**Review #2:**

**General Comment:** High-spatial and high-temporal resolution rice yield datasets are lack especially over large regions. The manuscript employed machine learning algorithms to generate long-term high-resolution rice yield over the South Asia, Southeast Asia, and East Asia. Undergoing a study at continental scales like this is a huge project. The 5km rice yield map over the major rice producing countries in Asia from 1995 to 2015 fills the data gap for assessing the impacts of climate change and the sustainable development. However, I have a few major concerns to be addressed so that the manuscript could be more solid.

**Response to general comment:** We are grateful for anonymous referee #2's recognition of this study's importance. We carefully revised our manuscript and provided a point-by-point response below. We have addressed all points raised in the revised manuscript.

------------------------------------------------------------------------

Note: The individual comments (shown in black) are listed below including our responses (shown in blue) and revised parts in the manuscript (shown in *red and italic font*). Line numbers (shown in **blue and bold font**) that we mention in this comment refer to our revised manuscript with all markup version.

**Comment 1:** (1) The rice cultivated area is the fundamental information for rice yield estimation. The manuscript used rice map for each year from 2000 to 2020 while the yield model was developed and used to estimate spatial distribution of rice yield during 1995 to 2015. Since most input dataset used for rice yield model in the study are available for the year 2000 to 2020, why not generating rice yield for 2000 to 2020 so that the map and the rice yield coincided with each other for the same year?

**Response to comment 1:**

Thank you very much for your comments and suggestions.

Our main objective in the study is to produce a long-term rice yield dataset with higher spatiotemporal resolutions and seasonal information across Asia. Although multi-sources data were used, available rice yield was the dominant factor for time span which was essential for model training and accuracy validation. After inputting our most efforts, we can only obtain the yield records of 1995~2015 for most countries (Table S2). Therefore, the time span of this study was selected from 1995 to 2015.

**Comment 2:** (2) Another concern is the way of predictor selection. The authors selected the predictors based on the correlation analysis between indicators and the yield at each administrative unit. While this is in general logic, it might be a problem when great differences existed in cropping patterns and the rice management in an administrative unit. The correlations may fail to achieve a significant level when an improper unit was targeted. This needs more clarification. Please also specify the administrative unit. Is it national level or sub-national level administrative units?

**Response to comment 2:**

Thanks very much for your constructive comment.

The administrative unit is the sub-national level unit which is at the minimum administrative division (including first, second and third levels, Table S1) scale in this study. Differences of the cropping patterns and the rice management do exist at a national level, while those in the minimum administrative division for each country are smaller.

Besides, the selected predictors in our study consistently indicated significant relationships with yield. To make our manuscript clearer, we have added more descriptions for administrative units (**Lines 114-115**). The administrative division for each country were also listed in Table S1 according to your advice.

**Comment 3:** (3) The authors only used one vegetation indicator LAI as the inputs. It is assessed by several research that LAI products are of high uncertainty even for the improved GLASS LAI products. The product still has some abnormal values and unrealistic seasonality especially in winter. From my understanding, using LAI products might introduce high uncertainty in yield model which is unable to be solved.

**Response to comment 3:**

Thanks very much for your constructive comment.

Considering the spatial and temporal resolutions, GLASS LAI products are more appropriate for our research than other latest public products even with higher spatial resolution. GLASS LAI products have the highest accuracy and the lowest uncertainty compared with other available LAI products according to Xiao et al. (2016) and Liang et al. (2021). In addition, the abnormal values and unrealistic seasonality of LAI are always over the northern high latitudes and the equatorial belt due to cloud/snow coverage and low solar zenith angle in winter (Garrigues et al., 2008; Jin et al., 2017). However, for northern areas at high latitudes area, there is no rice planting in winter. Only some rice planting area of Malaysia and Indonesia located in the equatorial belt may be affected by these problems. Moreover, only 5 LAI variables, accounting for one-tenth of all variables, were used and preprocessed to filter abnormal pixels and those without rice growth patterns (Sect. 2.3.1) to reduce the uncertainty from LAI. In the revised manuscript, we have suggested the process for LAI data (**Lines 211-212**) and added the uncertainty of GLASS LAI in section 4.3 (**Lines 439-442**).

References:

Garrigues, S., Lacaze, R., Baret, F., Morisette, J. T., Weiss, M., Nickeson, J. E., Fernandes, R., Plummer, S., Shabanov, N. V., and Myneni, R. B.: Validation and intercomparison of global Leaf Area Index products derived from remote sensing data, J. Geophys. Res. Biogeosciences, 113, 2008.

