The organic matter inputs in the rice fields consisted of various types of farm manure (green manure and animal feces), crop straw, and dead roots and stubble leftover from previous harvests. Root and straw biomass were calculated using the root/shoot ratio and harvest indices (Huang et al., 2007;Gao et al., 2002;Xie et al., 2010c). Stubble was assumed to represent one-tenth of the straw biomass (Huang et al., 2004). The proportions of incorporated straw and applied farm manure were derived from data obtained from two large-scale investigations, the First National Census of Pollution Sources conducted by China's Ministry of Environmental Protection (CFPC, 2011) and census~~ing~~ conducted by the Institute of Atmospheric Physics, Chinese Academy of Sciences. The proportion of straw and the amount of manure incorporated into the crop fields were summarized by province~~ (Table B1 in the supporting information)~~. Table B1 shows the statistical parameters of the PDF of organic matter incorporation in each province.

The irrigation in rice cultivation were grouped into five general irrigation patterns: 1) flooding-drainage-flooding-intermittent irrigation, 2) flooding-drainage-intermittent irrigation, 3) flooding-intermittent irrigation, 4) continuous flooding and 5) continuously intermittent irrigation (Gao and Li, 1992; Huang et al., 2004). Data pertaining to W_{ptn} were only very rarely available on a regional scale. The limited information provided in a few studies (Mao, 1981;Liang, 1983;Xiong et al., 1992;Cai et al., 2003;Ma et al., 2005;Ministry of Water Resources and Utilization of China (MWRUC), 1996) ~~enabled us to produce~~ could only yield rough estimates related to irrigation in regions of major rice cultivation. The PDF of field irrigation were defined by the occurrence percentage of each irrigation pattern (Table B2). ~~-~~(Table B2 in the supporting information).

The data pertaining to the rice grain yield and harvesting area as of 2005 were obtained from China's Statistical Yearbook (EBCAY, 2006) and the nation's agricultural database maintained by the Chinese Academy of Agricultural Sciences, respectively. The spatial distributions of all rice paddies in 2005 and the rice paddy area within each 1 km × 1 km grid were obtained from the Data Center for Resources and Environmental Sciences of the Chinese Academy of Sciences (RESDC, CAS). Daily mean air temperature data from 678 meteorological stations throughout China for 2005 were acquired from the National Meteorological Information Center (NMIC) of the China Meteorological Administration (CMA) (http://cdc.cma.gov.cn/). The temperatures were then spatially interpolated into 10 km × 10 km grids for each day according to the method described by Thornton et al. (1997). Details on the datasets used in this study can be found in Appendix ~~C~~B.

To preserve details related to spatial variations, all data input into the model were converted into 10 km × 10 km grids. The applied rasterization techniques and details of how the model was run on raster datasets were provided in previously published papers (Huang et al., 2006).

2.4 Combining uncertainty and spatial aggregation

In each 10 km × 10 km grid, the uncertainties in our estimates originated from ~~errors in~~ both the model ~~performance~~ fallacy (Equation 4) and error in the input data. Equation 5 was used to merge the two ~~components~~ uncertainty sources ~~(equation 4 and those in the section D of the Supporting Information)~~where MSE was again split into two parts as showed in Equation 4:

$$\sigma_{T,i}{}^2 = \sigma_{b,i}^2 + \sigma_{v,i}^2 + \sigma_{d,i}^2$$
$$= (F_i \times r_b)^2 + (F_i \times r_v)^2 + \sigma_{d,i}{}^2 \qquad (5)$$

where $\sigma_{T,\cancel{j}\,i}$ represents the uncertainty of the methane flux in grid ~~j~~i, and ~~F_j~~ F_i and $\sigma_{\cancel{D}d,\cancel{j}\,i}$ represent the mean and standard deviation of the Monte Carlo simulation results in grid ~~j~~i, respectively. ~~B_R~~ r_b and ~~CV~~ r_v represent the same entities as in Equations 2 and 3. $\sigma_{d,i}{}^2$ signifies the uncertainty caused by the error and availability of data, $(F_i \times r_b)^2$ represents the modelling bias, and $(F_i \times r_v)^2$ represents the rest parts of the model fallacy apart from $(F_i \times r_b)^2$. To produce the uncertainty of the national inventory, the three components ~~of the estimation uncertainties~~ ($(F_i \times r_b)^2$, $(F_i \times r_v)^2$ and $\sigma_{d,i}{}^2$ in Equation 5) of the estimation uncertainties in all grids were separately aggregated (Equation D2, D3, D4 and -D5 in ~~the section~~Appendix D of the Supporting Information) and summed ~~(Equation 6)~~ as follows:

$$\sigma_T{}^2 = \sigma_b^2 + \sigma_v^2 + \sigma_d^2 \qquad (6)$$

3 Results

3.1 Methane emissions and the uncertainties derived from field measurements

Among the 495 methane flux measurements (the accumulative methane emission from transplanting to harvesting), 184 (37% of all cases) came from paddies that were continuously flooded during the entire rice growing period; 50 (10% of all cases) came from paddies with single mid-season drainage; and 261 (53% of all cases) came from paddies under multi-drainage. The average methane fluxes associated with the three water regimes were 531.6 ± 512.6, 251.6 ± 231.1 and 224.1 ± 207.5 kg CH_4 ha^{-1}, respectively (Fig. 3a). The overall arithmetic average of the 495 measurements (represented hereafter by m_c) was 341.2 ± 383.2 kg CH_4 ha^{-1}. However, the simple arithmetic average might be a biased representation of the "true" mean methane flux of rice paddies in China, as far less than 37% of the rice paddies in China are continuously flooded. In the literature, 10%, 20% and 70% of the rice area was reported to be under continuous flooding, single drainage and multi-drainage water regimes, respectively (Xiong et al., 1992;Ministry of Water Resources and Utilization of China (MWRUC), 1996), and the harvested-area-weighted mean (Appendix A in the Supporting Information) of the measured fluxes (represented hereafter by m_w) was 260.4 ± 281.6 kg CH_4 ha^{-1} (Fig. 3a).

The 95% confidence intervals (CIs) of the methane flux measurements were 61.1–2,145.9 kg CH_4 ha^{-1}, 9.6–809.9 kg CH_4 ha^{-1} and 14.0–797.7 kg CH_4 ha^{-1}, respectively, for the three water regimes (continuous flooding, single drainage, and multi-drainage in Fig. 3a). The 95% CI of all combined area-weighted measurements (Appendix A in the Supporting Information) was 13.7–1,115.4 kg CH_4 ha^{-1}. The measurements were not normally or symmetrically distributed (Fig. 3b). The P-P plots (Fig. 4) showed that the parameterized gamma distributions matched the sample distributions. The 95% CIs calculated with the parameterized gamma functions were 16.8–1,900.8 kg CH_4 ha^{-1}, 10.4–863.4 kg CH_4 ha^{-1} and 8.9–774.2 kg CH_4 ha^{-1}, respectively, for the three water regimes; these values overlapped the CIs derived directly from the measurements by 88.2%, 99.9% and 97.0%, respectively.

The national methane emissions from rice agriculture calculated by multiplying the rice harvesting area (yearbook data in 2005) by the area-weighted mean flux (260.4 ± 281.6 kg CH_4 ha^{-1}) were 7.51 Tg CH_4 (Fig. 3a). ~~The uncertainty of the national emissions is usually represented by the standard error (SE).~~ When the measurements are statistically independent, the standard error (SE) of the summation is $n-1$ (n is the sample size of the measurements) times smaller than the standard deviation (± 281.6 kg CH_4 ha^{-1}), which consists of the representative and measurement errors of the measured fluxes (Van Bodegom et al., 2002a;Verburg et al., 2006). Assuming that the measurements were statistically independent, the 95% CI of the national inventory was 7.20–8.58 Tg CH_4 (Equation A3 in the Supporting Information). However, the independency assumption is questionable because of the spatial correlations between the spatially correlated background environmental conditions and agricultural activities (Legendre, 1993;Dormann et al., 2007). The equivalent sample size~~, n,~~ used to calculate SE may be smaller than 495, and the 95% CI of the national inventory is therefore larger than that with the independency assumption.

