nature portfolio

Peer Review File

The optimization of model ensemble composition and size can enhance the robustness of crop yield projections

Web links to the author's journal account have been redacted from the decision letters as indicated to maintain confidentiality

29th Jun 23

Dear Professor Yu,

Your manuscript titled "Optimizing ensemble configurations can enhance the robustness of crop yield projections" has now been seen by 3 reviewers, and we include their comments at the end of this message. They find your work of interest, but some important points are raised. We are interested in the possibility of publishing your study in Communications Earth & Environment, but would like to consider your responses to these concerns and assess a revised manuscript before we make a final decision on publication.

We therefore invite you to revise and resubmit your manuscript, along with a point-by-point response that takes into account the points raised. Please highlight all changes in the manuscript text file.

We are committed to providing a fair and constructive peer-review process. Please don't hesitate to contact us if you wish to discuss the revision in more detail.

Please use the following link to submit your revised manuscript, point-by-point response to the referees' comments (which should be in a separate document to any cover letter) and the completed checklist:
[link redacted]
** This url links to your confidential home page and associated information about manuscripts you may have submitted or be reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage first **

We hope to receive your revised paper within six weeks; please let us know if you aren't able to submit it within this time so that we can discuss how best to proceed. If we don't hear from you, and the revision process takes significantly longer, we may close your file. In this event, we will still be happy to reconsider your paper at a later date, as long as nothing similar has been accepted for publication at Communications Earth & Environment or published elsewhere in the meantime.

Please do not hesitate to contact us if you have any questions or would like to discuss these revisions further. We look forward to seeing the revised manuscript and thank you for the opportunity to review your work.

Best regards,

Alienor Lavergne, PhD
Associate Editor
Communications Earth & Environment

EDITORIAL POLICIES AND FORMATTING

We ask that you ensure your manuscript complies with our editorial policies. Please ensure that the following formatting requirements are met, and any checklist relevant to your research is completed and uploaded as a Related Manuscript file type with the revised article.

Editorial Policy: Policy requirements (Download the link to your computer as a PDF.)

Furthermore, please align your manuscript with our format requirements, which are summarized on the following checklist:
Communications Earth & Environment formatting checklist

and also in our style and formatting guide Communications Earth & Environment formatting guide .

*** DATA: Communications Earth & Environment endorses the principles of the Enabling FAIR data project (http://www.copdess.org/enabling-fair-data-project/). We ask authors to make the data that support their conclusions available in permanent, publically accessible data repositories. (Please contact the editor if you are unable to make your data available).

All Communications Earth & Environment manuscripts must include a section titled "Data Availability" at the end of the Methods section or main text (if no Methods). More information on this policy, is available at http://www.nature.com/authors/policies/data/data-availability-statements-data-citations.pdf.

In particular, the Data availability statement should include:
- Unique identifiers (such as DOIs and hyperlinks for datasets in public repositories)
- Accession codes where appropriate
- If applicable, a statement regarding data available with restrictions
- If a dataset has a Digital Object Identifier (DOI) as its unique identifier, we strongly encourage including this in the Reference list and citing the dataset in the Data Availability Statement.

DATA SOURCES: All new data associated with the paper should be placed in a persistent repository where they can be freely and enduringly accessed. We recommend submitting the data to discipline-specific, community-recognized repositories, where possible and a list of recommended repositories is provided at http://www.nature.com/sdata/policies/repositories.

If a community resource is unavailable, data can be submitted to generalist repositories such as figshare or Dryad Digital Repository. Please provide a unique identifier for the data (for example a DOI or a permanent URL) in the data availability statement, if possible. If the repository does not provide identifiers, we encourage authors to supply the search terms that will return the data. For data that have been obtained from publically available sources, please provide a URL and the specific data product name in the data availability statement. Data with a DOI should be further cited in the methods reference section.

Please refer to our data policies at <a
href="http://www.nature.com/authors/policies/availability.html">http://www.nature.com/authors/policies/availability.html.

REVIEWER COMMENTS:

Reviewer #1 (Remarks to the Author):

Thank you for providing the opportunity to review this important manuscript. I have carefully evaluated the content and have a few observations that need to be addressed before accepting it for publication. Please find my detailed comments below:
The manuscript 'Optimizing ensemble configurations can enhance the robustness of crop yield projections' by Li et al. is well-written, and the concept and problems are clearly described. I agree with the authors that it is important to evaluate how climate and crop simulation model ensemble configurations, such as model composition and size, influence crop yield projections and uncertainty. The authors present an optional framework based on a global-scale analysis to guide the selection of a representative ensemble subset from a large pool of models, such as AgMIP's GGCMI. They highlight the importance of considering the impact of specific models when selecting models for local-scale applications.