Jin, H., Li, A., Bian, J., Nan, X., Zhao, W., Zhang, Z., and Yin, G.: Intercomparison and validation of MODIS and GLASS leaf area index (LAI) products over mountain areas: A case study in southwestern China, Int. J. Appl. Earth Obs. Geoinformation, 55, 52–67, https://doi.org/10.1016/j.jag.2016.10.008, 2017.

Liang, S., Cheng, J., Jia, K., Jiang, B., Liu, Q., Xiao, Z., Yao, Y., Yuan, W., Zhang, X., and Zhao, X.: The global land surface satellite (GLASS) product suite, Bull. Am. Meteorol. Soc., 102, E323–E337, https://doi.org/10.1175/BAMS-D-18-0341.1, 2021.

Xiao, Z., Liang, S., Wang, J., Xiang, Y., Zhao, X., and Song, J.: Long-time-series global land surface satellite leaf area index product derived from MODIS and AVHRR surface reflectance, IEEE Trans. Geosci. Remote Sens., 54, 5301–5318, https://doi.org/10.1109/TGRS.2016.2560522, 2016.

**Comment 4:** (4) According to the importance of the indicators, static indicators (Year, Lat, Long, Ele) are much higher than other indicators. For some countries, the proportioned importance of CEC+TI indicators could be higher than 90%. And for the whole study area, the CEC+TI are the most important indicators. How to explain this? Does this mean there are no need to add other indicators for yield mapping?

**Response to comment 4:**

Thank you very much for your comments and suggestions.

Only about half of the models which the importance of CEC+TI are obviously more than 50%, which suggests the importance of other indicators (CGP, EGP, CEC) still account for approximately 50%. We have attempted to estimated rice yield only by predictors of CEC+TI for some high proportioned importance and low proportioned importance cases. All the results were worse than the original models. For the high proportioned importance cases, the accuracy of only input CEC+TI predictors decreased less than that of low proportioned importance ones.

It is generally accepted that CEC+TI shows a high proportioned importance on yield estimates, which is in agreement with the findings from Huntington et al. (2020), Cao et al. (2021) and Ray et al. (2019). The three static predictor, *Lat*, *Lon*, and *Ele*, are the most basic geographical environment for rice growing and TI is used to replace the influence on rice yield of long-term agronomic technology improvements and varieties renewal because of the management data at a larger scale unavailable. Agronomic technology and varieties renewal are essential for rice yield compared with climate change and can offset the negative impacts of climate change according to the related studies (Yu et al., 2012; Ladha et al., 2021). Therefore, it is reasonable to have a high proportion for the importance of CEC+TI.

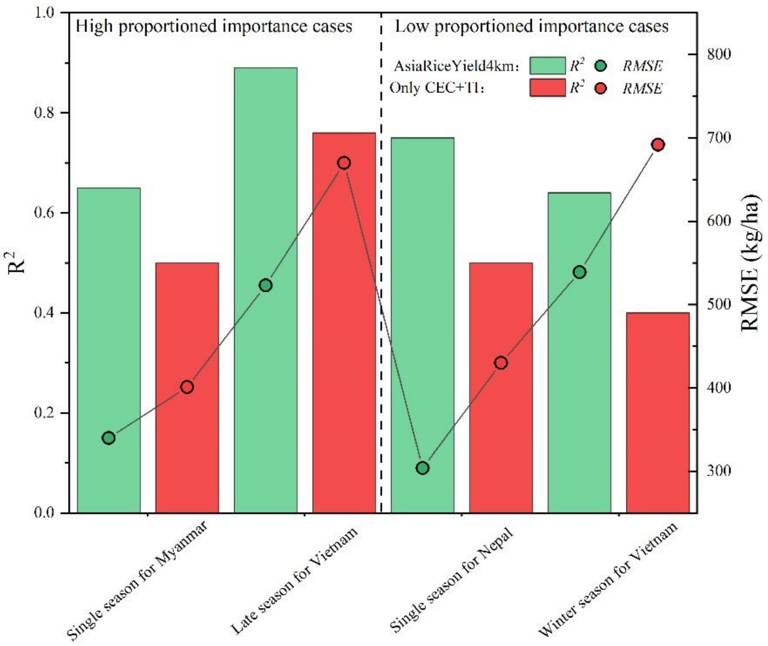


**Figure C1: Accuracy of AsiaRiceYield4km and only CEC+TI predictors for rice yield estimation.**

References:


Cao, J., Zhang, Z., Luo, Y., Zhang, L., Zhang, J., Li, Z., and Tao, F.: Wheat yield predictions at a county and field scale with deep learning, machine learning, and google earth engine, Eur. J. Agron., 123, 126204, https://doi.org/10.1016/j.eja.2020.126204, 2021.

