Response to reviewer #3:

Many thanks for your thoughtful and valuable comments and suggestions, which are very helpful in improving our manuscript. We have conducted substantial new experiments and analyses to ensure that the study is more comprehensive and rigorous, and our maps are more reliable. Our responses to the comments point-by-point are included below in blue. The corresponding changes in the revised manuscript are shown in purple.

General comment: This manuscript developed a phenological- and pixel-based soybean area mapping (PPS) method to identify soybean on a large scale and generated a dataset of soybean planting areas across China. The topic is significant for sustainable soybean production and management. However, the proposed methodology lacks notable innovation when compared to prior studies. Given the intricate spectral variations within soybeans and the fragmentated nature of agricultural landscapes across China, the presented method fails to demonstrate its robustness across diverse regions and time periods, therefore raising concerns about the reliability of the resulting soybean map. Furthermore, certain descriptions of the proposed method lack essential details and specific contents are not easy to follow. Below, I have provided several detailed comments:

Reply: Thank you for the expert questions, which do encourage us to comb our method and add more detailed experiments to demonstrate the robustness of our approach and the reliability of the resulting soybean maps. We explained all these details by three sections (innovation, accuracy, and methods' details).

(1) Innovation: We firstly obtained the characteristic spectra and growth curves of soybean in different areas during the key observed growth periods, and then trained local unsupervised classifiers to self-adapt to cross-regional growth variability, which have avoided huge requirement for extensive ground samples. The Regional Adaptation Spectra-Phenology Integration (RASP) framework proposed is novel and can be repeatable into other major areas planted by soybean with simple inputs, providing a solution for mapping annual soybean dynamics with a higher resolution (please see the details in the revised method section of the new edition from line 231-312 in page 11-15).

(2) Accuracy: In our results of accuracy assessment, we have used so many data available from different years to verify the reliability of our soybean maps by several methods, including visual comparison, comparing soybean areas retrieved with county- and prefecture- statistical books, and point verification with confusion matrix separately by sub-zone. Very few previous studies have assessed comprehensively maps' accuracy by the three methods. To demonstrate the robustness of mapping across diverse regions, we added the accuracy results of statistical data verification and point verification in each sub-zone.

- **The variations in accuracy among sub-zones based on statistics validation:**

"The mapping accuracy in Zone I closely matched county-level statistics, showing high consistency (R^2=0.86). Zones II-IV also demonstrated reasonable agreement (R^2=0.50~0.69), despite relatively lower accuracy due to the scarcer planted areas (Fig. S5). No significant trend deviation from statistics was indicated for the mapping area in Zone I, with slight overestimates for Zone II and III, and underestimates for Zone IV (Fig. S5). These accuracy variations are acceptable, given the challenges in accurately identifying soybeans in regions where they are planted less prevalently. Specifically, maize is more dominant than soybeans in Zone II, while Zone III is characterized by diverse crops and complex planting patterns. Underestimation in Zone IV is possibly due to fewer clear observations in the southwest.

Nevertheless, the overall accuracy across the zones is acceptable."

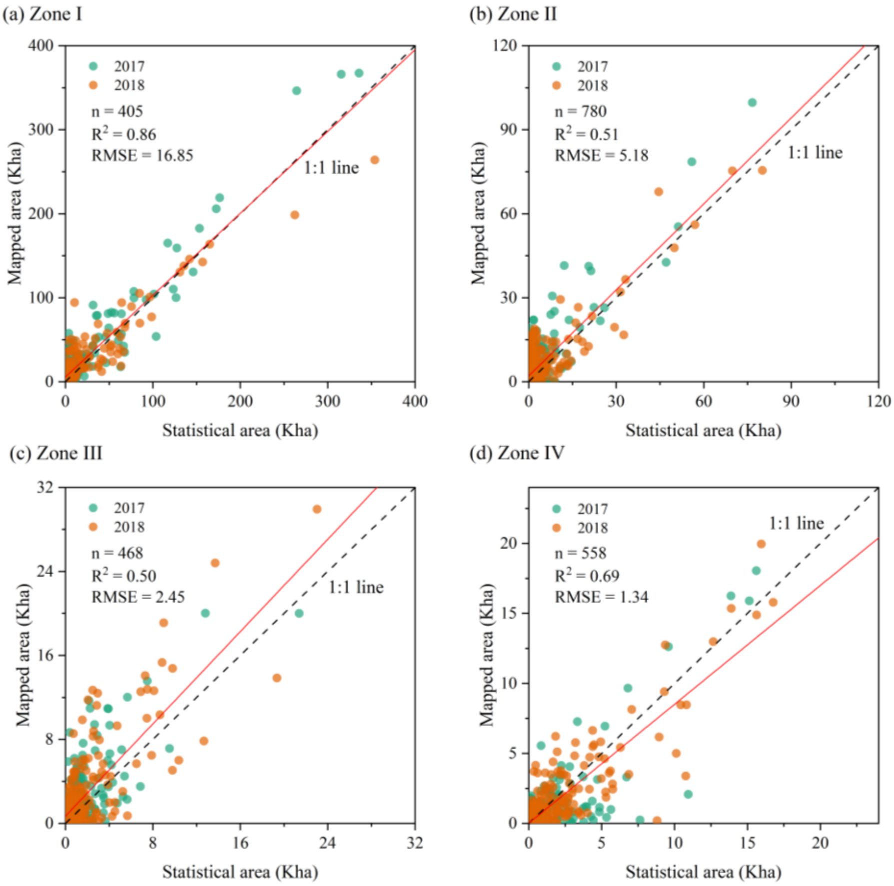

Figure S5. Comparison of soybean areas with county-level statistics in (a) Zone I, (b) Zone II, (c) Zone III, and (d) Zone IV in 2017 and 2018.

➤ The variations in accuracy among sub-zones based on samples validation:

"The overall accuracy for each sub-zone in 2019 varied from 83.58% to 90.67% (Table S1). Specifically, Zone I demonstrated the highest producer's accuracy for soybean at 88.31%, aligning with its high consistency with statistics. Zone III achieved the highest overall accuracy at 90.67%, attributed to its superior user's accuracy for soybean, indicating fewer misclassifications, and effective differentiation from non-soybean crops (Table S1). The producer's accuracy in Zone IV was relatively lower at 63.89%, possibly due to the limited samples, high heterogeneity, and fewer clear observations (Table S1)."