3.2 Model performance under different situations of data availability

~~R1 was more likely to overestimate the amount of methane emitted than R2 (Table 2), especially when more organic matter was incorporated (Fig. 5a). The estimated methane flux calculated by R1 was more than 6,000 kg CH_4 ha^{-1}, whereas the corresponding measured methane flux was less than 3,000 kg CH_4 ha^{-1} (Fig. 5a).~~ The averaged bias of the estimate obtained with R1 was 212.0 kg CH_4 ha^{-1} (Table 2) or 62.1% of the measured mean (m_c = 341.2 kg CH_4 ha^{-1}). The average bias of R2, in contrast, was −1.3 kg CH_4 ha^{-1}~~;~~ . R1 was more likely to overestimate the amount of methane emitted than R2 (Table 2), especially when more organic matter was incorporated (Fig. 5a). For example, in one case the modeled CH_4 flux was more than 6,000 kg CH_4 ha^{-1}, whereas the measured flux was less than 3,000 kg CH_4 ha^{-1} (Fig. 5a). T~~t~~he estimates obtained using R2 did not show significant variations and appeared to decline when the measured methane fluxes increased (Fig. 5b). The CH4MOD model also produced a small averaged bias, representing 7.1% of the measured mean. The ~~total estimation errors~~MSE ~~were~~ was 253.0, 407.8 and 596.0 kg CH_4 ha^{-1} for the M—S0, R2—S0 and R1—S0 scenarios, respectively (Table 2), which demonstrates that model performance improves when more factors are incorporated into the model. Although the CH4MOD model produced better simulation results than the simple regression equations, its performance fundamentally depends on data availability. When no case-specific data were available (as in scenario M—S3), ~~B_R~~ r_b was −32.2%, and ~~ERR_T~~ MSE was 122.1% of the mean flux; the results obtained under this scenario were even worse than the results obtained under the R2—S0 scenario (Table 2). For the M—S1 scenario, where the data pertaining to the soil properties and water regime were case-specific, the magnitude of ~~B_R~~ r_b decreased to 9.0% of the mean flux, and the ~~total error~~MSE decreased to 101.2% of the mean flux. The M—S0 scenario produced much better results than the other scenarios, as more data were available for the key model input variables (Table 2). Even no case-specific input data used in M—S3 had smaller r_b, r_v and MSE than R1—S0. In Table 2, larger r_v of R1—S0 than M—S3 might come from the too simple explanation of the influence from organic matter inputs on methane emission that added extra error on the estimation.

3.3 Inventory of rice paddy methane emissions and ~~spatial variations~~the uncertainties with different approaches

Because of the spatial heterogeneity in the climate, soil properties, organic matter incorporation and field irrigation in rice cultivation, the ~~simulated~~ methane fluxes simulated by CH4MOD varied spatially between 17.2 kg CH_4 ha^{-1} and 708.3 kg CH_4 ha^{-1} from grid to grid (Fig. 6). The national means for the simulated methane fluxes were 217.9 kg CH_4 ha^{-1}, 204.6 kg CH_4 ha^{-1} and 255.8 kg CH_4 ha^{-1} for single, early and late rice cultivation, respectively. The within-grid estimation error ($\sigma_{T,i}$, calculated with Equation 5) represented 81.2%–95.5% of the mean fluxes~~,~~ ~~F_j~~F_i~~,~~ in the grids. In the present study, model fallacy, represented by $U_{b,i}$+$U_{v,i}$, contributed 79.5%–88.9% to the uncertainty $\sigma_{T,i}^2$, with $\sigma_{d,i}^2$ accounting for the remaining 11.1%–20.5%. This implies that a model with better performance is needed to reduce the uncertainty of $\sigma_{T,i}$ in each grid.

As shown in Fig. 7, the highest levels of emitted methane occurred in southern China, with the northeast also representing a major source of methane, despite this region being climatically cool. The total amount of methane emitted, as calculated by

the M-S0~~modeling~~ approach, was 6.43 (3.79–9.77) Tg CH_4 yr^{-1} (Table 3), which is close to the 7.51 Tg CH_4 yr^{-1} derived from the experimental field measurements.

In Table 3, the estimated national CH_4 emissions ranged from 6.43 (3.79–9.77) Tg CH_4 yr^{-1} to 13.59 (1.45–38.98) Tg CH_4 yr^{-1} for the M–S0 scenario and R1–S0 scenario, respectively. The 95% CIs of the national estimation differed more greatly among the approaches than those among the data availability scenarios of each approach. With M–S0, The fallacy of CH4MOD contributed 56.6% of the total uncertainty, with the remaining 43.4% being attributed to errors and the scarcity of the spatial datasets of the model inputs (Table 4). As an indicator of the trade-off between the complexity of the approach and data availability, the σ_d/σ_{b+v} ratio in Table 3 was 0.87 for M–S0, closer to 1 than those for the other approaches and scenarios, which also yielded the narrowest 95% CI in Table 3.

4 Discussion

4.1 Contributions of different error sources to the uncertainties in the inventory

Methane fluxes in rice paddies varied extensively with environmental and agronomic factors. Certain factors, such as rice biomass (Bachelet and Neue, 1993), organic matter input (Kern et al., 1995), water management (Khosa et al., 2011;Mishra et al., 1997), paddy soil properties (Yao et al., 1999;Gaunt et al., 1997), climate (Sass et al., 1991) and rice varieties (Ding et al., 1999), have previously been recognized as having significant impacts on methane emissions. Other factors, such as atmospheric CO_2 and ozone contents (Bhatia et al., 2011;Inubushi et al., 2011), N fertilizer application (Banger et al., 2012;Xie et al., 2010a) and active soil organic C (Zhan et al., 2011), are also receiving increasing attention. Because so many factors affect the production, oxidation and emission of methane from rice cultivation, the observed methane fluxes varied extensively both spatially and temporally. In the experimental field measurements (Fig. 1), the variations in rice paddy methane fluxes ranged from 3.2 kg CH_4 ha^{-1} to 2,451.7 kg CH_4 ha^{-1}, averaging 341.2 ± 383.2 kg CH_4 ha^{-1}. The average simulated methane fluxes in the 10 x 10 km grids varied from 17.2 to 708.3 kg CH_4 ha^{-1} (Fig. 6). The extremely high methane fluxes obtained from experimental measurements were not reproduced by the model estimations. This ~~discrepancy~~ was partly due to the discrepancy ~~variations~~ in the spatial representativeness of the methane fluxes in field observations and model estimations (Verburg et al., 2006). The experimental measurements represented methane fluxes from an area of less than one hectare, while the modeled fluxes were the averages from 10 x 10 km grids. This mismatch in spatial representativeness might also be due to errors in the model input data as well as to the impacts of other unknown factors (Singh and Dubey, 2012;Bhatia et al., 2011;Zheng et al., 2010;Gauci et al., 2008). Methane emissions could be estimated using a limited number of factors and simplified equations to express the complex relationships between methane emissions and influential factors, but such simplification resulted in poor performance of the methods (Table 2). In Equation 5, $\sigma_{\cancel{D}d,\cancel{j}i}$ is the uncertainty due to errors in the input data. With an increasing number of explanatory factors, $\cancel{B_R}$ r_b and $\cancel{CV}$ r_v might decrease (which means better performance of the method), but $\sigma_{\cancel{D}d,i}$ might increase because of the cumulative errors resulting from the increasing number of factors incorporated in the models. To reduce uncertainties in the estimates and improve the performance of the model, the input data need to be available and of good quality. ~~In the present study, model fallacy, represented by the equation $(F_j \times B_R)^2 + (F_j \times CV)^2$, contributed 79.5%–88.9% to the uncertainty of $\sigma_{T,j}^2$, with $\sigma_{D,j}^2$ accounting for the remaining 11.1%–20.5%. This implies that a model with better performance is needed to reduce the uncertainty of $\sigma_{T,j}$ in each grid.~~

