The article "Ground cover rice production system facilitates soil carbon and nitrogen stocks at regional scale" by Liu et al. is based on sophisticatedly designed soil sampling from geographically representative field sites in Central China. I found it of good value to understand local soil responses to film coverage. Its novelty in regional scale may also provide supportive information to local policy makers. However, the data obtained in this article was largely devalued by its weak argument in Introduction, lack of rationalization in Method, as well by the far-fetched interpretation in Discussion.

>> Many thanks for your comments. In the revised version we have extended the introduction as suggested below. Also the chosen methodology is better justified. As we take from the reviewer comments, the rationale behind our argumentation in the discussion section was partly misunderstood. We have addressed the issues raised in the responses to the specific questions below, and have further clarified our argumentation. Also we have extended the introduction section of the revised version to provide a better framework for the later discussion of our results. However, in this context some suggestions of the reviewer, e. g., to move the biogeochemical framework we provide in the discussion for the interpretation of the results to earlier sections, appear contradictory to us. This would weaken the readability of the paper in our view (please also see our specific responses below). Finally, we believe that the interpretation is not far-fetched and provide the reasons for it in the answers to the specific comments.

Here are my general comments:

1) The authors very often cite a great length of literature in Discussion, which should have been reviewed and argued in Introduction to build up your own argument, clarify knowledge gap and rationalize your own research question.

>> In order to comply with this comment, we have extended the introduction section with more literature, thus guiding the reader more straightforward to the rationale behind the hypothesis we had developed to be tested in our experiment.

2) Why and how could you make a hypothesis (C, N stocks would reduce under GCRPS), but then observed completely opposite results? Do you indicate that you did sufficient literature review to guide you to such hypothesis? If yes, then how could you reject it later on with your own results? If no, then please take full use of literature review to thoroughly debate which factors could be relevant to increase or decrease C and N stocks under GCRPS.

>> Our hypothesis was based on general findings in the literature that aerobic soils have a higher organic carbon turnover and lower C content as compared to soils being predominantly anaerobic. Of course one can adapt their hypothesis to suit the findings, but we just report our initial hypothesis based on what is in the literature. What is wrong here, what would have been your starting hypothesis? For example, in the recent review article of Kögel-Knabner et al (2010, Geoderma), it is stated that the high soil organic matter content of Paddy soils may be associated with retarded decomposition under anaerobic conditions. Consequently, it appeared straightforward to hypothesize that the more aerobic conditions under GCRPS cultivation would increase the C loss rates and thus overall reduce the SOC stocks under GCRPS cultivation. Based on our extremely robust dataset as gained from regional sampling at 49 paired sites we found the opposite and thus rejected the hypothesis. All this illustrates

that soil organic matter dynamics of rice soils is still not well understood, as was also pointed out in this review article by Kögel-Knabner and colleagues.
Science is based on testing hypotheses and then either rejecting them or not.

3) Besides, if you decide to stay on the hypothesis of reducing C, N stocks under GCRPS, then it would be contradictory to use positive word such as “facilitate” in your article title.
>> Thank you for this comment, we agree that the title is ambiguous. Consequently we changed it to “Ground cover rice production systems increase soil organic carbon and nitrogen stocks at a regional scale”

4) The relevance of 13C, 15N, and respiration rates should have been clarified in Introduction, i.e. why these properties are relevant, what additional information can they provide than the total C and N, what they can tell you to support your argument? Otherwise, it would be lack of ground to just bring it up in Method and Results.
>> We added the necessary information based on your suggestion to the introduction section of the revised version.

5) Why did you air-dry all the soil samples before incubation? How much do you think such drying treatment will affect the mineralization potential? The community of microbes could change, I assume?
>> Soil sampling and final soil analysis (in this case C decomposition potentials) were done at different locations, with weeks between the sampling and the analysis. We only ensured a standardized treatment of sampling, as hundreds of studies before. This equal treatment of all samples allowed us to reduce storage bias resulting in a similar effect for all samples, as well as allowing us to adjust soil water content to make it consistent for all samples. The results are then referred to as “mineralization potential”, which makes it clear that they are the derived CO_2 fluxes from a standardized experiment and should not be confused with field measurements. Our procedure will likely increase mineralization rates, which is why it was termed “mineralization potential”. Nonetheless, this data can still provide qualitative information on differences in mineralization dynamics soil processes between Paddy and GCRPS soils.

