Dear RC2,
Thank you very much for your feedbacks.
I will reply with the format of *italic text from RC2* and regular text from author in 2 sections (major and minor comments).

About major comments,
"1. The authors' reparametrize an existing crop model for rice to extend its usage to soybean. However, soybean is a legume and therefore capable of symbiotic nitrogen fixation. It is my understanding that neither symbiotic nitrogen fixation nor nitrogen uptake are explicitly simulated in the model but are implicitly captured by the nitrogen limitation (SLN see also major comment 2). Even if symbiotic nitrogen fixation is captured implicitly reducing nitrogen stress, the respiratory costs of symbiotic nitrogen fixation are not considered. Yet, these costs are essential when determining the pathway of nitrogen uptake a plant follows depending on environmental conditions. For example, Fischer et al. (2006) show that the simulation of N uptake requires the optimization of the respiration costs of different uptake paths to reduce respiration losses. Neglecting these respiratory costs may lead to an overestimation of NPP and in turn crop yield under nitrogen limited conditions. When assessing Fig. 5 and 6 a), MATCRO-Soy indeed shows a systematic overestimation of FAO yields. It would be interesting to assess whether this overestimation is stronger in countries with lower fertilization rates which could indicate that the missing representation of the cost of N fixation contributes to the overestimation. Additionally, I would ask the authors to explain why they do not consider respiratory costs of biological nitrogen fixation and the implications of this assumption and also considers this when discussing the limitations of their model. What would be potential solutions to solve this issues in future model versions? "
Reply:

1. We agree about the importance of respiratory costs in symbiotic nitrogen fixation, particularly under nitrogen-limited conditions. In MATCRO-Soy, both nitrogen fixation and uptake are implicitly constrained through specific leaf nitrigen (SLN). The respiration costs linked to different uptake pathways, including biological nitrogen fixation, are not explicitly represented. This simplification may indeed contribute to the overestimation of yields under low-N input countries, as noted in Fig. 5 and 6a (e.g. Bolivia, Paraguay, Russia). Those countries has nitrogen fertilizer input less than 25 gNm^{-2} for the growing season.

   According to your comment, we deeply investigated again the cause of overestimation. We found current MATCRO-Soy overestimated the growing degree days (related to harvest time), where we should use different formulation for hourly growing degree days to fit with the model structure. Hence, we will update the results in revised manuscript. This revision has lower yield value in Argentina, Brazil, Canada, and Italy as the growing period become lower compared with the previous version.

   However, the lack of respiration cost consideration from nitrogen fixation is inevitable affecting the accuracy in the model prediction. We highlighted this limitation clearly in the revised manuscript and acknowledge the need to incorporate variable respiration costs for different nitrogen uptake pathways. (L712-718)"While nitrogen uptake and fixation are implicitly constrained through the SLN parameter, the additional carbon costs of BNF are not simulated. This simplification may partly explain the overestimation of yields in low-input systems, as seen in countries such as Bolivia, Paraguay, and Russia, where nitrogen fertilizer inputs are below 25 g N m^{-2} per growing season (Fig. 5 and 6a). This highlights an important opportunity for future model development to incorporate variable of respiratory costs for different nitrogen uptake pathways."

   About the question, we missed to adopt the method due to the simplification of biological nitrogen fixation process in the model. Quantifying respiration costs of BNF at a global scale require more detailed input (e.g. soil nitrogen). Until now, introducing this component without well-supported parameters may introduce more uncertainty than insight for MATCRO-Soy in my understanding. We might planning to consider this mechanism in the future into the model structure with further examination for its effect.

"2. The explanations of Eq. 8, 9 17, 18 and 19 are incomplete. For example the explanation of eq. 8 and 9 is missing the explanation of SLN_{Y0}, SLN_{Y1}, SLN_{Y2} and SLN_{Y3}, $SLN_{Y3,h}$, $SLN_{Y3,l}$ and $N_{fert,max}$. All of these seem to be parameters that are listed in tables later on but this is not clearly explained. In addition I think that the assumption that biological nitrogen fixation (BNF) is captured through SLN needs a thorough explanation. Do the authors assume that BNF is captured by N_{fert}?. If so this should be stated and its implications need to be discussed. This is also related to major comment 1 regarding the assumption to not represent the respiratory costs of BNF. Similar in eq. 17, 18 and 19

p_{leaf0}, p_{leaf1}, p_{leaf2}, p_{pod1} *and* p_{pod2} *are only listed in the tables but not explained."*

Reply:

2. Thank you for pointing this out. In the current model version, we implicitly assume that biological nitrogen fixation (BNF) is captured the empirical nitrogen limitation factor in the leaf (SLN) across developmental stages using experimental data by Menza et al. (2023). This simplification allows the model to simulate yield responses under varying nitrogen availability without explicitly separating soil N uptake and BNF.
   We revised in the manuscript (L216-217): Nitrogen uptake including biological nitrogen fixation is implicitly captured through SLN that influence Vcmax in Eq. (7) and (8), while the effect of applied fertilizers in Eq. (8) and (9).