Table S1. Confusion matrix of the soybean maps in each sub-zone in 2019.

	Reference	Map		Producer's	User's	F1	Overall
		Soybean	Non-Soybean	Accuracy	Accuracy	Score	Accuracy
I	Soybean	922	122	88.31%	81.09%	0.85	87.12%
	Non-Soybean	215	1358	86.33%	91.76%	0.89	
II	Soybean	233	74	75.90%	86.30%	0.81	83.58%
	Non-Soybean	37	332	89.97%	81.77%	0.86	
III	Soybean	101	26	79.53%	98.06%	0.88	90.67%
	Non-Soybean	2	171	98.84%	86.80%	0.92	
IV	Soybean	23	13	63.89%	92.00%	0.75	87.18%
	Non-Soybean	2	79	97.53%	85.87%	0.91	

(3) Method descriptions in details: We have added the necessary details to the method we proposed in method section of our MS. To fully and positively respond all your valuable and suggestive comments, we also further listed them point by point in the follows.

Specific comments:

Comment 1: Line83-93: The expression and logic are not clear. I suggest that the authors reorganize "method (5)" to emphasize its key theory, advantages and disadvantages. Additionally, Line93-98 should be revised to describe the fundamental theory and performance of those method proposed by prior researchers. Furthermore, in Introduction section, the authors didn't introduce the fundamental concept behind the proposed method, nor highlighted the current issues faced by previous efforts in large-scale soybean mapping.

Reply: Thank you for your instructive suggestion. We have followed you to modify all these sentences thoroughly in the Introduction section:

(1) We have reorganized the previous researches and divided the remote sensing-based crop classification methods used widely into four categories. Method 5 in the original text has been incorporated into supervised classification. Additionally, we revised the corresponding section, as well as discussing the advantages and disadvantages of each method:

"Mapping crops by remote sensing can be categorized into four methods : 1) supervision classification based on a large number of field samples or high quality training labels (Song et al., 2017; You et al., 2021; Shangguan et al., 2022; Li et al., 2023); 2) developing some composite indexes based on the feature bands and determining the binary classification using appropriate thresholds (Huang et al., 2022; Chen et al., 2023; Zhou et al., 2023); 3) threshold segmentation based on prior knowledge such as phenology or spectra (Zhong et al., 2016); 4) combining unsupervised classification with cluster assignment (Wang et al., 2019; You et al., 2023). Supervision classification methods relied on ground samples heavily, while the 2nd and 3rd methods are both based on reliable and accurate thresholds. However, mapping soybean by these methods was mainly applied in small areas, very few covering over a larger region. Because of sufficient field samples, supervision classification can achieve maps with a higher accuracy, which is relatively mature method used widely. However, collecting sufficient field samples is extremely time, money, and labor consumed, and unsuitable for long-term years over larger areas (Luo et al., 2022). Furthermore, the threshold-based methods (the 2nd and 3rd) have been applied into large areas, however, determining the thresholds will inevitably bring significant uncertainty, especially for the areas with high heterogeneity in climate, environment, and planting patterns. Thus, these methods show low reproducibility, further hindering their application across diverse geographic areas. As for mapping soybean, it is still a big challenge due to their similar growth characteristics with many other summer crops (Wang et al., 2020; Di Tommaso et al., 2021). The thresholds that work well in some areas did not perform well in other areas (Graesser and Ramankutty, 2017; Guo et al., 2018). These limitations restrict accurate soybean maps available, especially over large regions in China."

(2) We have described the fundamental theory and the performance of prior researches mentioned in the original line 93-98 as you suggested:

"For example, the phenological observations at the agricultural meteorological stations were employed as a reference to detect the critical phenological dates of pixels through inflexion- and threshold-based methods, thereby generating planting areas for three major crops in China with R^2 greater than 0.8 compared to county statistics (Luo et al., 2020). The time-weighted dynamic time warping method based on the similarity of phenological curves of Normalized Difference Vegetation Index (NDVI) has successfully estimated the planting area of maize in China, with provincial averages for producer's and user's accuracies at 0.76 and 0.82, respectively (Shen et al., 2022). Phenological-based Vertical transmit Horizontal receive (VH) polarized time series accurately captured temporal characteristics of soybeans, thus were used for an unsupervised classifier to map the seasonal soybeans, achieving an overall accuracy over 80% in Ujjain district (Kumari et al., 2019)."

(3) We have highlighted the issues faced by previous studies in large regional soybean mapping and supplement the theoretical basis of our approach proposed:

"... These limitations restrict accurate soybean maps available, especially over large regions in China. Given the challenges of collecting sufficient field samples over larger region and the limited adaptability to environmental variations of threshold-based method, previous researches have yet to achieve multi-year, high-resolution soybean maps nationwide."

...

"... By integrating unsupervised classification's regional scalability with specific local soybean growth signs from phenological data, we fully leverage soybean's characteristic spectra and vegetation indices during key growth periods across different areas. Through training the local unsupervised classifier to accommodate the crop growth variability across regions, and avoiding extensive jobs on collecting samples, the approach provides an effective solution for regional adaptive large-area crop mapping."

Reference:

Chen, H., Li, H., Liu, Z., Zhang, C., Zhang, S., and Atkinson, P. M.: A novel Greenness and Water Content Composite Index (GWCCI) for soybean mapping from single remotely sensed multispectral images, Remote Sens. Environ., 295, 113679, https://doi.org/10.1016/j.rse.2023.113679, 2023.

Di Tommaso, S., Wang, S., and Lobell, D. B.: Combining GEDI and Sentinel-2 for wall-to-wall mapping of tall and short crops, Environ. Res. Lett., 16, 125002, https://doi.org/10.1088/1748-9326/ac358c, 2021.

Graesser, J. and Ramankutty, N.: Detection of cropland field parcels from Landsat imagery, Remote Sens. Environ., 201, 165–180, https://doi.org/10.1016/j.rse.2017.08.027, 2017.

Guo, W., Ren, J., Liu, X., Chen, Z., Wu, S., and Pan, H.: Winter wheat mapping with globally optimized threshold under total quantity constraint of statistical data, J. Remote Sens., 22, 1023–1041, https://doi.org/10.11834/jrs.20187468, 2018.