6) The Results are better reorganized to first deliver the most primary results, link them with logics, and then the secondary results. For instance, information such as soil texture, pH and bulk density could be moved below, unless you can reasonably link them to your primary results C and N stocks. On the other hand, the average C and N assimilation of aboveground biomass could be considered to be moved up directly following the C and N stocks. This may make a better reading flow.
>> We have reorganized the results section according to your suggestion. However we would like to keep soil physical and chemical properties after the SOC and N stocks. Our logic is to firstly verify that the significant difference on SOC and N stocks does not come from soil differences between paired GCRPS and Paddy.

7) In the Discussion part, authors tended to use a lot of observations from other reports to

interpret the results observed in this study. This makes the Discussion less convincing. Peer reports should be used to compare and discuss, not to explain your own results.

>> We disagree that using the literature to back up and explain our results makes the discussion less convincing. We believe that it is common practice to use information from and existing theories developed in other studies to unravel our results explain observed results.

8) The authors attributed the greater total C and N in GCRPS to more residues returns. Have those newly returned C and N been converted to stable form? Or they are just less decomposed litter buried or simply mixed into topsoils?

>> Plant biomass production (aboveground and belowground) was higher for GCRPS. The fraction of aboveground residues left on the field and finally being ploughed is the same for Paddy and GCRPS. This results in higher total amounts (in kg) of residue returns. Figure 6 in the revised version shows comparatively more particulate organic matter (LF) in GCRPS than in Paddy systems, which supports the above statement.

9) With respect to C stabilization/liability, what does the 13C and 15N show? What could be captured from the 13C, 15N and mineralization rates? For instance, Figure 5 shows that Paddy soils are less depleted in 15N than GCRPS. This indicates that soils from GCRPS are less decomposed than that from Paddy, suggesting greater mineralization potential in GCRPS soils (I am not expert in stable isotope. Excuse me, if I am not correct here.). Then, why did Paddy soils show greater cumulative mineralization? What could be the factors? Local aeration, temperature, community or accessibility of microbes?

>> As outlined in the discussion, $\delta^{15}N$ of bulk soil N is a proxy for N loss pathways. These pathways are characterized by clearly pronounced discrimination of ^{15}N vs ^{14}N. Consequently, NH_3 volatilization or denitrification result in relative enrichment of ^{15}N in the remaining N compounds. Therefore, the higher $\delta^{15}N$ values in Paddy soils indicate a higher N loss. Hence, our observation of lower ^{15}N in the soils and plant leaves indicates less N volatilization along gaseous pathways (mainly NH_3 volatilization) for GCRPS than for Paddy fields. We did not show ^{13}C data in this manuscript and we cannot unequivocally answer the question why mineralization potential was higher in Paddy soils (this was a statistically insignificant trend only). As you already mentioned, the microbial community may have been changed after conversion from Paddy to GCRPS.

10) Why heavy fractions have significant differences before and after incubation, but other fractions do not. Does it have anything to do with the stabilization mechanism of SOC? And how? How does this then affect the mineralization, and SOC stocks?

>> HF is the fraction containing micro- and meso-aggregates, which together with the s+c fraction confer physical protection to SOM. However, the HF can potentially be very sensitive to changes in land use and management (Baldock and Skjemstad, 2000). During the incubation experiment, all soils were exposed to non-saturated conditions (60% WHC). It is therefore plausible that these large changes in soil redox conditions may affect Paddy systems more significantly that GCRPS (particularly in this sensitive fraction) as the latter ones generally occur under higher redox potentials (Liu et al. 2013; Tao et al., 2015). The relative decrease of SOC in the HF fraction is the

result of the disruption of micro- and meso-aggregates. This unprotected OM may get mineralized, comminuted, and thus incorporated in the s+c fraction, which after incubation invariably showed an increase in their relative contribution to total SOM.