   However, we acknowledge that this approach does not differentiate the source of nitrogen (from soil and nitrogen fixation) and therefore omits the associated respiration costs specific to BNF. This may contribute to yield overestimation under nitrogen-limited conditions, as discussed in our response to major comment 1. We have clarified this assumption in the revised manuscript and added a discussion on its implications in the model limitation section. Future improvements will involve representing nitrogen sources separately and incorporating the respiratory cost of BNF to better capture the carbon–nitrogen trade-off.Implication is written in L712-718, the same as major comment 1.

   Regarding Equations 17–19, we revised the text to explicitly define the variables (L251-253).

"3. *Eq. 25 and 26: Water stress is a factor to reduce the yield at harvest. It is not clear whether only the value for the water content of each soil layer at the day of harvest is used to calculate water stress or if this is integrated over the entire growing season but only applied at harvest. In both cases only applying water stress at harvest has its limitations (e.g. such an approach misses propagation effects of early season water stress as well as drought mortality) which should be discussed."*

Reply:

3. In this model, we apply the water stress to be integrated over the entire growing season, not only in the harvest time. Then we believe it still captures the effects of early season of water stress. I added explanation of (L289)"during the crop growth" for water stress factor calculation and sentence (L284-285)"The initial value of WSL is set equal to the soil porosity, assuming the soil is saturated at the beginning of the simulation."

"*4. The authors use segmented linear models to estimate several parameters in different developmental stages (DVS). However, they do neither provide the software they used to create these models nor how the models were trained. Therefore, I think the answers to the following questions need to be added to section 3.2: How where the initial breakpoints estimated? What kind of optimization method was used?"*

Reply:

4. We trained the segmented linear models (Fig. 2 and 3) by estimating the glucose partitioning ratio with biosynthetic process simulation (Penning de Vries et al. 1989). The initial breakpoints were determined based on assumption based on phenological growth characteristics. For example, we assume the glucose ratio is mostly growth for leaf toward the shoot (about 60%) compared to other organ in the initial stage, then decline after emergence and stop growing after pod development. These breakpoints were then optimized by fitting the partitioning functions to the calibration data. We have clarified this method and added the explanation in Section 3.2.
   (L352-356): The segmented linear models for glucose partitioning were captured from the pattern from biochemical framework of Penning de Vries et al. (1989), with initial breakpoints estimated from phenological patterns (e.g., leaf allocation is high in early stage and declining after emergence). These breakpoints were then optimized by fitting the functions to calibration data.

"*5. Detrending: It looks like the authors apply linear detrending which removes both the slope and intercept of the linear regression from the observations and simulations respectively. This should at least be explained when the term "detrending" is introduced. In addition, I think the authors should highlight that the detrended comparison is useful to evaluate interannual variability and sensitivity to climate variability but mention that linear detrending removes important signals from the data."*

Reply:

5. Thank you for the helpful suggestion. We applied linear detrending by removing the slope and intercept of the linear regression from both observed and simulated yield time series for each country. We agree with reviewer that this approach removes important trend-related signals, and we added this when the term "detrending" is first introduced in the result (L476-480) in section 5.1: We applied linear detrending for both observed and simulated yield in time-series by removing the slope and intercept of the linear regression. It isolates the variability primarily driven by climate fluctuations to evaluate interannual variability independent of long-term trends. However, it also removes longer-term signals (e.g. effect of technological improvements or increasing CO_2 concentrations).

   We also state that detrended comparisons are useful for evaluating the model sensitivity to climate variability, but not for assessing its ability to capture long-term yield trends or absolute yield levels in L594-596 (Section 6.1): While detrended comparisons are useful for evaluating interannual variability and the model sensitivity to climate fluctuations, detrending removes important long-term signals related to technological improvements, management changes, or increased CO_2 effects.