The aggregated uncertainty of the national inventory depended not only on the magnitude of $\sigma_{\cancel{CV}V,\cancel{j}i}$ and $\sigma_{\cancel{D}d,\cancel{j}i}$ in each grid ($\cancel{j}i$) but also on the spatial correlation between these variables (Equation C2 in the Supporting Information). The spatial correlation of $\sigma_{\cancel{D}d,\cancel{j},i}$ depends on the availability of input data for the model and on spatial aggregation (Table C1 in the Supporting Information). However, the spatial correlation of $\sigma_{\cancel{CV}V,i\cancel{j}}$ could not be assessed analytically because it was a result of model ~~imprecision~~fallacy and~~,~~ ~~random noise~~errors in measurements~~ and/or unknown factors~~. In the case of a strong correlation of $\sigma_{\cancel{CV}V,\cancel{j}i}$ values, the aggregated $\sigma_{\cancel{CV}V}^2$ will account for a large proportion of σ_T^2 (right side in Fig. 8). However, if the spatial correlation is confined to a short distance, such as less than four grids (Dormann et al., 2007;Dray et al., 2006), the contribution of $\sigma_{\cancel{CV}V}^2$ to σ_T^2 will be negligible (left side in Fig. 8). At the mid-point of D_C (Equation C2, 30 grids, equal to 300 kilometers), as shown in Fig. 8, the model uncertainty ($\sigma_{\cancel{BR}r}^2$ + $\sigma_{\cancel{CV}V}^2$) accounted for 56.6% of the uncertainty in σ_T^2 (Table 4).

4.2 Consistency of errors between model validation and model up-scaling

Up-scaling a site-scale model (e.g., CH4MOD in this study) to a ~~regional~~national scale poses enormous challenges when data are scarce. Enhancing the spatial abundance of the input data minimizes the propagation of data error into the aggregated uncertainties. Many environmental and agricultural factors impact methane emissions from rice paddies. In the CH4MOD model, the key factors were parameterized as model inputs (Huang et al., 2004). However, when assessing the uncertainty of a model, the explanatory variables are arbitrarily included (Verburg et al., 2006). Li et al. (2004) found that soil properties were the "most sensitive factor" and therefore used this parameter in the uncertainty analysis. The inclusion of as many of the highly sensitive key factors as possible in the uncertainty analysis should generate more accurate and reliable results (right part in Fig. 2).

Experimental field studies have shown that the rice variety has substantial impacts on methane emissions (Aulakh et al., 2008;Inubushi et al., 2011;Jia et al., 2002). A study of field observations (Su et al., 2015) showed that transfer of the barley gene *SUSIBA2* to rice favors the allocation of photosynthates to the aboveground biomass over allocation to the roots and, moreover, that less biomass allocation to root exudates results in reduced methane emissions. The impact that the rice variety has on methane emissions was parameterized as the variety index (*VI*) in CH4MOD. According to Huang et al. (1998), *VI* ranges from 0.5 to 1.5 and averages 1.0 for most rice varieties. To validate the CH4MOD model (left portion of Fig. 2) using the 495 methane emission measurements included in the present study, *VI* was assigned a default value of 1.0 regardless of the rice variety because, until now, no dedicated attempts have been made to quantify the *VI* of different rice varieties. Therefore, the ~~*BR*~~ r_b and ~~*CV*~~ r_v values presented in Table 2 incorporate the uncertainty in model performance that can be attributed to different rice varieties ($M_f(x)$ in Equation C2 of the Supporting Information). To maintain consistency, *VI* was assigned the same default value (1.0) when the model was scaled-up to the national scale (right side of Fig. 2), and no PDF was built for the uncertainty calculation conducted with the Monte Carlo simulation. If a PDF had been incorporated into the uncertainty calculation when the model was scaled-up, the overall uncertainties (Table 4) would have been overestimated. However, if different *VI* values were assigned to rice varieties during model validation, the error caused by the inaccuracy of *VI* would also need to be considered during the scaling-up of the model to prevent underestimation of the overall uncertainty.

5 Conclusion

Due to the remarkable spatial variation in rice paddy methane emissions, the uncertainties in ~~regional~~national estimates obtained either through field measurements or modeling remain considerably large. For field measurements, the reduction in uncertainty achieved by increasing the number of observations was shown to be inversely related to the spatial correlation between the measurements. To reduce the estimation bias, the number of measured emission fluxes should be proportional to the paddy area where the corresponding agronomic activities and environmental conditions occur homogenously.

Model performance depends not only on the effectiveness of the models themselves but also on the availability of the data needed to drive the model. We found that without a sufficient quantity of high-quality data, a well-developed model ~~will~~ may perform even more poorly than simple regression approaches. When modeling methane emissions, uncertainties in the performance of the model remain the major obstacle to reliably estimating methane emissions. Estimate uncertainty could be reduced at the ~~regional~~national scale by increasing the availability of input data and decreasing spatial correlations among the residues of the model output.

Modelling by CH4MOD with all the available data, the national methane emission from rice paddies was 6.43 (3.79–9.77) Tg CH_4 yr^{-1} in China. Comparing to other options, balancing between the uncertainties caused by the model fallacy and data scarcity produced national estimations of least total uncertainty.

References

Aulakh, M., Wassmann, R., and Rennenberg, H.: Pattern and amount of aerenchyma relate to variable methane transport capacity of different rice cultivars, Plant Biol., 2, 182-194, 2008.

Bachelet, D., and Neue, H.: Methane emissions from wetland rice areas of Asia, Chemosphere, 26, 219-237, 1993.

Bachelet, D., Kern, J., and Toelg, M.: Balancing the rice carbon budget in China using spatially-distributed data, Ecol. Model., 79, 167-177, 1995.

Banger, K., Tian, H., and Lu, C.: Do nitrogen fertilizers stimulate or inhibit methane emissions from rice fields?, Glob. Change Biol., 18, 3259-3267, 2012.

Bhatia, A., Ghosh, A., Kumar, V., Tomer, R., Singh, S., and Pathak, H.: Effect of elevated tropospheric ozone on methane and nitrous oxide emission from rice soil in north India, Agr. Ecosyst. Environ., 144, 21-28, 2011.

Bodelier, P. L. E., and Laanbroek, H. J.: Nitrogen as a regulatory factor of methane oxidation in soils and sediments, FEMS Microbiol. Ecol., 47, 265-277, 2006.

Butenhoff, C., Frolking, S., Li, C., Houweling, S., Milliman, T., Khalil, A., and Zhuang, Q.: Intercomparison of models to estimate methane emissions from rice agriculture using common data sets, AGU Fall Meeting, 2009, 283,

Butterbach-Bahl, K., Papen, H., and Rennenberg, H.: Impact of gas transport through rice cultivars on methane emission from rice paddy fields, Plant Cell Environ., 20, 1175-1183, 1997.

Cai, Z., Tsuruta, H., Gao, M., Xu, H., and Wei, C.: Options for mitigating methane emission from a permanently flooded rice field, Glob. Change Biol., 9, 37-45, 2003.

Cao, M., Dent, J., and Heal, O.: Modeling methane emissions from rice paddies, Global Biogeochem. Cy., 9, 183-195, 1995.

Cao, M., Gregson, K., Marshall, S., Dent, J., and Heal, O.: Global methane emissions from rice paddies, Chemosphere, 33, 879-897, 1996.

CFPC: Datasets of China Pollution Source Census, China Environmental Sciences Press, Beijing, China, 2011.

Chen, H., Zhu, Q. a., Peng, C., Wu, N., Wang, Y., Fang, X., Jiang, H., Xiang, W., Chang, J., and Deng, X.: Methane emissions from rice paddies natural wetlands, and lakes in China: synthesis and new estimate, Glob. Change Biol., 19, 19-32, 2013.

Conrad, R., Klose, M., Noll, M., Kemnitz, D., and BODELIER, P. L. E.: Soil type links microbial colonization of rice roots to methane emission, Glob. Change Biol., 14, 657-669, 2007.