Huang, Y., Qiu, B., Chen, C., Zhu, X., Wu, W., Jiang, F., Lin, D., and Peng, Y.: Automated soybean mapping based on canopy water content and chlorophyll content using Sentinel-2 images, Int. J. Appl. Earth Obs. Geoinformation, 109, 102801, https://doi.org/10.1016/j.jag.2022.102801, 2022.

Kumari, M., Murthy, C. S., Pandey, V., and Bairagi, G. D.: Soybean Cropland Mapping Using Multi-Temporal Sentinel-1 Data, Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci., XLII-3-W6, 109–114, https://doi.org/10.5194/isprs-archives-XLII-3-W6-109-2019, 2019.

Li, H., Song, X.-P., Hansen, M. C., Becker-Reshef, I., Adusei, B., Pickering, J., Wang, L., Wang, L., Lin, Z., Zalles, V., Potapov, P., Stehman, S. V., and Justice, C.: Development of a 10-m resolution maize and soybean map over China: Matching satellite-based crop classification with sample-based area estimation, Remote Sens. Environ., 294, 113623, https://doi.org/10.1016/j.rse.2023.113623, 2023.

Luo, Y., Zhang, Z., Li, Z., Chen, Y., Zhang, L., Cao, J., and Tao, F.: Identifying the spatiotemporal changes of annual harvesting areas for three staple crops in China by integrating multi-data sources, Environ. Res. Lett., 15, 074003, https://doi.org/10.1088/1748-9326/ab80f0, 2020.

Luo, Y., Zhang, Z., Zhang, L., Han, J., Cao, J., and Zhang, J.: Developing High-Resolution Crop Maps for Major Crops in the

European Union Based on Transductive Transfer Learning and Limited Ground Data, Remote Sens., 14, 1809, https://doi.org/10.3390/rs14081809, 2022.

Shangguan, Y., Li, X., Lin, Y., Deng, J., and Yu, L.: Mapping spatial-temporal nationwide soybean planting area in Argentina using Google Earth Engine, Int. J. Remote Sens., 43, 1724–1748, https://doi.org/10.1080/01431161.2022.2049913, 2022.

Shen, R., Dong, J., Yuan, W., Han, W., Ye, T., and Zhao, W.: A 30 m Resolution Distribution Map of Maize for China Based on Landsat and Sentinel Images, J. Remote Sens., 2022, 2022/9846712, https://doi.org/10.34133/2022/9846712, 2022.

Song, X.-P., Potapov, P. V., Krylov, A., King, L., Di Bella, C. M., Hudson, A., Khan, A., Adusei, B., Stehman, S. V., and Hansen, M. C.: National-scale soybean mapping and area estimation in the United States using medium resolution satellite imagery and field survey, Remote Sens. Environ., 190, 383–395, https://doi.org/10.1016/j.rse.2017.01.008, 2017.

Wang, S., Azzari, G., and Lobell, D. B.: Crop type mapping without field-level labels: Random forest transfer and unsupervised clustering techniques, Remote Sens. Environ., 222, 303–317, https://doi.org/10.1016/j.rse.2018.12.026, 2019.

Wang, S., Di Tommaso, S., Deines, J. M., and Lobell, D. B.: Mapping twenty years of corn and soybean across the US Midwest using the Landsat archive, Sci. Data, 7, 307, https://doi.org/10.1038/s41597-020-00646-4, 2020.

You, N., Dong, J., Huang, J., Du, G., Zhang, G., He, Y., Yang, T., Di, Y., and Xiao, X.: The 10-m crop type maps in Northeast China during 2017–2019, Sci. Data, 8, 41, https://doi.org/10.1038/s41597-021-00827-9, 2021.

You, N., Dong, J., Li, J., Huang, J., and Jin, Z.: Rapid early-season maize mapping without crop labels, Remote Sens. Environ., 290, 113496, https://doi.org/10.1016/j.rse.2023.113496, 2023.

Zhong, L., Hu, L., Yu, L., Gong, P., and Biging, G. S.: Automated mapping of soybean and corn using phenology, ISPRS J. Photogramm. Remote Sens., 119, 151–164, https://doi.org/10.1016/j.isprsjprs.2016.05.014, 2016.

Zhou, W., Wei, H., Chen, Y., Zhang, X., Hu, J., Cai, Z., Yang, J., Hu, Q., Xiong, H., Yin, G., and Xu, B.: Monitoring intra-annual and interannual variability in spatial distribution of plastic-mulched citrus in cloudy and rainy areas using multisource remote sensing data, Eur. J. Agron., 151, 126981, https://doi.org/10.1016/j.eja.2023.126981, 2023.

Comment 2: Fig.1 shows that there are more soybean agrometeorological observation stations in Jiangxi Province than in Sichuan Province. So, why does the study area not include regions in South China, especially prefectures in the Jiangxi Province?

Reply: Yes, more soybean AMSs are located in Jiangxi Province, but we did not retrieve soybean areas there because of the quality limitations of Sentinel images available nowadays. Moreover, soybean planted in Southern China are generally scattered in fragmented and more complicated fields. It will be a very big challenge for smoothly selecting the specific features of a certain minor crop among many dominant crops. We excluded Southern China, including Jiangxi province, considering the above difficulties and their minor roles relative to the overall soybean production in China. All our responses to this comment are showed specifically as follows:

(1) According to the provincial statistics, the soybean planting area of the top 13 provinces accounts for over 90% of the whole national production, with only below 10% from other provinces. Therefore, despite of phenological observations available, we excluded the province from our analyses because of their minimal contribution.

(2) In Southern China, soybeans can be cultivated in multiple patterns, including double, triple, or even year-round cropping (Wang and Gai, 2002). Moreover, such cropping patterns are characterized by different intercropping and cropping rotation between soybean and other crops. Thus, how and when soybean is planted there are decided by local farmers optionally. This means that the growth phases of soybeans are inconsistent, consequently the standard curves are very hard to identify. Moreover, phenological data from local AMS could not be representative, and can't reveal local reality. The larger complexity in cropping patterns and more inputs required for

accurately identifying soybean, therefore, make us exclude the Southern China from our studied areas.

Reference:

Wang, Y. and Gai, J.: Study on the ecological regions of soybean in China II · Ecological environment and representative varieties, Chinese Journal of Applied Ecology, 71–75, 2002.