"*6. L622-627: The authors describe the role of CO2-fertilization effects for simulations results and hypothesize that this is the reason for the positive bias compared to observations. I would appreciate a more thorough explanation how the authors come to this conclusion. What is their explanation for the models ability to capture the temporal trend of the yield increase well but systematically overestimate yields? If this is all explained by CO2- fertilization, this indicates that the model captures the temporal development of the effect well (slope) but not the overall magnitude of the role of CO2 (bias) for yield formation. Additionally, while the CO2 -fertilization is one possibility, I do not think that this is the only possible explanation that is supported by their results. As stated in major comment 1 and 3, underestimation of respiratory costs for symbiotic nitrogen fixation and water stress are also likely explanations that should be discussed.*"

Reply:

6. Thank you for the critical point. We agree that CO_2 fertilization is not the sole explanation for the observed yield overestimation. In our results, while the model successfully captured the temporal trend (slope) in yield increase, it systematically overestimated absolute yield levels (bias). This suggests that the overestimation is not due to the CO_2 fertilization effect under current conditions. However, we acknowledge that elevated CO_2 could contribute to greater overestimation in future projections if its effects are overrepresented. We also agree that other factors, such as underestimated respiratory costs of symbiotic nitrogen fixation and insufficient water stress representation are likely contributors. We have revised the discussion to reflect this explanation more clearly.

Original manuscript: (L719-724) "The simulated yield increases throughout the year, driven by the positive effects of increased atmospheric CO_2, a phenomenon known as the CO_2 fertilization effect, as observed in studies by Long et al. (2005) and Sakurai et al. (2014). Compared with simulations using statistical radiation use efficiency, process-based models have this tendency because of the greater effect of CO_2 on the photosynthesis process (Ai and Hanasaki, 2023). This result is expected, as most of the simulated yield values were overestimated compared with the reference data, except for the yield in Canada, which was due to the low temperature conditions".

Revised: (L720-729) The simulated yield increases throughout the year driven by the positive effects of increased atmospheric CO2, a phenomenon known as the CO2 fertilization effect, has been observed in studies by Long et al. (2005) and Sakurai et al. (2014). The CO_2 fertilization response may become a more prominent source of overestimation in future projections if the model overestimates the crop response to elevated CO_2. Compared with simulations using statistical radiation use efficiency, process-based models have this tendency because of the greater effect of CO2 on the photosynthesis process (Ai and Hanasaki, 2023). While the model captures the temporal trend in yield increase reasonably well, it tends to overestimate the absolute yield levels. This indicates that the overestimation is not necessarily attributable to the CO_2 fertilization effect under current conditions. Other contributing factors are underestimation of respiratory costs associated with symbiotic nitrogen fixation and insufficient representation of water stress responses. Moreover, the accuracy of data input also affects the model performance. For example, most of grid cells in Brazil show low yields due to input data indicating a short growing period.

About minor comments,

Recommendation from RC2 of grammar change for point 1-3, 5-8, 12-13, 15-20, 24, 26-27, and 29 each are adopted in the revised version (total 29 points).

Other points:

- Point 4: *I believe it should be "AgMIPs efforts" instead of "AgMIP efforts"*
  I personally think the acronym of AgMIP is robust for "Agricultural Model Intercomparison Project" as a noun, hence I will change the phrasing from "AgMIP efforts have demonstrated" to "AgMIP have demonstrated" (L63) than using "AgMIPs efforts". If you think it is not correct, please let me know.

- Point 9: *Is Masutomi et al., 2019 referring to the ozone implementation? If yes this is not clear from the current sentence structure*
  Yes, it is about the ozone implementation in the model
  original manuscript:
  "Crop growth mechanisms that consider photosynthesis and stomatal conductance, which are widely used to assess the impact of greenhouse gases on carbon and water fluxes (e.g. ozone) in Masutomi et al. (2019), have been incorporated."
  revised (L69):
  "The mechanisms that consider photosynthesis and stomatal conductance to assess the impact of greenhouse gases on carbon and water fluxes (e.g. ozone) have been incorporated into MATCRO-Rice as described in Masutomi et al. (2019)."

- Point 10: *It is not clear to me what the authors mean by carbon allocation driven by photosynthetic activity. To me carbon allocation is the distribution of carbon between different plant organs – here called carbon partitioning – while photosynthetic activity drives carbon assimilation. Please clarify*
  Sorry for the confusion,
  revised:
  (L93-94) The photosynthesis and carbon partitioning modules are closely linked, as carbon assimilation from photosynthesis is subsequently allocated to different plant organs.

- Point 11: *I do not understand what is meant by "The phenology module serves as a time dimension"*
  Original: The phenology module serves as a time dimension
  revised: (L87) The phenology module serves as a time-dependent framework

- Point 14: *Is de Vries et al., 1989 the reference to "school of De Wit" or MACROS? Also the acronym MACROS has not yet been introduced.*
  reply:
  de Vries et al. (1989) is not the original reference for "school of De Wit" but using its concept and described parameterization for MACROS (Modules for an Annual CROp Simulator). Pardon for the confusion, I deleted MACROS and revised as follows:
  (L108) The carbon partitioning module distributes the glucose into each organ (i.e. leaf, stem, root, and storage organ) following the method derived from the school of de Wit by simulating the biosynthetic process (Penning de Vries et al., 1989).