Major point:
The authors mention that some studies have analyzed the source of uncertainty based on a small number of crop models, potentially impacting the robustness of results if the crop models used have similar processes (Huang et al., 2022; Xiong et al., 2019). However, in some clusters (e.g., Figure 1b: cluster 2 N=3, cluster 3 N=2, Figure S2b: cluster 3 N=3, Figure S2 N=1), the number of crop models for different crops is small, with some cases having only one model (Figure S2 N=1). The authors should provide a justification for how their proposed model ensemble may not have the same impact on the robustness of results.

Minor points:
L142-144: The statement "As for the CO_2 fertilization effect, the maize yield changes showed a stronger decrease signal compared with wheat, rice, and soybean" (Figure S1) is incorrect. In fact, according to Figure S1, maize shows the least changes.

L178-189: It is difficult to identify how the GGCM emulator was mostly grouped by GGCM from Figure 1a. I suggest the authors provide a closer look at this particular case to help readers understand the grouping.

L210-213: In Figure 1d, cluster 2 should have n=56 and nGGCM=3 (not 2), and cluster 3 should have n=64 and nGGCM=2 (not 3). Additionally, in Figure S2, the ensemble size of cluster 2 for rice is n=72, and it has 4 GGCMs (models 2, 3, 4, and 5), not 3. Similarly, cluster 3 has 3 GGCMs. Therefore, the subsequent claims must be corrected accordingly.

Reviewer #2 (Remarks to the Author):

This manuscript investigates the influences of ensemble configurations on crop yield projections and modelling uncertainty arising from Global Gridded Crop Models (GGMs) and Global Climate Models (GCMs) under climate change. The analyses of evaluations are simple, easy to understand/interpret, and they are also so for the results. The research topic has significance and can provide reference for the application and improvement of crop model.

Special comments
1. The authors examine influences of ensemble configurations on crop yield in this manuscript. I would like to ask the authors to explain what may happen or how different the results might be when other model outputs/variables are analysed?

2. Additionally, I would like to ask the authors if quantifying uncertainty using only the mean proportion would be sufficient?

Reviewer #3 (Remarks to the Author):

This work focuses on ensemble configurations of Global Gridded Crop Models and Global Climate Models to investigate uncertainties in predicting the impact of climate change on several crops (wheat, maize, soybean and rice) at different spatial scales. This study identified the optimization of the ensemble configuration to increase modelling performances. For this, the Authors divided GGCMs into different clusters, then quantified the main source of uncertainty among simulations and finally investigated the proportion of uncertainty by individual GGCMs.

This study highlighted that the selection of ensemble configuration had a huge impact on the simulation outputs. A considerable effort has been made to compare all of these ensemble combinations and drawn conclusions for the modeling of major crops, in particular to find an optimum that will minimize uncertainty. Also this study highlighted an interesting point related to the spatial scale of modeling recurrent in spatialization processes (in addition to the selection of a relevant ensemble configuration). Questioning on how selecting the best combination and spatial scale should be of first interest in most studies involving crop modeling. Thus, your study is a valuable work to highlight this statement.

Concerning the calibration process of GGCMs, which strategy has been used to carry the modeling? At which spatial scale this calibration has been made? At the same resolution than the GCM? The resolution of each GCM is presented in Table S1 and highlighted that these resolutions were not the same for each of them. Downscaling processes have been applied to match with the finer one to have modeling at same spatial scale for each ensemble configurations?

All the objectives are enounced in the manuscript. However, for a faster understanding and to gain in clarity, objectives could be clearly highlight at the end of the contextualization.

For hierarchical clustering, on Fig 1. and S2, all the GGCM are not coupled with all the GCM, and each GGCM do not have the same number of GCM. Is it because a preselection was made upstream from the modelling to optimize the performances? If yes, on which criteria this performance was made?

For clustering to divide the sub-region, several type of variables were used. However, for the management variable, with only the N fertilizer input was taken into account?