Dijkstra, F. A., Prior, S. A., Runion,G. B., Torbert, H. A., Tian, H., Lu, C., and Venterea R. T.,: Effects of elevated carbon dioxide and increased temperature on methane and nitrous oxide fluxes: Evidence from field experiments, Front. Ecol. Environ., 10(10), 520–527, 2012

Ding, A., Willis, C., Sass, R., and Fisher, F.: Methane emissions from rice fields: effect of plant height among several rice cultivars, Global Biogeochem. Cy., 13, 1045-1052, 1999.

Dormann, F. C., M McPherson, J., B Araújo, M., Bivand, R., Bolliger, J., Carl, G., G Davies, R., Hirzel, A., Jetz, W., and Daniel Kissling, W.: Methods to account for spatial autocorrelation in the analysis of species distributional data: a review, Ecography, 30, 609-628, 2007.

Dray, S., Legendre, P., and Peres-Neto, P. R.: Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM), Ecol. Model., 196, 483-493, 2006.

EBCAY: China Agriculture Yearbook, China Agriculture Press, Beijing, China, 2006.

Gao, X. Z., Ma, W. Q., Ma, C. B., Zhang, F. S., and Wang, Y. H.: Analysis on the current status of utilization of crop straw in China (in Chinese with English abstract), J. Huazhong Agr. Univ., 21, 242-247, 2002.

Gauci, V., Dise, N. B., Howell, G., and Jenkins, M. E.: Suppression of rice methane emission by sulfate deposition in simulated acid rain, J. Geophys. Res., 113, G00A07, 2008.

Gaunt, J., Grant, I., Neue, H., Bragais, J., and Giller, K.: Soil characteristics that regulate soil reduction and methane production in wetland rice soils, Soil Sci. Soc. Am. J., 61, 1526-1531, 1997.

Goovaerts, P.: Geostatistical modelling of uncertainty in soil science, Geoderma, 103, 3-26, 2001.

Huang, Y., Sass, R. L., and Fisher Jr, F. M.: A semi‐empirical model of methane emission from flooded rice paddy soils, Glob. Change Biol., 4, 247-268, 1998.

Huang, Y., Zhang, W., Zheng, X., Li, J., and Yu, Y.: Modeling methane emission from rice paddies with various agricultural practices, J. Geophys. Res., 109, D08113, 2004.

Huang, Y., Zhang, W., Zheng, X., Han, S., and Yu, Y.: Estimates of methane emissions from Chinese rice paddies by linking a model to GIS database, Acta Ecologica Sinica, 26, 980-987, 2006.

Huang, Y., Zhang, W., Sun, W., and Zheng, X.: Net primary production of Chinese croplands from 1950 to 1999, Ecol. Appl., 17, 692-701, 2007.

Inubushi, K., Cheng, W., Mizuno, T., Lou, Y., Hasegawa, T., Sakai, H., and Kobayashi, K.: Microbial biomass carbon and methane oxidation influenced by rice cultivars and elevated CO2 in a Japanese paddy soil, Eur. J. Soil Sci., 62, 69-73, 2011.

Ito, A., and Inatomi, M.: Use of a process-based model for assessing the methane budgets of global terrestrial ecosystems and evaluation of uncertainty, Biogeosciences, 9(2), 759–773, 2012

Jia, Z., Cai, Z., Xu, H., and Tsuruta, H.: Effects of rice cultivars on methane fluxes in a paddy soil, Nutr. Cycl. Agroecosys., 64, 87-94, 2002.

Kennedy, M.C., and O'Hagan, A.: Bayesian calibration of computer models. J. R. Stat. Soc. Ser. B Stat. Methodol. 63, 425–464, 2001

Kern, J., Bachelet, D., and Tölg, M.: Organic matter inputs and methane emissions from soils in major rice growing regions of China, Soils and Global Change, 189-198, 1995.

Kern, J. S., Zitong, G., Ganlin, Z., Huizhen, Z., and Guobao, L.: Spatial analysis of methane emissions from paddy soils in China and the potential for emissions reduction, Nutr. Cycl. Agroecosys., 49, 181-195, 1997.

Khalil, M., Rasmussen, R., Wang, M. X., and Ren, L.: Methane emissions from rice fields in China, Environ. Sci. Technol., 25, 979-981, 1991.

Khalil, M., Shearer, M., and Rasmussen, R.: Methane sources in China: historical and current emissions, Chemosphere, 26, 127-142, 1993.

Khosa, M. K., Sidhu, B., and Benbi, D.: Methane emission from rice fields in relation to management of irrigation water, J. Environ. Biol., 32, 169-172, 2011.

Legendre, P.: Spatial autocorrelation: trouble or new paradigm?, Ecology, 74, 1659-1673, 1993.

Li, C.: Modeling trace gas emissions from agricultural ecosystems, Nutr. Cycl. Agroecosys., 58, 259-276, 2000.

Li, C., Qiu, J., Frolking, S., Xiao, X., Salas, W., Moore III, B., Boles, S., Huang, Y., and Sass, R.: Reduced methane emissions from large-scale changes in water management of China's rice paddies during 1980–2000, Geophys. Res. Lett., 29, 1972, 2002.

Li, C., Mosier, A., Wassmann, R., Cai, Z., Zheng, X., Huang, Y., Tsuruta, H., Boonjawat, J., and Lantin, R.: Modeling greenhouse gas emissions from rice-based production systems: sensitivity and upscaling, Global Biogeochem. Cy., 18, GB1043, 2004.

Liang, G.: Rice Ecology (in Chinese), China Agricultural Press, Beijing, 1983.

Liu, T. L., Juang, K. W., and Lee, D. Y.: Interpolating soil properties using kriging combined with categorical information of soil maps, Soil Sci. Soc. Am. J., 70, 1200-1209, 2006.

Ma, X., Zhu, B., Du, D., and ZHENG, X.: CH_4, CO_2 and N_2O emissions from the year-round flooded paddy field at fallow season (in Chinese with English abstract), Journal of Agro-Environment Science, 24, 1199-1202, 2005.

Mao, L.: Irrigation in Rice Paddies (in Chinese), China Agricultural Press, Beijing, 1981.

Matthews, E., Fung, I., and Lerner, J.: Methane emission from rice cultivation: geographic and seasonal distribution of cultivated areas and emissions, Global Biogeochem. Cy., 5, 3-24, 1991.

Matthews, R., Wassmann, R., and Arah, J.: Using a crop/soil simulation model and GIS techniques to assess methane emissions from rice fields in Asia. I. Model development, Methane Emissions from Major Rice Ecosystems in Asia, 141-159, 2001.

Ministry of Water Resources and Utilization of China (MWRUC): National Program for Conservation of Irrigation Water in Chinese Agriculture (in Chinese), edited by: Division of Rural Water Resources and Utilization, China Agriculture Press, Biejing, China, 1996.

Mishra, S., Rath, A., Adhya, T., Rao, V., and Sethunathan, N.: Effect of continuous and alternate water regimes on methane efflux from rice under greenhouse conditions, Biol. Fert. Soils, 24, 399-405, 1997.

Neue, H., Becker-Heidmann, P., and Scharpenseel, H.: Organic matter dynamics, soil properties, and cultural practices in rice lands and their relationship to methane production, Soils and the Greenhouse Effect, 457-466, 1990.

Oberthür, T., Goovaerts, P., and Dobermann, A.: Mapping soil texture classes using field textuing, particle size distribution and local knowledge by both conventional and geostatisical methods, Eur. J. Soil Sci., 50, 457-479, 1999.

Ogle, S. M., Breidt, F., Easter, M., Williams, S., Killian, K., and Paustian, K.: Scale and uncertainty in modeled soil organic carbon stock changes for US croplands using a process - based model, Glob. Change Biol., 16, 810-822, 2010.

Penman, J.: Good practice guidance and uncertainty management in national greenhouse gas inventories, Institute for Global Environmental Strategies (IGES) for the IPCC, 2000.