Comment 3: As stated in lines 149-151, the regions chosen to validate the classification results didn't include samples from fragmented planting regions with small soybean cultivation areas. Could this validation approach potentially lead to an overestimation of the overall validation accuracy? Additionally, there is a lack of a spatial distribution map for these field samples.

Reply: Many thanks for this valuable suggestion. We have followed you to include more samples (2019 soybean and other crop reference points from the openly products available, including fragmented planting regions with small soybean cultivation areas) to validate the accuracy of our soybean maps in other growing regions.

In provinces without ground samples, we manually selected reference points on large soybean plots based on GLAD (https://glad.earthengine.app/view/china-crop-map, last access: March 2024) soybean layer. The criterions selected are: (1) located in large plots; (2) false color composite image (R: NIR, G: SWIR2, B: SWIR1) at the peak of growing season (Song et al., 2017; You and Dong, 2020); (3) phenological characteristics similar to local observations. Additionally, the reference points of maize, single-cropping rice and double-cropping rice in 2019 were selected based on GLAD maize layer, high resolution single-season rice map (https://doi.org/10.57760/sciencedb.06963, last access: March 2024), and double-season rice map (https://doi.org/10.12199/nesdc.ecodb.rs.2022.012, last access: March 2024) with the same principle to explore the spectral characteristics of crops in each sub-zone of the studied areas. The overall accuracy of all available maps in 2019 is above 85% (Pan et al., 2021; Li et al., 2023; Shen et al., 2023).

We plotted the spatial distribution of ground samples and reference points as showed by Figure 1 below and modified Figure 1 in the edited MS. We have added the details of the reference points to the data section in revised manuscript.

Reference:

Pan, B., Zheng, Y., Shen, R., Ye, T., Zhao, W., Dong, J., Ma, H., and Yuan, W.: High Resolution Distribution Dataset of Double-Season Paddy Rice in China, Remote Sens., 13, 4609, https://doi.org/10.3390/rs13224609, 2021.

Li, H., Song, X.-P., Hansen, M. C., Becker-Reshef, I., Adusei, B., Pickering, J., Wang, L., Wang, L., Lin, Z., Zalles, V., Potapov, P., Stehman, S. V., and Justice, C.: Development of a 10-m resolution maize and soybean map over China: Matching satellite-based crop classification with sample-based area estimation, Remote Sens. Environ., 294, 113623, https://doi.org/10.1016/j.rse.2023.113623, 2023.

Shen, R., Pan, B., Peng, Q., Dong, J., Chen, X., Zhang, X., Ye, T., Huang, J., and Yuan, W.: High-resolution distribution maps of single-season rice in China from 2017 to 2022, Earth Syst. Sci. Data, 15, 3203–3222, https://doi.org/10.5194/essd-15-3203-2023, 2023.

Song, X.-P., Potapov, P. V., Krylov, A., King, L., Di Bella, C. M., Hudson, A., Khan, A., Adusei, B., Stehman, S. V., and Hansen, M. C.: National-scale soybean mapping and area estimation in the United States using medium resolution satellite imagery and field survey, Remote Sens. Environ., 190, 383–395, https://doi.org/10.1016/j.rse.2017.01.008, 2017.

You, N. and Dong, J.: Examining earliest identifiable timing of crops using all available Sentinel 1/2 imagery and Google Earth

Engine, ISPRS J. Photogramm. Remote Sens., 161, 109–123, https://doi.org/10.1016/j.isprsjprs.2020.01.001, 2020.

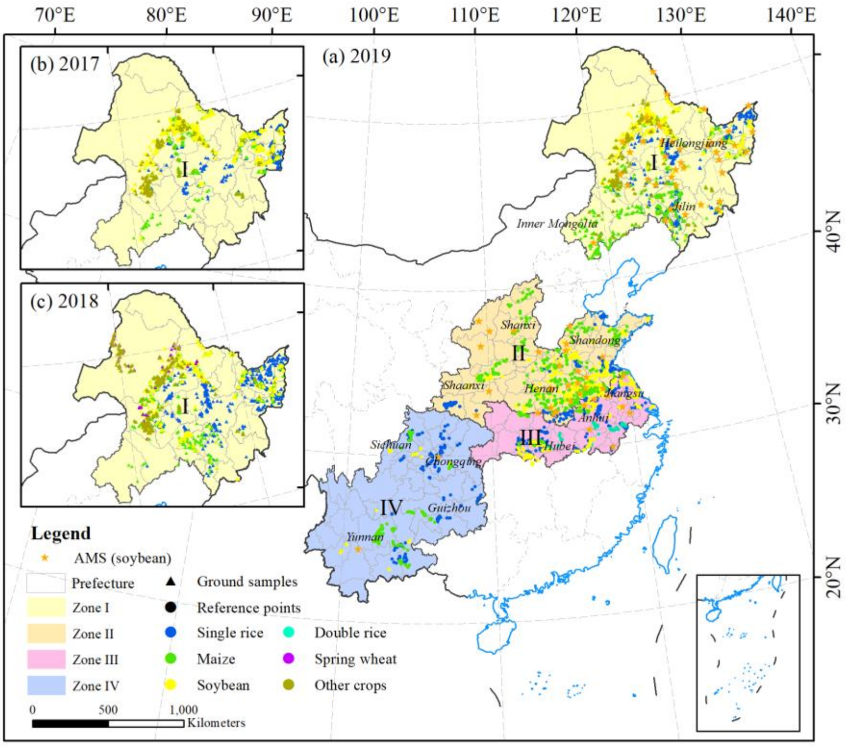

Figure 1. The study area including 14 provinces (including Chongqing Municipality) and spatial distribution of ground samples and reference points across China in (a) 2019, (b) 2017, and (c) 2018.

In addition, we updated the point validation results (Table 2) and inserted the validation for each sub-zone to the supplemental material (Table S1 above). We further explain the differences in accuracy between regions in the revised manuscript (see Reply for General comment above).

Table 2. Confusion matrix of the soybean maps during 2017-2019.

	Reference	Map		Producer's	User's	F1	Overall
		Soybean	Non-Soybean	Accuracy	Accuracy	Score	Accuracy
2017	Soybean	679	352	65.86%	72.47%	0.69	77.08%
	Non-Soybean	258	1372	84.17%	79.58%	0.82	
2018	Soybean	799	246	76.46%	74.19%	0.75	85.16%
	Non-Soybean	278	2208	88.82%	89.98%	0.89	
2019*	Soybean	1279	235	84.48%	83.32%	0.84	86.77%
	Non-Soybean	256	1940	88.34%	89.20%	0.89	

* Including ground samples and nationwide reference points based on existing datasets.