The English used in the article is correct but the sentence structure is not usual. Indeed, it is more frequent to see in the scientific literature the use of the passive voice which is little used here.
You should correct these minor points:
- l. 490 : in the equation caption, there is a n explained, but no n in the equations, please clarify this.
- The references are not conformed to the Communication Earth and Environment, please refer to the Guide for Submission, you should correct it.

Reviews of Manuscript No.: COMMSENV-23-0601A

Title: Optimizing ensemble configurations can enhance the robustness of crop yield projections

We greatly appreciate the comments from the reviewers. We have fully considered the reviewers' comments in this revised manuscript. Our responses to the reviewers' comments are detailed below. We hope that you find our paper acceptable now for publication in Communications Earth & Environment. The original comments are followed by our responses in blue.

Reviewer #1

Thank you for providing the opportunity to review this important manuscript. I have carefully evaluated the content and have a few observations that need to be addressed before accepting it for publication. Please find my detailed comments below:

The manuscript 'Optimizing ensemble configurations can enhance the robustness of crop yield projections' by Li et al. is well-written, and the concept and problems are clearly described. I agree with the authors that it is important to evaluate how climate and crop simulation model ensemble configurations, such as model composition and size, influence crop yield projections and uncertainty. The authors present an optional framework based on a global-scale analysis to guide the selection of a representative ensemble subset from a large pool of models, such as AgMIP's GGCMI. They highlight the importance of considering the impact of specific models when selecting models for local-scale applications.

Thank you for your positive comments.

Major point:

The authors mention that some studies have analyzed the source of uncertainty based on a small number of crop models, potentially impacting the robustness of results if the crop models used have similar processes (Huang et al., 2022; Xiong et al., 2019). However, in some clusters (e.g., Figure 1b: cluster 2 N=3, cluster 3 N=2, Figure S2b: cluster 3 N=3, Figure S2 N=1), the number of crop models for different crops is small, with some cases having only one model (Figure S2 N=1). The authors should provide a justification for how their proposed model ensemble may not have the same impact on the robustness of results.

We agree with this sentiment. Our clustering approach aims to provide different possible conditions of yield changes under climate change. However, we understand that each cluster, especially those with fewer models, might not alone provide a robustness projection. Therefore, we use different clusters, each representing a unique potential outcome, to achieve a more comprehensive and robust understanding of the various ways that crop yields might change in response to climate change. This highlights the importance to consider different subsets, such as ours, to capture a broad spectrum of potential outcomes. We recognize that some of our initial statements may not have been entirely appropriate in this context. To clarify it, we rewrote this paragraph in our revised manuscript.

"Several studies have investigated the source of uncertainty based on different sets of crop models [19,22,34,35]. However, the impact of model compositions, which can significantly influence the uncertainty of crop yield projections, has often been overlooked. For example, some studies that focused on a particular region with a similar number of crop models showed varied results regarding the source of uncertainty [35,36]. Thus, we highlight the need to carefully select ensemble members in agricultural impact studies, given that different model composition can lead partial representation of overall uncertainty in the sample and thus to attributing varying relevant to different sources of uncertainty. In our study, although clusters are comprised of different numbers of models, each cluster represents a unique potential trajectory of yield changes under climate change. Such approaches can offer a more comprehensive understanding of the complexity and uncertainties involved in projecting crop yields under climate change." (L214-223)

Minor points:

L142-144: The statement "As for the CO2 fertilization effect, the maize yield changes showed a stronger decrease signal compared with wheat, rice, and soybean" (Figure S1) is incorrect. In fact, according to Figure S1, maize shows the least changes.

Thanks, we revised this sentence as follows

"As for the CO_2 fertilization effect, the wheat, rice, and soybean yield changes showed a stronger decrease signal compared with maize." (L131-132)

L178-189: It is difficult to identify how the GGCM emulator was mostly grouped by GGCM from Figure 1a. I suggest the authors provide a closer look at this particular case to help readers understand the grouping.

We have rewritten the associated paragraph to provide more clarity on this particular case.