- Point 17: *Similarly why are b, h and o used for minimum, maximum and optimum? Why not use min, max and opt which are far more intuitive?*
  Revised:
  The parameters Tb, To, and Th are crop-specific and represent the base, optimum, and highest temperatures for crop development, respectively.

- Point 21: How can the developmental stage (DVS) be negative? I would assume it starts at 0.
  Reply: Yes, it starts at 0. We revised the Eq. 17

- Point 22: *I am a bit confused by the specific leaf weight parameter, this seems to be the same as leaf mass per area (LMA) trait. If so why not use the common definition? Or this not per leaf area but per agricultural area? If this is the case this should be mentioned. There also seems to be a discrepancy regarding the unit …*
  Reply: The unit is per agricultural area, we have revised in (L255) "LAI is computed using the specific leaf weight per agricultural area (SLW in kg ha-2) expressed as", with unit of SLW in Table 2 is kgha^{-1}

- Point 23. *Fig. 2a): How is the first segment fitted to the data if there are only values for DVS of around 0.2 or larger?*
  Reply: It is based on assumption, I added this explanation in the revision "The first segment of the fitted line (DVS < 0.2) was based on model assumptions due to sparse observational data in early stage." (L350-351)

- Point 25: (L431)*Why does the model have a higher accuracy for countries with high production levels? Are fertilization levels in these countries higher and water stress levels lower? If so this should also be discussed in the context of N-fixation (major comment 1) and water stress (major comment 3).*
  Reply:
  We consider the top producer's countries have high accuracy primarily due to the better data quality, well-managed production systems, and calibration data from these countries. Among the top producers of 10 countries, 3 of them (China, India, Italy) have relatively high N fertilization (>30 gNm^{-2} per growig season). In contrast, most of top producers (US, Brazil, Argentina) rely more on the biological nitrogen fixation (represented by Specific Leaf Nitrogen in the model), and do not have lower soil moisture. Therefore, we don not consider high N fertilization or lower water stress to be the primary drivers of model accuracy in high-producing countries, hence we don't include additional discussion on this aspect.

- Point 28: *This paragraph is very general and I am missing the connection to the MATCRO- Soy's strengths and application.*

  Original manuscript:

  Furthermore, the integration of crop models with remote sensing data will open new possibilities for monitoring and predicting crop productivity at finer spatial scales (Basso et al., 2001). Climate change may shift favourable conditions for high yields in the United States or worsen the challenge of low yields in India for producing soybean yields, making it essential for the model to project these trends accurately for future agricultural planning. In addition to climatic factors, variations in yield may be attributed to technological advancements, shifts in agricultural practices, and changes in crop management strategies that have not been considered in the model. The consideration of more detailed mechanisms of soybean growth can be considered for more accurate results as climate change affects the pest populations (Chen and Mccarl, 2001). However, it is important to acknowledge the limitations of the model, particularly its ability to predict yield variations under extreme or rapidly changing climatic conditions. Continuous updates of the experimental dataset are necessary to maintain its relevance and accuracy in predicting future soybean yields.

  I rewrite the paragraph by first highlighting the strength of the model by describing the specific important parameters in L633-635. Then talking about application in L636-643.

  Revised:
  (L633) "The strength of MATCRO-Soy lies in its ability to simulate key physiological processes of soybean growth (e.g. photosynthesis, phenology, and biomass partitioning), under varying climatic conditions. Its process-based structure allows for sensitivity analysis for further environmental impacts evaluation, such as effects of elevated CO_2 and temperature stress. The model has been shown to reasonably capture the temporal dynamics of yield formation. The integration of crop models with remote sensing data enhances its capability for monitoring and predicting crop productivity at finer spatial scales (Basso et al., 2001). Climate change may shift favourable conditions for high yields in the United States or worsen the challenge of low yields in India for producing soybean yields, making it essential **(L640)** for the model to project these trends accurately for future agricultural planning by using more detailed input. In addition to climatic factors, variations in yield may be attributed to technological advancements, shifts in agricultural practices, and changes in crop management strategies outside the scope of model can further improve the accuracy at the local scale. For example, including pest and crop interaction may enhance the model's capability to reflect local yield response to climate-driven pest dynamics (Chen and Mccarl, 2001). However, it is important to acknowledge the limitations of the model, particularly its ability to predict yield variations under extreme or rapidly changing climatic conditions. Continuous updates of the experimental dataset are necessary to maintain its relevance and accuracy in predicting future soybean yields."