Huntington, T., Cui, X., Mishra, U., and Scown, C. D.: Machine learning to predict biomass sorghum yields under future climate scenarios, Biofuels Bioprod. Biorefining, 14, 566–577, https://doi.org/10.1002/bbb.2087, 2020.

Ladha, J. K., Radanielson, A. M., Rutkoski, J. E., Buresh, R. J., Dobermann, A., Angeles, O., Pabuayon, I. L. B., Santos-Medellín, C., Fritsche-Neto, R., and Chivenge, P.: Steady agronomic and genetic interventions are essential for sustaining productivity in intensive rice cropping, Proc. Natl. Acad. Sci., 118, e2110807118, 2021.

Ray, D. K., West, P. C., Clark, M., Gerber, J. S., Prishchepov, A. V., and Chatterjee, S.: Climate change has likely already affected

global food production, PloS One, 14, e0217148, https://doi.org/10.1371/journal.pone.0217148, 2019.
Yu, Y., Huang, Y., and Zhang, W.: Changes in rice yields in China since 1980 associated with cultivar improvement, climate and crop management, Field Crops Res., 136, 65–75, 2012.

**Comment 5:** (5) When the model is applied for yield estimation during different growing season, does the pixel level cropping intensity map used or it is mainly based on the majority of rice cropping patterns in each administrative unit? The uncertainty of season rice yield might exceeded the uncertainty of the model due to the biased seasonal rice map.

**Response to comment 5:**

Thanks very much for your constructive comment.

In this study, we identified rice cropping intensity at the administrative scale due to the unavailable of suitable gridded rice cropping intensity maps. Although several large scale gridded cropping intensity maps were generated recently, they still cannot distinguish different crops especially for rice (Han et al., 2022; Liu et al., 2021). In this study, RiceAtlas, the most comprehensive and detailed spatial dataset on rice cropping intensity, was used. This dataset is nearly ten times more spatially details and has nearly seven times more spatial units compared with others (Laborte et al., 2017) which can reflect more explicit rice cropping intensity at administrative scale. At gridded scale, only pixels of rice area located in these minimum administrative unit with available seasonal rice yield were mapped. Besides, according to Response to comment 4 and Sect. 2.3.1, the pixels passed the inflection, and threshold detection were used for model training which suggest that these pixels are relative pure and can mitigate the uncertainty of rice seasons. We admit that the administrative rice crop intensity will introduce uncertainty for yield estimation, and such uncertainty of rice crop intensity had been added into Sect. 4.3 (**Lines 439-442**).

Reference:
Han, J., Zhang, Z., Luo, Y., Cao, J., Zhang, L., Zhuang, H., Cheng, F., Zhang, J., and Tao, F.: Annual paddy rice planting area and cropping intensity datasets and their dynamics in the Asian monsoon region from 2000 to 2020, Agric. Syst., 200, 103437, https://doi.org/10.1016/j.agsy.2022.103437, 2022.
Laborte, A. G., Gutierrez, M. A., Balanza, J. G., Saito, K., Zwart, S. J., Boschetti, M., Murty, M. V. R., Villano, L., Aunario, J. K., Reinke, R., Koo, J., Hijmans, R. J., and Nelson, A.: RiceAtlas, a spatial database of global rice calendars and production, Sci. Data, 4, 170074, https://doi.org/10.1038/sdata.2017.74, 2017.
Liu, X., Zheng, J., Yu, L., Hao, P., Chen, B., Xin, Q., Fu, H., and Gong, P.: Annual dynamic dataset of global cropping intensity from 2001 to 2019, Sci. Data, 8, 283, https://doi.org/10.1038/s41597-021-01065-9, 2021.

**Comment 6:** (6) Any possibility to use some in-situ collected actual yield data to validate the yield map?

**Response to comment 6:**

Thanks very much for your constructive suggestions.

Although it's difficult to collect in-situ yield for such a larger area, a fraction of in-situ single rice yield data is available from 1995 to 2015 in China. These data are obtained from China agro-meteorological stations, which are maintained by China Meteorological Administration (CMA) (http://data.cma.cn/). Fig. C2 presents the locations of the 47 agro-meteorological stations.