Ren, W., Tian, H., Xu, X., Liu, M., Lu, C., Chen, G., Melillo, J., Reilly, J., and Liu, J.: Spatial and temporal patterns of CO2 and CH4 fluxes in China's croplands in response to multifactor environmental changes, Tellus B, 63, 222-240, 2011.

Sanchis, E., Ferrer, M., Torres, A. G., Cambra-López, M., and Calvet, S.: Effect of water and straw management practices on methane emissions from rice fields: a review through a meta-analysis, Environ. Eng. Sci., 29, 1053-1062, 2012.

Sass, R., Fisher, F., Turner, F., and Jund, M.: Methane emission from rice fields as influenced by solar radiation, temperature, and straw incorporation, Global Biogeochem. Cy., 5, 335-350, 1991.

Sass, R., Fisher, F., Wang, Y., Turner, F., and Jund, M.: Methane emission from rice fields: the effect of floodwater management, Global Biogeochem. Cy., 6, 249-249, 1992.

Sass, R., Fisher, F., Lewis, S., Jund, M., and Turner, F.: Methane emissions from rice fields: effect of soil properties, Global Biogeochem. Cy., 8, 135-140, 1994.

Sass, R., Fisher Jr, F., Ding, A., and Huang, Y.: Exchange of methane from rice fields: national, regional, and global budgets, J. Geophys. Res., 104, 26, 1999.

Sass, R. L., and Fisher, F. M.: Methane emissions from rice paddies: a process study summary, Nutr. Cycl. Agroecosys., 49, 119-127, 1997.

Shi, X., Yu, D., Warner, E., Pan, X., Petersen, G., Gong, Z., and Weindorf, D.: Soil database of 1: 1,000,000 digital soil survey and reference system of the Chinese genetic soil classification system, Soil Survey Horizons, 45, 129-136, 2004.

Singh, A., and Dubey, S. K.: Temporal variation in methanogenic community structure and methane production potential of tropical rice ecosystem, Soil Biol. Biochem., 48, 162-166, 2012.

Smith, P., Martino, D., Cai, Z., Gwary, D., Janzen, H., Kumar, P., McCarl, B., Ogle, S., O'Mara, F., and Rice, C.: Greenhouse gas mitigation in agriculture, Philos. T. Roy. Soc. B, 363, 789-813, 2008.

Su, J., Hu, C., Yan, X., Jin, Y., Chen, Z., Guan, Q., Wang, Y., Zhong, D., Jansson, C., and Wang, F.: Expression of barley SUSIBA2 transcription factor yields high-starch low-methane rice, Nature, 2015.

Taylor, J. A., Brasseur, G., Zimmerman, P., and Cicerone, R.: A study of the sources and sinks of methane and methyl chloroform using a global three-dimensional Lagrangian tropospheric tracer transport model, J. Geophys. Res., 96, 3013-3044, 1991.

Thornton, P. E., Running, S. W., and White, M. A.: Generating surfaces of daily meteorological variables over large regions of complex terrain, J. Hydrol., 190, 214-251, 1997.

Tian, H. Q., Lu, C.Q., Ciais, P., et al. : The terrestrial biosphere as a net source of greenhouse gases to the atmosphere, Nature, 531(7593), 225–228, 2016

Van Bodegom, P., Wassmann, R., and Metra-Corton, T.: A process-based model for methane emission predictions from flooded rice paddies, Global Biogeochem. Cy., 15, 247-263, 2001.

Van Bodegom, P., Verburg, P. H., and van der Gon, H. A. C. D.: Upscaling methane emissions from rice paddies: problems and possibilities, Global Biogeochem. Cy., 16, 1014, 2002a.

Van Bodegom, P. M., Verburg, P. H., Stein, A., Adiningsih, S., and Denier Van Der Gon, H. A. C.: Effects of interpolation and data resolution on methane emission estimates from rice paddies, Environ. Ecol. Stat., 9, 5-26, 2002b.

Verburg, P. H., van Bodegom, P. M., van der Gon, H. A. C. D., Bergsma, A., and van Breemen, N.: Upscaling regional emissions of greenhouse gases from rice cultivation: methods and sources of uncertainty, Plant Ecol., 182, 89-106, 2006.

Watanabe, A., Kajiwara, M., Tashiro, T., and Kimura, M.: Influence of rice cultivar on methane emission from paddy fields, Plant Soil, 176, 51-56, 1995.

Wei, H. P.: Statistical analysis of methane emissions from Chinese rice paddies from 1987 to 2010, Master, College of Resources and Environmental Sciences, Nanjing Agricultural University, Nanjing, China, 65 pp., 2012.

Weller, S., Janz, B., Jörg,L., Kraus, D., Racela, H. S., Wassmann, R., Butterbach-Bahl, K., and Kiese R.: Greenhouse gas emissions and global warming potential of traditional and diversified tropical rice rotation systems, Global Change Biol., 22(1), 432–448,2016

Xie, B., Zheng, X., Zhou, Z., Gu, J., Zhu, B., Chen, X., Shi, Y., Wang, Y., Zhao, Z., and Liu, C.: Effects of nitrogen fertilizer on CH 4 emission from rice fields: multi-site field observations, Plant Soil, 326, 393-401, 2010a.

Xie, B. H., Zhou, Z. X., Zheng, X. H., Zhang, W., and Zhu, J. G.: Modeling methane emissions from paddy rice fields under elevated atmospheric carbon dioxide conditions, Adv. Atmos. Sci., 27, 15, 2010b.

Xie, G. H., Wang, X. Y., and Ren, L. T.: China's crop residues resources evaluation (in Chinese with English abstract), Chinese J. Biotechn., 26, 855-863, 2010c.

Xiong, Z., Cai, H., Min, S., and Li, B.: Rice in China (in Chinese), China Agricultural Science and Technology Press, Beijing, 1992.

Yagi, K., Tsuruta, H., and Minami, K.: Possible options for mitigating methane emission from rice cultivation, Nutr. Cycl. Agroecosys., 49, 213-220, 1997.

Yan, X., Yagi, K., Akiyama, H., and Akimoto, H.: Statistical analysis of the major variables controlling methane emission from rice fields, Glob. Change Biol., 11, 11, 2005.

Yao, H., Conrad, R., Wassmann, R., and Neue, H.: Effect of soil characteristics on sequential reduction and methane production in sixteen rice paddy soils from China, the Philippines, and Italy, Biogeochemistry, 47, 269-295, 1999.

Zhan, M., Cao, C., Wang, J., Jiang, Y., Cai, M., Yue, L., and Shahrear, A.: Dynamics of methane emission, active soil organic carbon and their relationships in wetland integrated rice-duck systems in Southern China, Nutr. Cycl. Agroecosys., 89, 1-13, 2011.

Zhang, B. W., Tian, H. Q., Ren, W., Tao, B., Lu, C.Q., Yang, J., Banger, K., Pan, S.F.: Methane emission from global rice fields: Magnitude, Spatiotemporal patterns, and environmental controls, Glob., Biogeochem. Cycles, dio: 10.1002/2016GB005381, 2016

Zhang, W., Yu, Y. Q., Huang, Y., Li, T. T., and Wang, P.: Modeling methane emissions from irrigated rice cultivation in China from 1960 to 2050, Glob. Change Biol., 17, 3511-3523, DOI 10.1111/j.1365-2486.2011.02495.x, 2011.

Zhang, W., Zhang, Q., Huang, Y., Li, T. T., Bian, J. Y., and Han, P. F.: Uncertainties in estimating regional methane emissions from rice paddies due to data scarcity in the modeling approach, Geosci. Model Dev., 7, 1211-1224, 10.5194/gmd-7-1211-2014, 2014.

Zheng, F., Wang, X., Lu, F., Hou, P., Zhang, W., Duan, X., Zhou, X., Ai, Y., Zheng, H., and Ouyang, Z.: Effects of elevated ozone concentration on methane emission from a rice paddy in Yangtze River Delta, China, Glob. Change Biol., 17, 898-910, 2010.

