**Response to reviewer 3 questions**

Reviewer #3:

Major comments:  
This paper on 'Risk-based evaluations of competing agronomic climate adaptation strategies: The case of rice planting strategies in the Indo Gangetic Plains' has developed customized planting dates of rice in IGP, in a quite big area ranging from latitude 22-32 °N and longitude 75-90 °E. Actually, this study has made a lot of effort, however as they hugely depended on secondary and simulated data all the time and tried to develop their framework and recommendation on the basis of those datasets without ground validation, the reliability, applicability, and trustfulness of the results presented is very low. Authors mapped rice and wheat area without following proper method of mapping (excluding water bodies, other vegetation, infrastructure, etc). Also, the novelty is not sufficient to publish in Agricultural Systems.

Response: The question of huge area is inconsequential. There are many global models. Similarly on mapping out water bodies, infrast, et.c., is inconsequential on the same basis. We could put a crop mask e.t.c, but all gridded crop simulations report similarly.

On the trustfulness, reliability and applicability, it depends on whether the reviewer trusts any crop growth model results. If not, then our paper indeed will not help him/her. However, agricultural systems journal has published many other crop growth model based paper including recently by Wang et al 2024 focusing on the area of interest.

We use APSIM growth model which has already been well calibrated and validated with already published papers that we cite having discussed these aspects of the model.

Other major comments:  
1. Too much dependency on simulated results, and computer programming and programmer (Lines 204-205) having/including no or very low real ground data.

Response: We thank the reviewer for the comment and agree with the importance of ground data for informing agricultural studies, For the purpose of this paper, the model inputs are based on ground data and the validated model is already published in peer reviewed articles (Balwinder, Urfels). In addition, the purpose of this data is about methods innovation on how to evaluate production systems from a risk perspective. Since long-term data of different management strategies does not exist at scale – we rely on simulated results to produce a dataset that is handy to illustrate our method for the IGP. Of course, one may apply the same method for long-term trials or other longitudinal datasets but remain outside the scope of this paper and do not cover spatial variation of climate.

2. Model calibration and validation work was dependent on previous work by Balwinder-Singh from CIMMYT. For how many varieties and agro-ecological conditions had he calibrated for rice and wheat crops? As the study area covers a huge area (spreading from latitude 22-32 °N and longitude 75-90 °E), I do not believe the users do not need to re-calibrate and validate the models for different biophysical and management conditions. When the scenarios for planting strategies were made, for how many climatic conditions was the APSIM model run?

Response: This again is a question on the crop model not necessary on what our focus in the paper is. First, the mode outputs were evaluated for their performance across the area of interest and the reviewer may refer to the original paper for these evaluations which is published in a peer-reviewed journal. Second, we could of course work further developing the model but this remains outside the scope of this paper. In principle, we could have used any crop model or any hypothetical results to illustrate our case but using peer-reviewed and published results provide more useful context and help to refine the conclusions that were drawn from the initial paper and thus improve the literature by improving the evaluations. The focus of the paper is on making risk based evaluations that consider risk aversion of the farmers. Something that is not well done in the literature.

3. Highly coarse mapping work. Proper mapping needs to disaggregate rice, wheat, and rice-wheat systems area systematically in the 1st step. And then using current farmers' practice data as the baseline, further scenarios and willingness to pay can be implemented. Mapping everything (buildings, other vegetation, water bodies, etc, in rice or wheat areas) is not the correct approach. Even for the disaggregated maps, validation (matching between observed and simulated) is required.

Response: We thank the reviewer for the comment. This again is not needed for this paper. If there is a building, surely no one will plant rice or wheat there. There is no need for us to go into that unnecessary granularity for the sake of illustrating our novel method and results.

As specified in the Methods section, we do not use farmer’s practice as baseline because it did not have data for all pixels as such it would have affected crop model comparisons for all the other scenarios in those pixels. We matched pixels, the choice of the baseline does not matter to our results. We have included edits to reflect more clearly this choice of the baseline.

4. Planting data is one of the components for risk analysis (as one of the factors of production); what about the effect/risks of a series of other factors and their interaction on rice as well as wheat cultivation? Risk analysis based only on planting date might be misleading as there are several other biophysical and socio-economic factors that affect the production process, which are not included in this study.

Response: We have included this as a limitation of using crop models but this is not just a problem with our approach and addressing the whole realm of possible factors affecting farmers choice of planting dates is outside the scope of this paper.

5. Confusing write-up in many places and difficult to follow, mixing methodology in the introduction, see line 111-116, similarly in several places methods are included in results.

Response: These repetitions are meant to remind readers as this paper advances a methodology that is not well known and our key contribution is the approach.

6. Missing clear objective at the end of the introduction section.

Response: We have added an objective statement at the end of the first paragraph.

7. Missing logical discussion section

Response: We have added a discussion section as suggested.

8. Line 130-131: To assess the economic return from rice and wheat, simulated yield (without using properly calibrated model) multiplied by price data (interpolated using random forest from Landscape Crop Assessment Survey, Line 212-213, without validation). As IGP is a huge territory, the values vary across location and socio-economic conditions. Whole computation of economic analysis and costing without ground information, for me is very hard to believe.

Response: We thank the reviewer for the comment. We used ground truth data from the Landscape Crop Assessment Survey and interpolated using the approach that Cedrez and Chamberlain developed. We also assumed irrigation costs for each scenario. While these are approximations, the aim is to showcase how the approach can use existing data. A more comprehensive economic costing and analysis is an area of future research but given that we apply the same interpolations and costs across scenarios, a better approximation of these values will not change the final comparisons because whatever the errors of approximation are, they can nullified when making relative comparisons of the scenarios.

10. Line 231-233- is methodology, not the result.

Response: We have edited accordingly.

11. Line 231 and Table 1: Baseline should be the farmers' practice, not the government recommendation. One cannot imagine farmers using exactly the recommended planting date (and other practices), especially under highly variable rainfall conditions and also the uncertainty of input supplies.

Response: We thank the reviewer for the comment. We have chosen not to use farmer practice as the baseline because unlike the other scenarios it does not have data for all the pixels. We did robustness checks for Bihar (not reported in the paper) that showed that the results are consistent for the matching pixels. We have added edits to explain our choice.

12. No ground verification of all those obtained riskiness results.

Response: See comment above regarding the use of field level data. Our results are based on clear theory and explicit assumptions for definitions of riskiness. Datasets for validating these across the simulated area of interest and time scales do not exist. Experimental trials for long-term evaluation of different management strategies would be a useful next step for further research in the areas we identified. We thus provide useful evidence for stratification future cropping trials.