Fig. C3 shows that AsiaRiceYield4km was well consistent with in-situ yield as the average $R^2$ was

0.55 during 1995-2015. Moreover, the $R^2$ at the specific years could be as high as 0.72 at 2000, followed by 0.69 at 2005 and 0.68 at 2010. Besides, the *RMSE* was lower than 600 kg/ha at 2005, followed by 714 kg/ha at 2000 and 899 kg/ha at 2010. Therefore, the in-situ validation results were well satisfactory.

The *RMSE* for all years (Fig C3a) was somehow large (1019 kg/ha). Several reasons might cause such bias: the rice area planted at the agro-meteorological stations was generally lower than 0.015km$^2$, largely smaller than our pixel size (4×4km, 16 km$^2$). Besides, rice at agro-meteorological stations was well managed, thus such in-suit yields failed in characterizing those records at an administrative scale. Overall, the scale differences might be attributed as the main reason for the validation uncertainties.

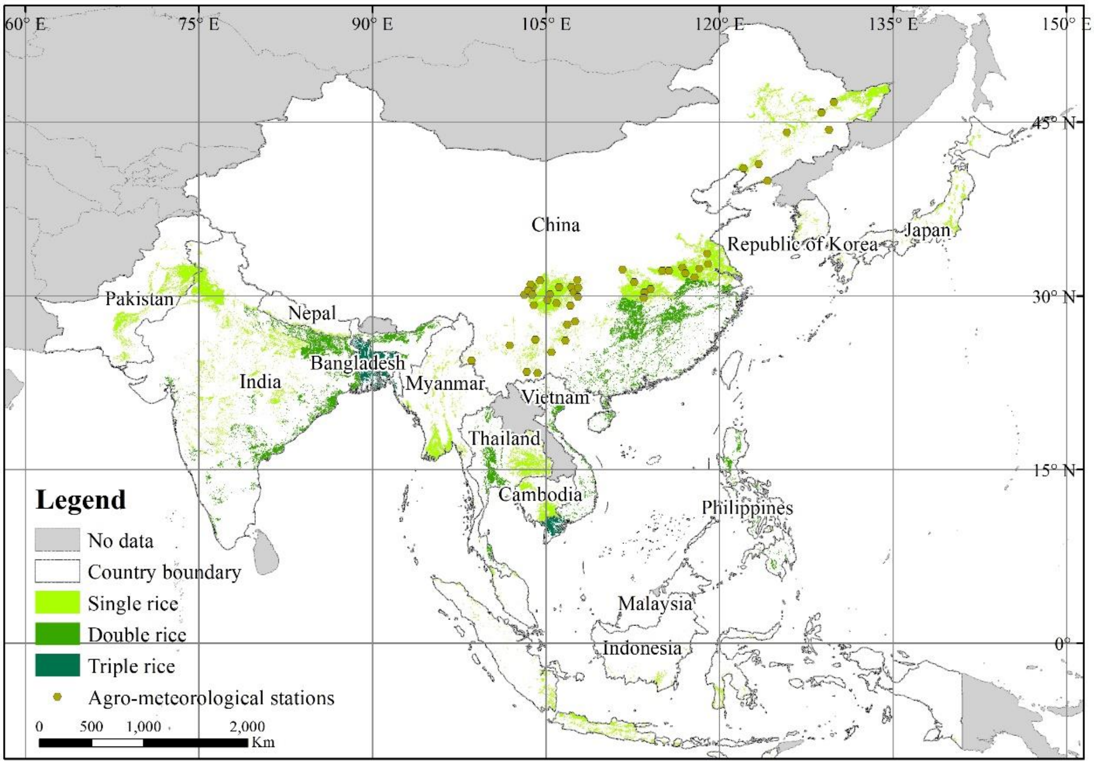


**Figure C2: Location of the selected agro-meteorological stations.**

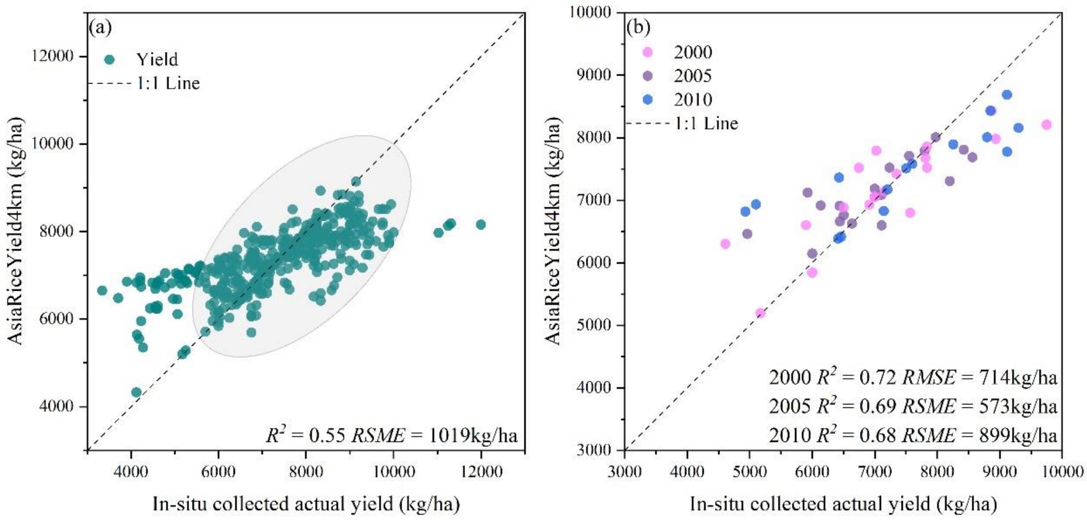

**Figure C3: (a) Accuracy between AsiaRiceYield4km and in-situ yield for all years. (b) The accuracy between AsiaRiceYield4km and in-situ collected actual yield in 2000, 2005 and 2010.**

Specific comments:

**Specific comment 1:** (1) Page 4 Line 106, what do you mean by 27 seasons?

**Response to specific comment 1:**

Thank you.

Here, 27 seasons refers to 27 different rice-cropping periods in 14 countries. However, we have to admit that "season" might confuse readers. Therefore, we've changed "season" into "case" and added the explanation for "case" (one specific rice-cropping period in a country). Relevant sentences in the manuscript were also modified.

**Specific comment 2:** (2) The authors collected many rice yield data from different sources. Please add more detailed information of the yield data including the spatial units, temporal extent, etc.

**Response to specific comment 2:**

Thanks very much for your constructive comment.