Figure Legends

Figure 1 Locations of the experimental sites (red stars). The background map represents the spatial distribution of rice paddies in China. The size of the red stars is proportional to the number of measured methane fluxes at the site. The polygons show zones of different crop rotation systems involving rice: I—Double rice rotation; II—Mixed zone of rice/rice rotation and rice/upland crop rotation; III & IV—Rice/upland crop rotation or rice/fallow rotation; V & VI—Rice/fallow rotation; and VII—No rice.

Figure 2 Flowchart for estimating regional/national methane emissions and the uncertainties associated with field measurements and modeling

Figure 3 Statistical representations of the measured methane fluxes. (a) Statistical parameters and (b) histogram of the measurements. The solid circles represent the sample mean, and the vertical bars are the 95% confidence intervals of the samples, from the 2.5% percentile to the 97.5% percentile. The dashed line indicates the arithmetic average of all measured fluxes (m_c). The solid line is the area-weighted mean of the methane fluxes (m_w), in reference to the areal proportion of each water regime in the national total rice harvesting area: 10% continuous flooding (Flooded), 20% single drainage (Single-D) and 70% multi-drainage (Multi-D) (Xiong et al., 1992; MWRUC, 1996;Li et al., 2001;Zou et al., 2009).

Figure 4 P-P plots of the cumulative probability of the measured methane fluxes versus the gamma distribution. (a) Single drainage irrigation cases, (b) multi-drainage irrigation cases, (c) continuous flooding irrigation cases, and (d) all cases after being area weighted (Appendix A). *n*, *avg.* and *std.* represent the sample size, statistical mean and standard deviation of the sample methane fluxes, respectively. α and β represent the shape and scale parameters of the gamma distribution, which were calculated with the statistical mean and variance of the measured methane fluxes; $\beta = (std.)^2/(avg.)$, and $\alpha = (avg.)/\beta$. The diagonal line is the 1:1 straight line for a perfect gamma distribution match.

Figure 5 Methane fluxes in the experiments plotted against the respective simulation results through different methods. (a) R1—S0, (b) R2—S0 and (c) M—S0, which are described in Table 1.

Figure 6 Histograms and their fitting gamma probability lines for the calculated methane fluxes (via CH4MOD) of the 10 km × 10 km rice paddy grids in China. (a) Single rice rotations, including rice-fallow rotations, and rotations of rice with upland crops; (b) early and (c) late rice in double rice rotations. The vertical bars are the histograms of the calculated F_j (Equation 5), and the solid line is the theoretic gamma PDF line, the parameters for which were derived from the statistics for F_j via momentum methods.

Figure 7 Spatial distributions of rice paddy methane emissions ($\times 10^6$ g CH_4 per 10 km × 10 km grid).

Figure 8 Composition of the aggregated uncertainty of the national inventory along with the spatial autocorrelation of the variances of the model residues in grids. Distance criteria (D_c) are used to define the step functions of spatial autocorrelation: if two grids diverge by a distance beyond D_c, the autocorrelation is 0; otherwise, it is 1. The step function is a simplified version and represents the upper limit of the true spatial autocorrelation. With the step function, a larger D_c indicates stronger autocorrelation.

Tables

Table 1 Methods and their input scenarios

Methods	Input scenario	Reference
R1: $C_{CH4} = 0.3\times C_{input}$	R1 - S0: Case-specific C input, adjusted with the water regime[†].	Neue et al., (1990)
R2: $CH_4 = -0.006\times C_{input} + 0.078\times N_{input} + 0.885\times R_{C/N} + 21.15$	R2 - S0: Case-specific C and N input.	Kern et al., (1997)
	R2 - S1: Case-specific C input, averaged N input in all cases.	
M: CH4MOD model	M - S0: Case-specific inputs of all model variables: e.g., organic matter amendments, soil properties and water regimes[‡]	Huang et al., (1998, 2004); Xie et al., (2010)
	M - S1: Case-specific inputs of soil properties and water regimes; other model variables use averaged values for all 495 cases	
	M - S2: Case-specific inputs of organic matter amendments; all other model variables use averaged values for all 495 cases, the water regime was assumed to be multi-drainage irrigation.	
	M - S3: No case-specific inputs used for soil properties or organic matter amendments, the water regime was assumed to be multi-drainage irrigation.	

† Regression equation R1 was developed according to measurements conducted in continuously flooded fields, and the calculated flux was therefore adjusted by a scaling factor of 1.00, 0.65 or 0.56 for continuous flooding, single drainage or multi-drainage irrigation, respectively (IPCC, 2006).

‡ The water regimes in the CH4MOD model (Huang et al., 2004) are more specifically defined and differ from that of the IPCC (2006).

Table 2 Performance of the methods under different scenarios of data availability

Method	Bias of the estimation (~~B_R~~r_b)	Std. of the estimation residues (~~CV~~r_v)	~~Total error‡~~root of MSE (RMSE)
R1－S0	212.0 (62.1%)†	577.1 (163.3%)	596.0 (174.7%)
R2－S0	−1.3 (−0.4%)	407.8 (119.5%)	407.8 (119.5%)
R2－S1	−4.9 (−1.4%)	415.7 (121.8%)	415.7 (121.9%)
M－S0	−24.2 (−7.1%)	251.8 (73.8%)	253.0 (74.1%)
M－S1	−30.8 (−9.0%)	343.9 (100.8%)	345.2 (101.2%)
M－S2	−120.7 (−35.4%)	341.3 (100.0%)	362.9 (106.1%)
M－S3	−109.8 (−32.2%)	401.8 (117.8%)	416.6 (122.1%)

† Percentages in parentheses indicate the magnitude of the error relative to the overall average methane flux (341.2 kg CH_4 ha^{-1}~~F_c~~) for all cases~~, and F_c = 341.2 kg CH_4 ha^{-1} (Fig. 2a)~~.

~~‡ $Total\ \text{error} = \sqrt{B_R^{\,2} + CV^2} \times F_c$, same as Equation 4~~

Table 3 Methane emissions inventory and the uncertainties caused by model imperfection and errors in model input data

~~Rice~~	~~Harvesting area ($\times 10^6$ ha)~~	~~CH_4 emission (Tg)~~	~~σ_T(Tg)‡~~	~~95% CI§ (Tg)~~		
~~Early rice~~	~~5.96~~	~~1.22~~	~~0.39~~	~~0.58~~	~~—~~	~~2.08~~
~~Late rice~~	~~5.96~~	~~1.52~~	~~0.40~~	~~0.85~~	~~—~~	~~2.39~~
~~Single rice~~	~~16.93~~	~~3.69~~	~~0.75~~	~~2.37~~	~~—~~	~~5.30~~
~~All rice~~	~~28.85~~	~~6.43~~	~~1.53~~	~~3.79~~	~~—~~	~~9.77~~

Scenario	CH_4 emission (Tg)	σ_{b+v}§	σ_d	$\frac{\sigma_d}{\sigma_{b+v}}$	σ_T(Tg)	95% CI (Tg)
R1—S0	13.59	9.89	1.11	0.11	9.99	1.45—38.98
R2—S0	10.37	2.74	0.14	0.05	2.74	5.71—16.39
R2—S1	10.24	2.91	0.07	0.02	2.91	5.83—17.16
M—S0	6.43	1.15	1.00	0.87	1.53	3.79 — 9.77
M—S1	7.94	1.89	0.97	0.51	2.13	4.33 —12.62
M—S2	7.40	3.16	0.56	0.18	3.12	2.56 —14.75
M—S3	9.23	3.79	0.00	0.00	3.79	3.37 —18.01

~~‡ Calculated with Equation 6;~~

~~§ 95% CIs were calculated by assuming the gamma probability distributions, for which the shape and scale parameters were estimated via momentum methods.~~

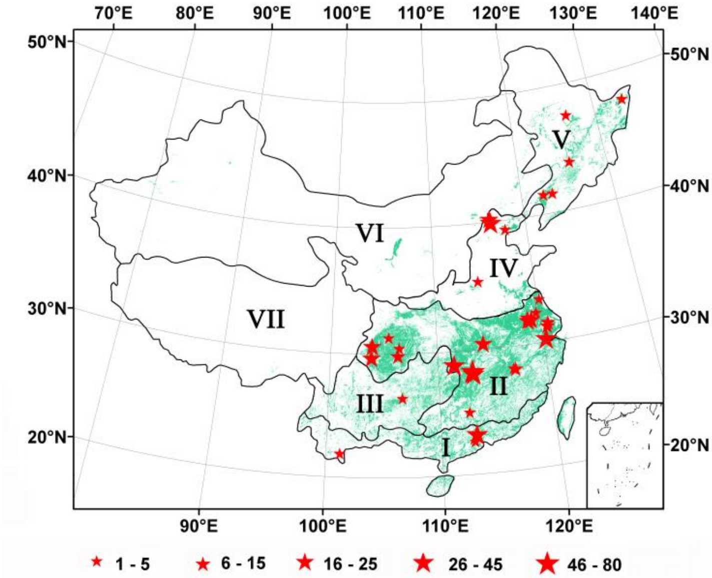
§ Root of U_b+U_v, uncertainty owing to model fallacy in the national inventory.