Refs:

Baldock, J.A., Skjemstad, J.O.: Role of the soil matrix and minerals in protecting natural organic materials against biological attack, Org. Geochem., 31, 697-710, 2000.

11) When choosing the sampling sites, you also considered the time spans since adoption of the GCRPS technique. Then, did you do any analysis against the time variable? Any patterns of total C and N stocks over adoption time? Are the increase C and N stocks consistently observed in different adoption years? Are the increasing rates constant over different years? Could it be possible that the benefits of C and N increase only occur for the first several years and then cease when soil C and N stocks approach their maximum capacities?

>> These are interesting questions that we have also asked ourselves. We had included the factor "time since conversion to GCRPS" in our statistical analysis, but no significant effect was found despite a trend towards increasing differences between GCRPS and Paddy with time since conversion. Statistical insignificance in this case may be a result of the still relatively short times since conversion from Paddy to GCRPS (5-20 years since conversion) in the light of the slow rates of change for SOC stocks. Furthermore, we may not have had a sufficient sample size for all years of the 5-15 years period, preventing proper testing of the time factor. This also may have prevented clear responses to the question "Are the increasing rates constant over different years?" We did such an analysis, but no significant difference was observed. The last question remains our research goal in the future, however a thorough addressing of this question requires the availability of longer time series of Paddy-GCRPS conversion.

12) If out of practical reasons, it is just not feasible to investigate root biomass for all field sites. Then, why did you choose this particular site? How well this site could represent all other sites of different soil types, and varying altitudes?

>> Yes, it is just not feasible to investigate root biomass for all field sites. This particular site was chosen because it is a well-managed long-term monitoring site with well-documented agronomic history (e.g., Tao et al., 2015 European Journal of Agronomy). Hence, the risk for unrepresentative effects at this intensively studied site was very low.

13) In Conclusion part, it is better to summarize the key results first before relating to implications. The ideal case would be that the readers can get the most valuable information from just reading your conclusions.

>> This is what we did in the Conclusion part, first to summarize the key results and then relating to our research goal in the future. Nonetheless, we have slightly changed the conclusion section of the revised version to guide the reader more straightforward to the most important conclusion from our experiment.

Specific comments:

Page 3650

L13-18: lack of literature reference.

>> We added additional literature references in the revised version.

L18-20: “As with conventional paddy rice systems: : :as compared to Paddy systems: : :”: Either grammatically incorrect, or convoluted expressed.

>> This was changed in the revised version.

Page 3651

L5 to 30: There should be less description on general effects of SOM on soil properties, but more related to rice system and what could possibly be the effects of GCRPS to SOM.

>> Revised as suggested – we have added more rice-specific information, and describe potential pathways how GCRPS could alter SOM.

Page 3652

L5-7: Such detailed description should be moved to Method.

>> This sentence was moved to the Methods section in the revised version.

L20-22: Lack of literature reference or data source.

>> We added references in the revised version.

L23: What does “implications” mean here?

>> Due to additional field work, labor demand and costs (e.g., the need for buying the PE foliage), not all farmers have adopted this technique even though GCRPS has clear advantages.

Page 3653

L10: “180kgfertilizerNha-1”: improper way to express measurement unit. At least, there should be space between numbers and measurement unit. And, is it different from the above “150kgNha-1”?

>> We corrected this. Furthermore, we provide detailed information on field management and fertilization in the revised version in order to clarify the management regimes for GCRPS and Paddy. This includes a better clarification of the differences in fertilizer application rates between Paddy (180 kg N ha^{-1}) and GCRPS (150 kg N ha^{-1}).

Page 3654

L1-13: It would be much more convinced if you could provide some literature references for all the methods you used here.

>> Revised as suggested, we have added references to the revised version.