We have added detailed administrative scale and temporal information in the revised Supplement Table S1.

**Specific comment 3:** (3) Page 10, Line 229 – 234, the dataset was first divided into two parts according to the administrative units. 80% of the administrative units were randomly selected as training and validation among which 70% of samples were used for training and 30% were used as validation sets. In this case, the training samples were not 56% of the whole dataset. Same for validation and testing. Please make it more clear for readers.

**Response to specific comment 3:**

Thanks very much for your constructive comment.

The dataset division in the original manuscript is incomprehensible. For each case, 20% of the subset was selected randomly by administrative units. The rest of the 80% was split into 70% for training and 30% for validation. One sample is one administrative unit in one year with several predictors. Therefore, the training, validation, and testing dataset were 56% (80%×70%), 24% (80%×30%), and 20% (20%×100%) of the whole data, respectively (Fig. C4). The original flowchart was misleading. We have reorganized the expression of the dataset division (**Lines 241-245**) and redrawn the flowchart of the dataset division (Fig. 2 step3).

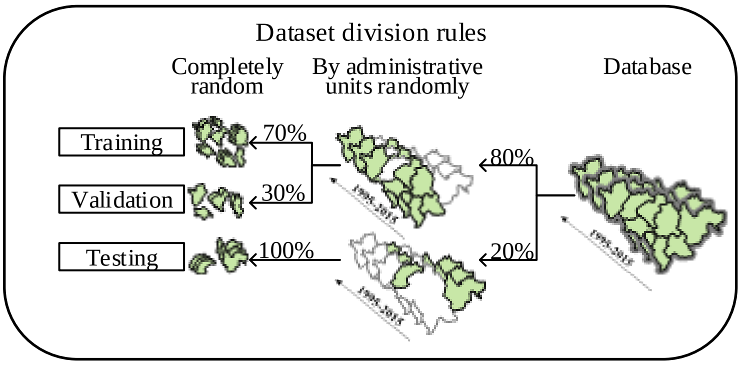


**Figure C4: Dataset division rules**

**Specific comment 4:** (4) Add more testing results for other years. The authors estimated rice yield for Asia for 1995-2015 but was insufficiently validated and tested for different years. Also, the temporal changes of rice yield should be added to result and discussion sections.

**Response to specific comment 4:**

Thanks very much for your constructive comment.

We have added the temporal comparison of AsiaRiceYield4km and observed yields in Sect. 3.2 **(Lines 338 to 345)** for all years to validate the temporal accuracy of our results. Moreover, temporal variation analysis of rice yield was also included in Sect. 3.4 **(Lines 381-387)**.