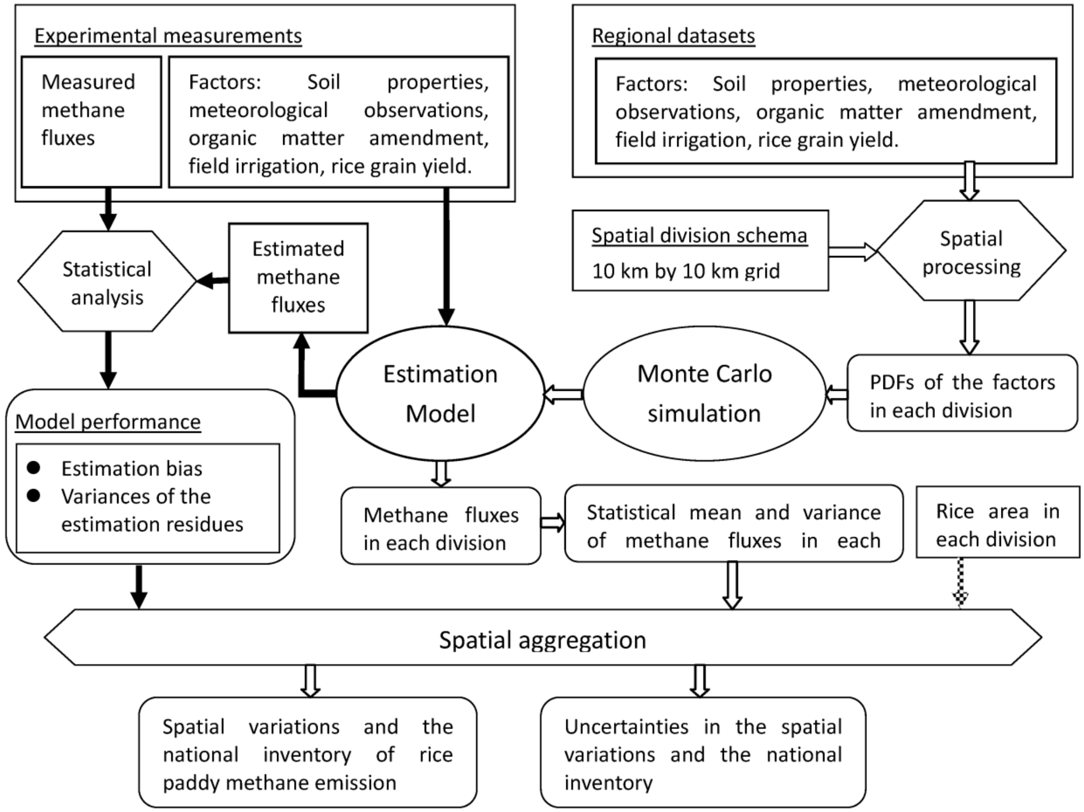
Table 4 Components of the uncertainty in the national inventory

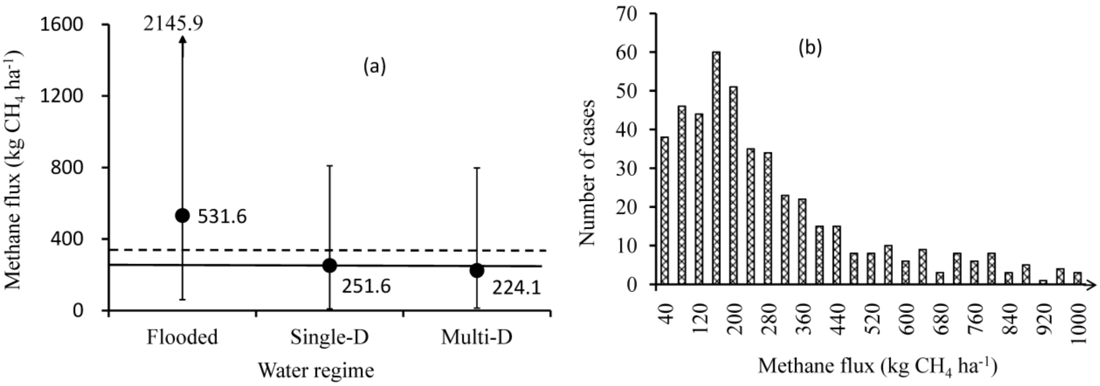
Rice	Due to model performance		Due to data quality and availability,	Total	
	$\cancel{U_B}U_b$	$\cancel{U_V}U_v$	$\cancel{U_D}U_d$	U_T	σ_T
Early rice	0.01	0.06(0.00—0.81)[‡]	0.08	0.15	0.39
Late rice	0.01	0.10(0.00—1.28)	0.05	0.16	0.40
Single rice	0.07	0.25(0.00—5.15)	0.24	0.56	0.75
All rice	0.21	1.12(0.00—22.56)	1.00	2.35	1.53

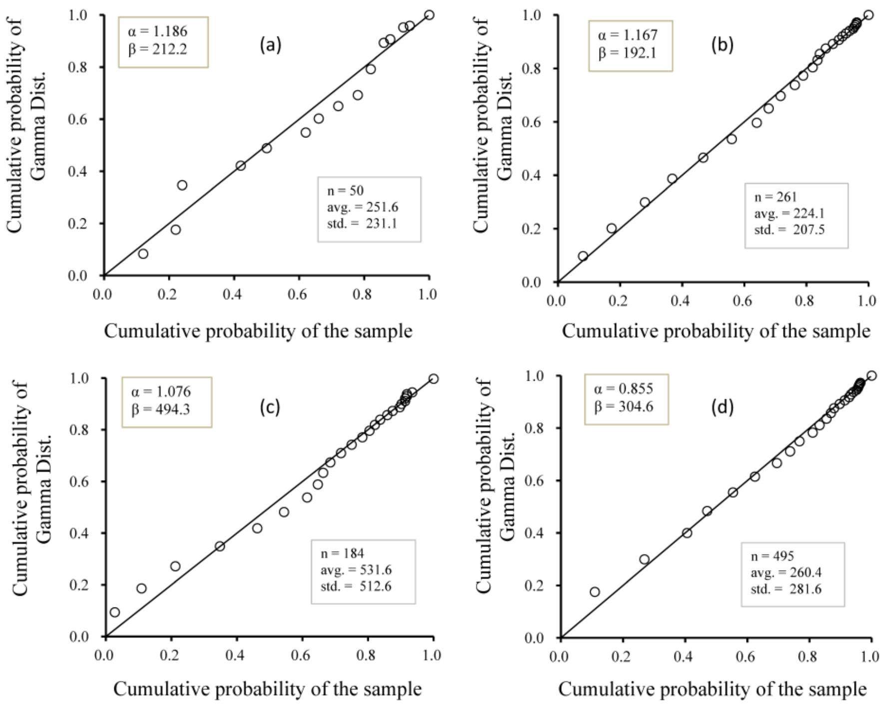
‡ Numbers in parentheses represent the range of $\cancel{U_V}$ U_v depending on the spatial correlation of the model simulation residuals. Long-distance correlation results in a large aggregated $\cancel{U_V}U_v$, whereas short-distance correlation results in a small aggregated $\cancel{U_V}U_v$.