Page 3657

L12: “: : :no differences in average potential C mineralization rates: : :”: how did you calculate the average? You mean, averaging the mineralization rates over 200 days? Then, it seems meaningless to me. And why there are no differences in average but a higher value in cumulative mineralization rates?

>> We apologize for unclear writing of this section. The section was rephrased in order to clarify that mean cumulative C mineralization rates were not significantly different between soils of the Paddy and GCRPS systems. Paddy showed an insignificant trend towards higher C mineralization rates. We only used cumulative rates calculated for the 200 days period.

L21-25: These sentences should belong to Introduction.

>> We do not agree here – we believe that this sentence improves readability because it puts the subsequent discussion of our findings in a biogeochemical context.

Page 3658

L2-5: These sentences should belong to Method.

>> Again, we disagree. We provide this information here to explain why our findings may differ from results of earlier studies. This clearly belongs to the discussion in our view.

L10-14, L19-30: They should be used in Introduction.

>> Also here we do not agree. This sentence establishes an important link between our results and other research. We feel that it would be inappropriate to discuss all these details in the Introduction.

L15-19: These sentences are just repeating your description in Results.

>> This sentence contains results, yes, but this appears hard to avoid in order to establish the context between our results and earlier publications and thus serves to improve readability.

Page 3659

L1-11: If these sentences are moved to Introduction, then it could be a good literature review.

>> Also here we do not agree. As mentioned above, this sentence describes essential information required to explain our results. We feel that it would be inappropriate to discuss all of these details in the Introduction. We have however, made some changes to this in the revised version.

L14-19: Just from "higher cumulative C loss rates", you cannot directly get the conclusion that SOM under GCRPS is more effectively persevered. Besides, you did not do aggregate fractionation, you could not simply relate your interpretation to the conceptual model of Six et al., 2004.

>> We have eliminated such a conclusion, as we agree that the data provided does not justify that GCRPS provides greater SOM stability than Paddy systems. The section has been reorganized.

L20-25: Too much observations from other reports rather than your own observations.
Such interpretations are far-fetched.

>> This needs to be seen in the light of literature. We are not in agreement with this statement. Here we refer to Figure 6 in the revised version, which shows that the C content of the heavy

fraction significantly declines throughout the lab incubation experiment for Paddy soils only but not for GCRPS soils. This provides hints on the physicochemical protection mechanisms we discuss – and the cited literature explicitly deals with Paddy soils.

Page 3660

L2-15: Such discussion or information should have been either discussed in Introduction, or clarified in Method.

>> We disagree. We feel that it would be inappropriate to discuss details of the interpretation of natural abundance of ^{15}N in bulk soil N in the introduction of this paper. Again, having this in the discussion seems to us to be essential to the rationale behind our discussion. Omitting this here would allow only experts in the field of isotopic fractionation to follow our line of thought for the nitrogen turnover processes. From the previous comment above (comment 9) we conclude that it is essential to clarify why and how we interpret our $d^{15}N$ data.

L17-25: Most of these sentences should be mentioned earlier in Introduction or Method.

>> We do not necessarily agree with this. The introduction is a section that introduces the problem and place it in a general research context, but is not a section where one would go into every fine specific detail.

L29: It is not readily convinced to simply attribute "less loss of ammonia" to "the covering of soil immediately after fertilizing". More in-depth interpretation may be needed.

>> We have added a sentence with a more detailed outlined rationale on the mechanism of how covering the soil reduces NH_3 emissions. But this is well accepted in the current literature.

Technical comments:

Figure 1: I would suggest to place SOC content above and SOC stock below, as, logically, SOC stock is calculated from SOC content.

>> We have changed this in the revised version.

Figure 3: What does CAGB represent here? You did not explain it in your text body. The text body and figures should be consistent.

>> Carbon (CAGB) and nitrogen (NAGB) assimilated in aboveground biomass were calculated as the sum of grain and straw dry matter multiplied by grain and straw C or N concentration at harvest. We omitted these abbreviations in the revised version

Figure 4: Y-axis is missing.

>> Y-axis was added in the revised version.