Comment 4: L206-207: References are needed to support these statements.

Reply: Thank you for your suggestion. The selection of threshold values is based on our analysis of ground samples and previous studies. We have supplemented the references here.

"Based on the cropland extracted, we filtered out the pixels exhibiting an Enhanced Vegetation Index (EVI) maximum value during the growing season less than 0.4 to remove fallow land according to the analysis of ground samples (Fig. S1) and previous studies, which found that almost all crops had maximum EVI values above 0.4 (Li et al., 2014; Zhang et al., 2017; Han et al., 2022)."

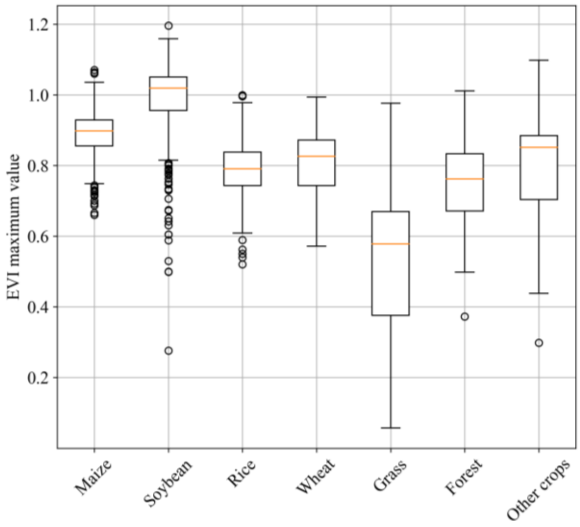

Figure S1. Box plot of the EVI maximum in 2019 based on all ground samples.

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.

Li, L., Friedl, M. A., Xin, Q., Gray, J., Pan, Y., and Frolking, S.: Mapping Crop Cycles in China Using MODIS-EVI Time Series, Remote Sens., 6, 2473–2493, https://doi.org/10.3390/rs6032473, 2014.

Zhang, G., Xiao, X., Biradar, C. M., Dong, J., Qin, Y., Menarguez, M. A., Zhou, Y., Zhang, Y., Jin, C., Wang, J., Doughty, R. B., Ding, M., and Moore, B.: Spatiotemporal patterns of paddy rice croplands in China and India from 2000 to 2015, Sci. Total Environ., 579, 82–92, https://doi.org/10.1016/j.scitotenv.2016.10.223, 2017.

Comment 5: The main crop types and cropping intensity vary across regions with different climate conditions. However, Fig.3 (a-i) only presents spectral curves for soybean planting in Northern China. Are the phenological characteristics described in "(2) Feature selection" also applicable to soybeans planted in Southwestern China? I suggest that the authors also provide spectral curves of soybean and main crops planted in South China.

Reply: Yes, the main crop types and cropping intensity do vary across regions with different climate conditions. Using the reference points described in Reply for Comment 3, we explored the spectral and vegetation indices characteristics of major crops in each region. All these selected crops grow the similar season as those of soybeans, which further are proved by the temporal consistent profiles across different sub-zones (Fig. S2-S4). We found notable differences in SWIR1, SWIR2, and SIWSI indices between soybean and rice during the early growth period. In mid and late growth phases, EVI, NIR, Red Edge2 and Red Edge3 values of soybean fields were significantly higher than other crops. The consistent differences are basis mentioned in the feature selection section, which further substantiate that the selected features can be applicable and potentially repeatable into various regions. We have added the following figure S2-S4 to the supplementary materials and stated in the revised manuscript:

"<u>All these spectral-phenological characteristics are also applicable to soybeans planted in other sub-zones (Fig. S2-S4).</u>"

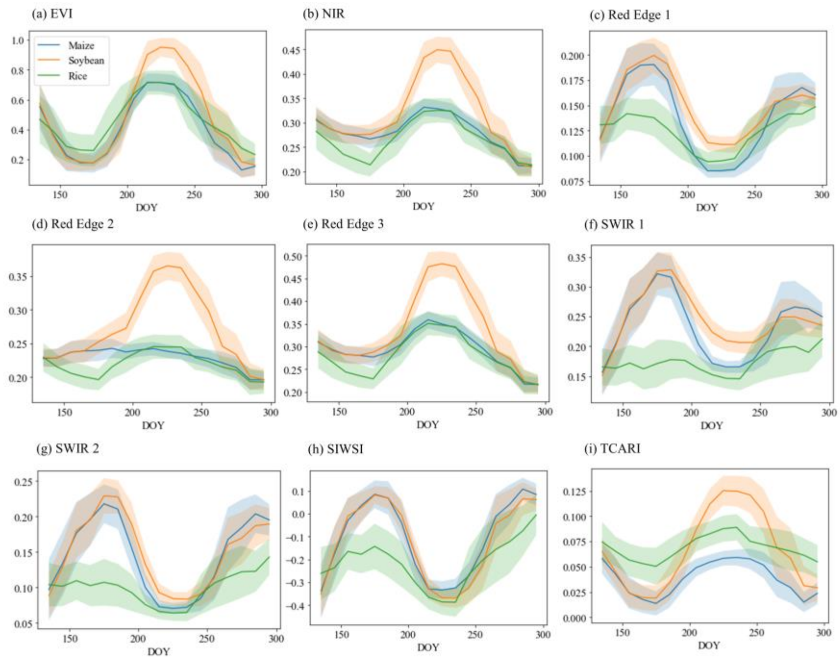

Figure S2. Temporal profiles of (a-i) for major crops in Zone II.