"The clustering tended to group results by GGCM, indicating that, in general, differences between GGCMs were responsible for differences between clusters. For instance, cluster 2 in Figure 1a illustrates that EPIC-TAMU, GEPIC, PEPIC, and PROMET (around 44%) are clustered together due to similar yield changes, yet they include the projections from all 32 GCMs. This result suggests that in this cluster, the crop models are the primary factors influencing the projections of crop yield changes. However, it is worth noting that despite this crop model dominance in distinguishing ensemble members, different GCMs also contributed to the diversity of projections within each crop model group. This is evident as the associated GCMs in clusters 1 and 3 of Figure 1a do not consistently group by crop models. For example, both cluster 1 and cluster 3 contain projections from pDSSAT, LPJ-GUESS, and LPJml, each driven by different GCMs." (L142-152)

We have revised this sentence to rectify certain oversights in our original writing:

"For instance, in Figure 1a, the ensemble members (GGCMs driven by GCMs) are mainly grouped by the GGCMs, with only pDSSAT, LPJ-GUESS, and LPJmL appearing concurrently in clusters 1 and 3. For GGCM phase 3, the wheat yield changes within cluster 1 were primarily grouped by two specific GCMs (IPSL-CM6A-LR and UKESM1-0-LL)." (L177-180)

L210-213: In Figure 1d, cluster 2 should have n=56 and nGGCM=3 (not 2), and cluster 3 should have n=64 and nGGCM=2 (not 3). Additionally, in Figure S2, the ensemble size of cluster 2 for rice is n=72, and it has 4 GGCMs (models 2, 3, 4, and 5), not 3. Similarly, cluster 3 has 3 GGCMs. Therefore, the subsequent claims must be corrected accordingly.

To correct this, we modified these sentences as:

"For example, the ensemble size of cluster 2 (n=56 and n_{GGCM}=3, Figure 1d) and cluster 3 (n=64 and n_{GGCM}=2, Figure 1d) for maize were similar, but the dominant source of uncertainty exhibited large differences at different spatial scales (Figure 2c and d). For instance, GCM was the dominant source of uncertainty in Northern Europe, Eastern Asia, and USA for cluster 2 (Figure 2c). However, in cluster 3, GGCM was the dominant source of uncertainty in these regions. In addition, the ensemble size of cluster 2 (n=72 and n_{GGCM}=4, Figure S2) and cluster 3 (n=85 and n_{GGCM}=3, Figure S2) for rice had the similar number of GGCMs, but they show differences in the sources of uncertainties across sub-regions (Figure S8)." (L205-212)

Reviewer #2

This manuscript investigates the influences of ensemble configurations on crop yield projections and modelling uncertainty arising from Global Gridded Crop Models (GGMs) and Global Climate Models (GCMs) under climate change. The analyses of evaluations are simple, easy to understand/interpret, and they are also so for the results. The research topic has significance and can provide reference for the application and improvement of crop model.

Thank you for your positive comments.

Special comments

1. The authors examine influences of ensemble configurations on crop yield in this manuscript. I would like to ask the authors to explain what may happen or how different the results might be when other model outputs/variables are analysed?

Indeed, the selection of model outputs and variables could potentially impact our analysis results. If we were to analyze other model outputs or variables, the configurations of the ensemble group might have a slightly change, but the differences would be small enough not to change our overall conclusions as this is that yield is a function of the other variables. However, it's worth noting that crop yield is a common and key variable in crop modeling and impact studies. Despite the potential for slight variations when different variables are considered, the focus on yield provides a highly relevant basis for our analysis. Thus, we believe our choice of yield as the primary variable is appropriate given our research objectives. We have incorporated related statements in our revised manuscript.

"Since crop yield is a common and key variable in crop modeling, the investigation of different variables might result in only minor variations in the outcomes." (L474-475)

2. Additionally, I would like to ask the authors if quantifying uncertainty using only the mean proportion would be sufficient?

Thank you for your question. In our study, we have found that the use of the mean proportion provides a straightforward and easily interpretable measure of central tendency, which is suitable for our analysis. In addition, we also evaluated the standard deviation of GGCM-induced crop yield uncertainty (Figure S13) to enhance our uncertainty analysis. We understand that other uncertainty quantification methods, such as the use of variance or range, might offer additional perspectives in different scenarios. However, we think the mean proportion coupled with the standard deviation to be suitable for our study. We will consider more complex measures of uncertainty in future research where appropriate. We have noted this in the revised manuscript.

"In this study, we use the mean proportion of uncertainty as it offers a straightforward and readily interpretable measure of central tendency." (487-488)

Reviewer #3

This work focuses on ensemble configurations of Global Gridded Crop Models and Global Climate Models to investigate uncertainties in predicting the impact of climate change on several crops (wheat, maize, soybean and rice) at different spatial scales. This study identified the optimization of the ensemble configuration to increase modelling performances. For this, the Authors divided GGCMs into different clusters, then quantified the main source of uncertainty among simulations and finally investigated the proportion of uncertainty by individual GGCMs.