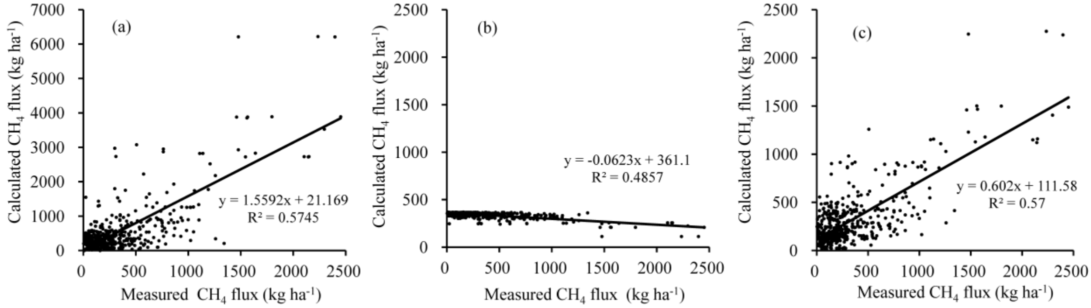
5 Figures

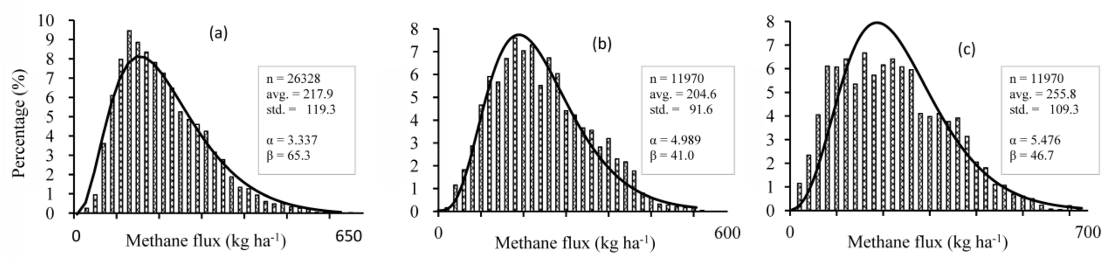

70°E
80°E
90°E
100°E
110°E
120°E
130°E
140°E
50°N
40°N
30°N
20°N
90°E
100°E
110°E
120°E
I
II
III
IV
V
VI
VII
1 - 5
6 - 15
16 - 25
26 - 45
46 - 80

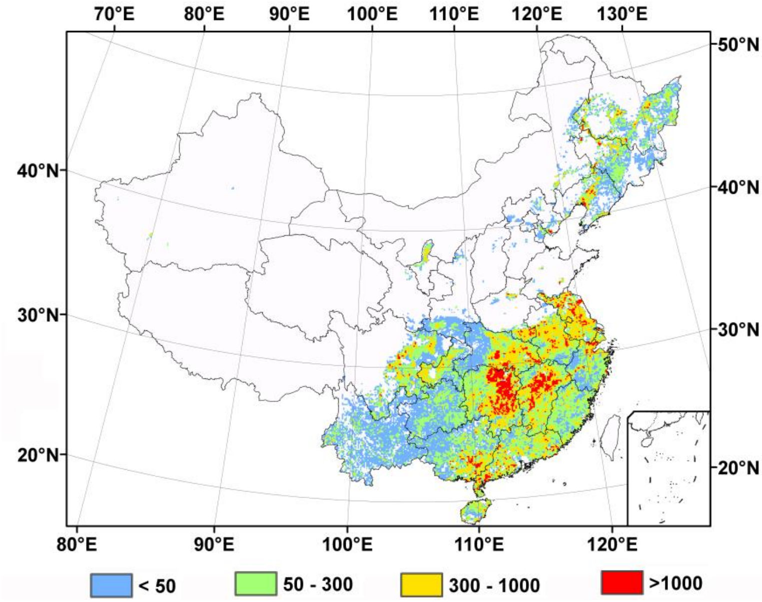

Experimental measurements
Measured methane fluxes
Factors: Soil properties, meteorological observations, organic matter amendment, field irrigation, rice grain yield.
Regional datasets
Factors: Soil properties, meteorological observations, organic matter amendment, field irrigation, rice grain yield.
Spatial division schema
10 km by 10 km grid
Spatial processing
Statistical analysis
Estimated methane fluxes
Estimation Model
Monte Carlo simulation
PDFs of the factors in each division
Model performance
Estimation bias
Variances of the estimation residues
Methane fluxes in each division
Statistical mean and variance of methane fluxes in each
Rice area in each division
Spatial aggregation
Spatial variations and the national inventory of rice paddy methane emission
Uncertainties in the spatial variations and the national inventory

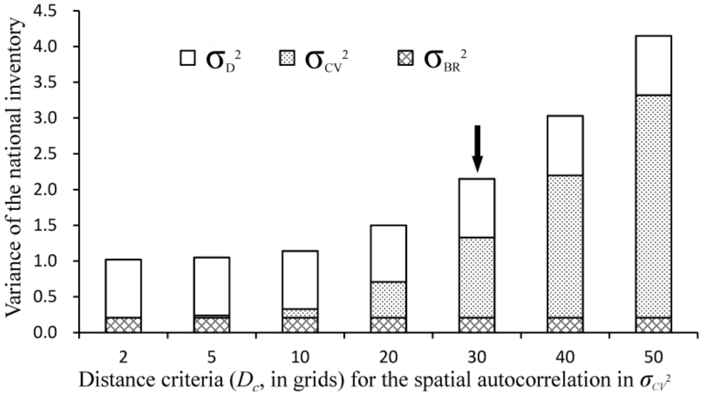

(a)
1600
1200
800
400
0
2145.9
531.6
251.6
224.1
Methane flux (kg CH4 ha-1)
Flooded
Single-D
Multi-D
Water regime
(b)
70
60
50
40
30
20
10
0
Number of cases
40
120
200
280
360
440
520
600
680
760
840
920
1000
Methane flux (kg CH4 ha-1)


(a)
α = 1.186
β = 212.2
n = 50
avg. = 251.6
std. = 231.1
(b)
α = 1.167
β = 192.1
n = 261
avg. = 224.1
std. = 207.5
(c)
α = 1.076
β = 494.3
n = 184
avg. = 531.6
std. = 512.6
(d)
α = 0.855
β = 304.6
n = 495
avg. = 260.4
std. = 281.6
Cumulative probability of Gamma Dist.
Cumulative probability of the sample
0.0
0.2
0.4
0.6
0.8
1.0


(a)
y = 1.5592x + 21.169
R² = 0.5745
Calculated CH4 flux (kg ha-1)
Measured CH4 flux (kg ha-1)
(b)
y = -0.0623x + 361.1
R² = 0.4857
Calculated CH4 flux (kg ha-1)
Measured CH4 flux (kg ha-1)
(c)
y = 0.602x + 111.58
R² = 0.57
Calculated CH4 flux (kg ha-1)
Measured CH4 flux (kg ha-1)


(a)
Percentage (%)
n = 26328
avg. = 217.9
std. = 119.3
α = 3.337
β = 65.3
0
650
Methane flux (kg ha-1)
(b)
n = 11970
avg. = 204.6
std. = 91.6
α = 4.989
β = 41.0
0
600
Methane flux (kg ha-1)
(c)
n = 11970
avg. = 255.8
std. = 109.3
α = 5.476
β = 46.7
0
700
Methane flux (kg ha-1)

5 
70°E
80°E
90°E
100°E
110°E
120°E
130°E
50°N
40°N
30°N
20°N
80°E
90°E
100°E
110°E
120°E
< 50
50 - 300
300 - 1000
>1000


σD2
σCV2
σBR2
Variance of the national inventory
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
2
5
10
20
30
40
50
Distance criteria (Dc, in grids) for the spatial autocorrelation in σCV2