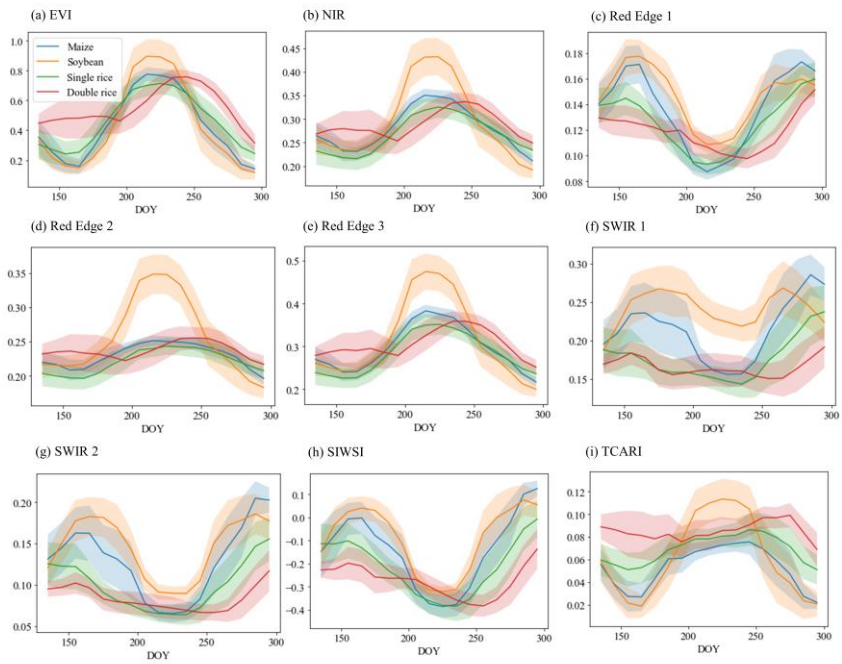

Figure S3. Temporal profiles of (a-i) for major crops in Zone III.

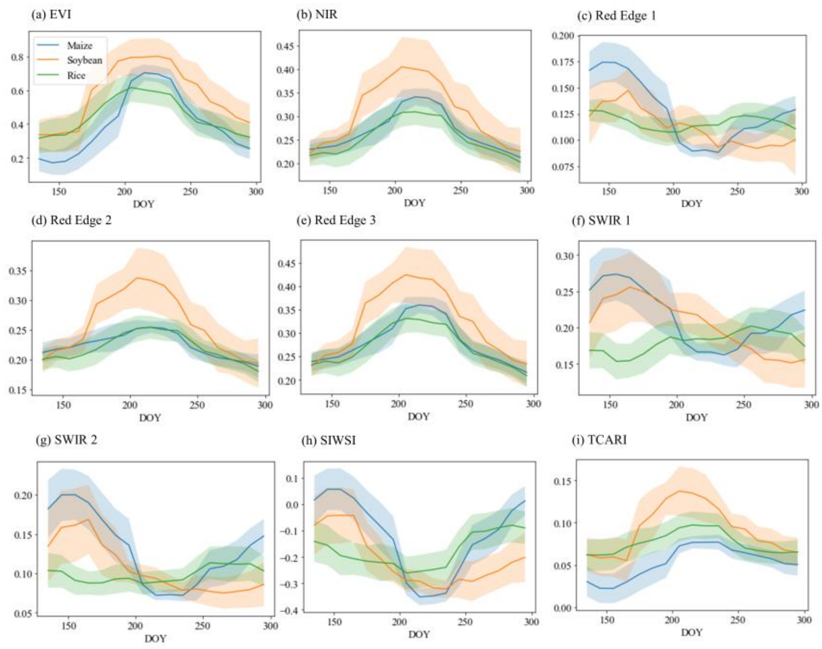

Figure S4. Temporal profiles of (a-i) for major crops in Zone IV.

Comment 6: The authors need provide example figures illustrating the result of "time window from 15 days before the podding date (DOYpodding) to 15 days after the full-seed date (DOYseed)"

Reply: Thank you for your insight suggestion. We have followed you to plot example figures (illustrating the result of "time window from 15 days before $DOY_{podding}$ to 15 days after DOY_{seed}) for identifying seasonal crops under single and double cropping patterns (Figure 2). We confined the peak value detected in soybean growing period to ensure the rationality of our method in the single or double cropping systems. We have added the figure into the revised manuscript.

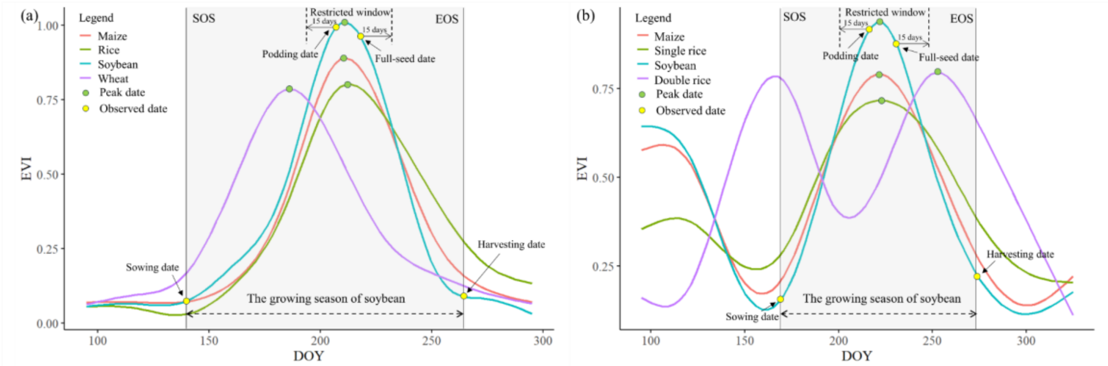

Figure 2. Schematic diagram of seasonal crop identification for (a) single - and (b) double - cropping systems.

Comment 7: L241-242: these contents are confusing, is there any typo?

Reply: We are so sorry for the confusion expression. We have revised it to the follow:

"Meanwhile, the timing of TCARI reaching saturation significantly differs among soybean, rice, and wheat (Fig. 4i)."

Comment 8: L255-256：How did you determine the number of K-mean clusters based on statistics? Further explanation is needed for clarity.

Reply: We collected the statistical area for seasonal crops (including rice, maize, soybean, cotton, peanuts, sesame, sweet potato, and sorghum) of each prefecture in 2018. We defined "major crops" as those species cumulatively representing 95% of the total seasonal crop area, with an additional category for all "other crops" to determine the number of clusters *k*. We have added the process determining the number of K-mean cluster into section 2.3.2 "(3) Unsupervised learning":

"The classifier was trained individually on each prefecture based on the number of clusters *k* input. The cluster number *k* is defined as the number of "major crops" that constituting 95% of the total area for seasonal crops (including rice, maize, soybean, cotton, peanuts, sesame, sweet potato, and sorghum) according to prefecture-level statistics, and plus one for "other crops"."