This study highlighted that the selection of ensemble configuration had a huge impact on the simulation outputs. A considerable effort has been made to compare all of these ensemble combinations and drawn conclusions for the modeling of major crops, in particular to find an optimum that will minimize uncertainty. Also this study highlighted an interesting point related to the spatial scale of modeling recurrent in spatialization processes (in addition to the selection of a relevant ensemble configuration). Questioning on how selecting the best combination and spatial scale should be of first interest in most studies involving crop modeling. Thus, your study is a valuable work to highlight this statement.

Thank you for your positive comments.

Concerning the calibration process of GGCMs, which strategy has been used to carry the modeling?

At which spatial scale this calibration has been made? At the same resolution than the GCM? The resolution of each GCM is presented in Table S1 and highlighted that these resolutions were not the same for each of them. Downscaling processes have been applied to match with the finer one to have modeling at same spatial scale for each ensemble configurations?

Thank you for your comments. The GGCMI Phase 2 models are not formally calibrated (Franke et al., 2020a; Franke et al., 2020b), and suitable calibration targets for global gridded crop model applications are not readily available. It is worth noting that, when focusing on relative yield changes, calibration can also have negative effects on model skill (Müller et al., 2017). We have added such information in the revised manuscript.

"The GGCMI Phase 2 models used in our study are not formally calibrated [14,31]. It is conceivable that a lack of calibration could influence the relative changes that are the basis of our study, but note that calibration can sometimes have negative impacts on model skill[15]." (L389-391)

In this study, we resampled this data to the 0.5° grid using a bilinear interpolation method to match the GGCM emulators (see Methods). Since we are only interested in the changes of multi-year average temperature and precipitation of the same GCM, no bias correction is necessary for the computation of ΔT and ΔP. We have noted this in the revised manuscript.

"We resampled this data to the 0.5° grid using a bilinear interpolation method. We used the change factor method to generate data for GGCM emulators. Since our study only focuses on the changes in future multi-year average temperature (ΔT) and precipitation (ΔP) in relation to the baseline of the same GCM, no bias correction is necessary for the computation of ΔT and ΔP." (L430-434)

Reference:

Franke, J.A., Müller, C., Elliott, J., Ruane, A.C., Jägermeyr, J. et al., 2020a. The GGCMI Phase 2 experiment: global gridded crop model simulations under uniform changes in CO2, temperature, water, and nitrogen levels (protocol version 1.0). Geoscientific Model Development, 13(5): 2315-2336.

Franke, J.A., Müller, C., Elliott, J., Ruane, A.C., Jägermeyr, J. et al., 2020b. The GGCMI Phase-2 emulators: global gridded crop model responses to changes in CO2, temperature, water, and nitrogen (version 1.0). Geoscientific Model Development, 13(9): 3995-4018.

Müller, C., Elliott, J., Chryssanthacopoulos, J., Arneth, A., Balkovic, J. et al., 2017. Global gridded crop model evaluation: benchmarking, skills, deficiencies and implications. Geoscientific Model Development, 10(4):

1403-1422.

All the objectives are enounced in the manuscript. However, for a faster understanding and to gain in clarity, objectives could be clearly highlight at the end of the contextualization.

We have rewritten the end of the introduction section as follows:

"The study investigates the influences of ensemble configurations on crop yield projections and modelling uncertainty from GGCMs and GCMs under future climate change. We analyze the importance of model composition and ensemble size in determining the primary sources of uncertainty in crop yield projections, with the aim of offering valuable guidance for optimizing ensemble configurations in climate-crop modeling studies." (L119-124)

For hierarchical clustering, on Fig 1. and S2, all the GGCM are not coupled with all the GCM, and each GGCM do not have the same number of GCM. Is it because a preselection was made upstream from the modelling to optimize the performances? If yes, on which criteria this performance was made?

The cluster analysis used in this study was not based on any preselection or performance optimization and 288 combinations (9 GGCMs, 32 GCMs) were used for wheat and maize and 256 (8 GGCMs, 32 GCMs) were used for rice and soybean. Instead, it was performed downstream, using the results of crop yield changes at every grid location. This allowed us to cluster ensemble members with high similarity in terms of spatial pattern and magnitude. This approach was specifically chosen to highlight the substantial dissimilarities in projected crop yield changes across different ensemble combinations. We have clarified this point in our revised manuscript.