Comment 9: The DTW step is not clearly described:

(1) I wonder whether the length and time coverage of S2 time series used for calculating DTW distance vary across different AEZs?

Reply: Yes, the length and time coverage of S2 time series is different for each sub-zone. According to the soybean sowing and maturity dates recorded at AMSs, we set the time coverage of Zone I-IV to April-September, May-October, June-October, and August-November, respectively. This selection of time spans ensures that the full growing season of soybeans is included in each sub-zone.

(2) Did the authors use averaged time series for 100 random points and those for all field samples around the whole China to calculate DTW distances? If so, it is important to note that the spectral differences between crops in North and South China may affect the validity of DTW calculation results. Have you considered the impact of intra-class spectral differences in soybean samples from different regions on the DTW calculation results and the final classification results?

Reply: Yes, the DTW distances are key parameters for distinguishing soybean from other crops. We determined the standard time series for each sub-zone separately. We randomly selected the 30% sample points (Dong et al., 2020) in each sub-zone and calculated the averages to determine the soybean standard curves, since the soybean growth periods and their related curves in same sub-zone do not differ hugely. For the classification results in each prefecture, we randomly select 100 points to calculate the averages and determine their standard curves for all crop category, and separately calculate the DTW distance of standard curves between the soybean and all crops.

As for your worry about the spectral differences between crops in North and South, our method proposed will not impact the DTW (calculated in a prefecture) validity because of their weak difference among a prefecture. Similarly for the intra-class spectral differences for soybean samples of different regions, such differences do not particularly impact DTW values and the final classification results because of soybean standard curves developed respectively in each sub-zone.

Reference:

Dong, J., Fu, Y., Wang, J., Tian, H., Fu, S., Niu, Z., Han, W., Zheng, Y., Huang, J., and Yuan, W.: Early-season mapping of winter wheat in China based on Landsat and Sentinel images, Earth Syst. Sci. Data, 12, 3081–3095, https://doi.org/10.5194/essd-12-3081-2020, 2020.

(3) Line 219-221: Did the authors use all the above 8 feature to calculate DTW distances? How did you integrated the 8 DTW distances into the final DTW value used for classification?

Reply: Yes, we calculated the DTW distances for these 8 features. The averaged DTW distance for all features was used to assess the similarity degree with the standard curve.

(4) "The cluster closest to the samples was identified as the soybean cluster." How did you determine the threshold?

Reply: We did not use a threshold here. Based on the DTW distance of their standard curves for each crop category and soybean, the cluster with the minimum distance among all categories is selected as soybean. Taking into account of all above questions you provided, we updated the method details in the Cluster assignment section in the revised manuscript:

"We then used dynamic time warping (DTW) method to measure the similarity between each cluster's eight features involved in classification and the soybean standard curves. We averaged the data of 30% samples in each sub-zone to establish the standard curves, reducing the impact of regional phenological variations. The time coverage of Zone I-IV was set to April-September, May-October, June-October, and August-November, respectively, which are corresponding with the soybean growing season. The cluster with the minimal average DTW value was identified as the soybean cluster."

Comment 10: Fig.8 (a1-3) depict false-color composite images composed of bands 4, 3, and 2. Distinguishing between soybeans and non-soybeans in these images is visually difficult. It is recommended to present images composited with other bands. The authors can refer to the following article, which uses the shortwave infrared band for false-color compositing.

Song X-P, Potapov P V, Krylov A, King L, Di Bella C M, Hudson A, Khan A, Adusei B, Stehman S V,Hansen M C. National-scale soybean mapping and area estimation in the United States using medium resolution satellite imagery and field survey. Remote Sens. Environ., 2017, 190: 383-395

You N,Dong J. Examining earliest identifiable timing of crops using all available Sentinel 1/2 imagery and Google Earth Engine. ISPRS-J. Photogramm. Remote Sens., 2020, 161: 109-123

Reply: Thank you for your advice. We updated false color composite images (R: NIR, G: SWIR2, B: SWIR1) to identify soybean plots more clearly. The reflectance differences between soybean and other crops in these bands do be greater than that in red, green and blue bands. Many thanks for your expert advice, which really encourage us to deepen our study！