"This approach was performed downstream (without any preselection), using the results of crop yield changes at every grid location. This allowed us to cluster ensemble members with high similarity in terms of spatial pattern and magnitude [41]." (L467-470)

For clustering to divide the sub-region, several type of variables were used. However, for the management variable, with only the N fertilizer input was taken into account?

In the GGCM-emulators, the management practices considered include water supply and nitrogen inputs. Precipitation was already considered as a separate variable in the cluster analysis. Therefore, in the context of the GGCM-emulator projections, only nitrogen fertilizer input was represented as a distinct management variable.

We have added such information in the revised manuscript.

"It's worth noting that the GGCM-emulators take into account management practices including water supply and nitrogen inputs. Given that precipitation is separately accounted for in the cluster analysis, within the framework of the GGCM-emulator projections, only nitrogen fertilizer input was represented as a distinct management variable." (L454-458)

The English used in the article is correct but the sentence structure is not usual. Indeed, it is more frequent to see in the scientific literature the use of the passive voice which is little used here.

We have made some revisions throughout the manuscript.

You should correct these minor points:

- l. 490 : in the equation caption, there is a n explained, but no n in the equations, please clarify this.

We apologize for the oversight. We removed the unassociated explanation of 'n'.

- The references are not conformed to the Communication Earth and Environment, please refer to the Guide for Submission, you should correct it.

We updated the reference format.

23rd Aug 23

Dear Professor Yu,

Your manuscript titled "Optimizing ensemble configurations can enhance the robustness of crop yield projections" has now been seen by our reviewers, whose comments appear below. In light of their advice we are delighted to say that we are happy, in principle, to publish a suitably revised version in Communications Earth & Environment under the open access CC BY license (Creative Commons Attribution v4.0 International License).

We therefore invite you to revise your paper one last time to address the remaining concerns of our reviewers. At the same time we ask that you edit your manuscript to comply with our format requirements and to maximise the accessibility and therefore the impact of your work.

EDITORIAL REQUESTS:

Please review our specific editorial comments and requests regarding your manuscript in the attached "Editorial Requests Table".

*****Please take care to match our formatting and policy requirements. We will check revised manuscript and return manuscripts that do not comply. Such requests will lead to delays. *****

Please outline your response to each request in the right hand column. Please upload the completed table with your manuscript files as a Related Manuscript file.

If you have any questions or concerns about any of our requests, please do not hesitate to contact me.

SUBMISSION INFORMATION:

In order to accept your paper, we require the files listed at the end of the Editorial Requests Table; the list of required files is also available at https://www.nature.com/documents/commsj-file-checklist.pdf .

OPEN ACCESS:

Communications Earth & Environment is a fully open access journal. Articles are made freely accessible on publication under a CC BY license (Creative Commons Attribution 4.0 International License). This license allows maximum dissemination and re-use of open access materials and is preferred by many research funding bodies.

For further information about article processing charges, open access funding, and advice and support from Nature Research, please visit https://www.nature.com/commsenv/article-processing-charges

At acceptance, you will be provided with instructions for completing this CC BY license on behalf of all authors. This grants us the necessary permissions to publish your paper. Additionally, you will be asked to declare that all required third party permissions have been obtained, and to provide billing information in order to pay the article-processing charge (APC).

Please use the following link to submit the above items:
[link redacted]
** This url links to your confidential home page and associated information about manuscripts you may have submitted or be reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage first **

We hope to hear from you within two weeks; please let us know if you need more time.

Best regards,

Alienor Lavergne, PhD
Associate Editor
Communications Earth & Environment

REVIEWERS' COMMENTS:

Reviewer #1 (Remarks to the Author):

Now the article is well-prepared for acceptance and publication. The raised queries have been thoroughly addressed, and the authors have made valuable revisions to enhance the quality of the article.

Reviewer #2 (Remarks to the Author):

The reviewer has no additional comment.

Reviewer #3 (Remarks to the Author):

The comments arisen in the first phase of reviewing has been corrected. Thanks for the clarification of the main points, especially concerning the objective (now clearly highlighted) and concerning the clustering in the Material and Methods part, this is clearer. Thanks for the modification of the reference format. I am satisfied with this version of the manuscript.