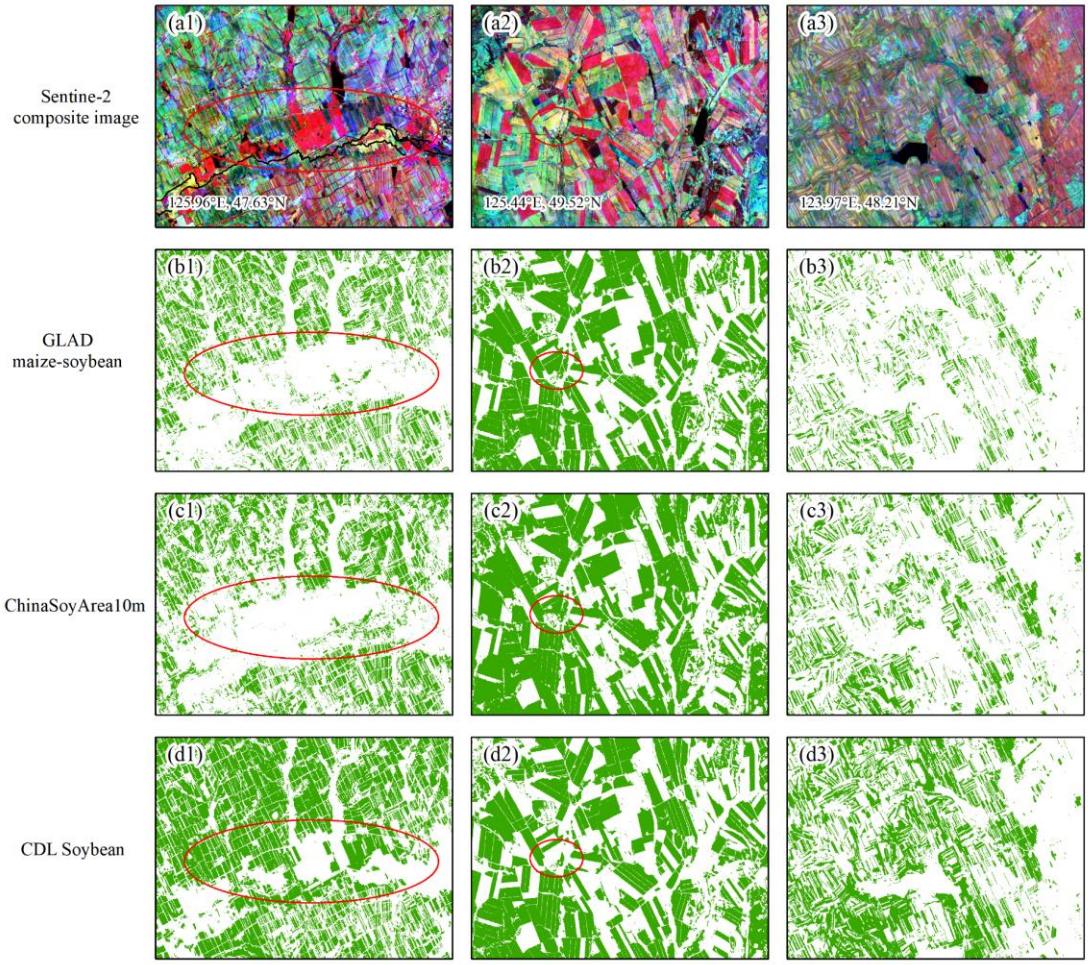

Figure 3. Visual comparison of our soybean maps and existing products in typical regions in 2019.

Comment 11: Fig. 9 indicates that there is a notably low frequency of clear observations in Sichuan Province, with the majority of areas showing zero clear observations per month. How can it be ensured that a complete 10-day composited time series is generated for DTW calculations in this region?

Reply: Yes, low frequency of clear observations was notably observed in Sichuan Province. For the areas with lower clear observations, beside the 10-day time series composite, we also conducted a gap-filling method on the composite time series by replacing the observations by the median of three adjacent observations (i.e., previous, current, and subsequent observations), to ensure the integrity of the time series as much as possible. We supplement in the "Data Processing" section:

"In areas with notably limited clear observations, a gap-filling method was conducted on the composite time series. This method involves substituting any given observation with the median value from three neighboring observations (i.e., previous, current, and subsequent observations) to maximize the continuity and completeness of time series."

Such fewer clear observations are inevitable, especially for a study over a larger region and long-term period. Although the 10-day composite time series were generated as far as possible, honestly, the uncertainty is inevitably introduced at times (such as 2017) and regions (such as the southwest) where there are particularly few clear observations. We added discussion to the “4.2 The uncertainty from image quality” section:

“In areas with quite lower clear observations, despite a gap-filling method was conducted to generate complete 10-day composite time series, higher uncertainty is inevitable. The gap-filling time series might contain duplicate values, which cannot accurately reflect the crop growth process in reality. Obviously, the total number of images available in 2017 over the study areas was significantly fewer than those of other years (Fig.10a1-e1) ... This might explain the lower user’s accuracy of soybean in Zone IV compared to other sub-zones (Table S1) and low overall accuracy based on sample verification in 2017 (Table 2).”

Comment 12: A considerable number of pixels corresponding to field ridges were inaccurately classified as soybeans in the 2020 map, particularly evident in East Heilongjiang, North Shandong and Henan Province. Can the authors consider the use of post-processing methods to eliminate this issue?

Reply: Thank you for pointing out the problem. We agree that ridge identification is a very important issue in remote sensing mapping, however it is still difficult to address the issue across a larger area. The main reasons are as follows:

(1) The ridge width is very narrow, and the 10m resolution image is often unable to accurately distinguish between the field and the ridge. It is generally accepted that the identification and elimination of the ridge is based on centimeter-level images (such as unmanned aerial vehicle images).

(2) We summarized the methods widely used to identify and eliminate the field ridges nowadays.

- Machine learning and deep learning methods. A labeled training dataset was used to train the model to identify planting areas and ridge area (Hamano et al., 2023).
- Point cloud processing technology. Point cloud data can reflect the height of the ground canopy, and the height of the ridge is often lower than that of the crop, so a suitable threshold can be adopted to distinguish the ridge from the crop (Liu et al., 2018)；
- Image processing and computer vision methods. The ridge has its special shape, such as a slender shape similar to a road or a closed border. Edge detection, morphological processing and other methods can extract features from remote sensing images to help identify and distinguish ridges (Li and Qu, 2019)。

Therefore, considering the relatively weaker impacts of the field ridge on crop mapping over a larger areas, and the complex image processing algorithms (which will consume huge computing power), we have not realized the field ridge identification after trading off the cartographic accuracy and calculation cost. In future studies, with the improvement of data accuracy and algorithm update, the identification of field ridge will be a key step in large-scale crop mapping. Following your suggestions, we added the discussion to “4.3 Limitations in small-scale planting areas” section:

“Our regional adaptive large-area crop mapping method in future will further be improved by the follows: … (4) Better post-processing of data. Misclassification of field ridges and image speckles is inevitable during mapping crops over large areas. With the progress of computing power, auxiliary data and image processing algorithms can further eliminate these issues (Liu et al., 2018; Li and Qu, 2019; Hamano et al., 2023). We are sure that integrating cloud computing platforms with advanced algorithms will provide substantial potential for accurate crop identification covering larger areas in future.”

Reference:

Hamano, M., Shiozawa, S., Yamamoto, S., Suzuki, N., Kitaki, Y., and Watanabe, O.: Development of a method for detecting the planting and ridge areas in paddy fields using AI, GIS, and precise DEM, Precis. Agric., 24, 1862–1888, https://doi.org/10.1007/s11119-023-10021-z, 2023.

Li, Y. and Qu, H.: LSD and Skeleton Extraction Combined with Farmland Ridge Detection, in: Advances in Intelligent, Interactive Systems and Applications, Cham, 446–453, https://doi.org/10.1007/978-3-030-02804-6_59, 2019.

Liu, H., Zhang, J., Pan, Y., Shuai, G., Zhu, X., and Zhu, S.: An Efficient Approach Based on UAV Orthographic Imagery to Map Paddy With Support of Field-Level Canopy Height From Point Cloud Data, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., 11, 2034–2046, https://doi.org/10.1109/JSTARS.2018.2829218, 2018.