Reviewer #1:

This study presents a novel data-driven framework for annual mapping of cropping patterns and produces maps of China's complex cropping systems. Overall, the study is carefully conducted and adds useful insights to the existing literature. However, several aspects of the manuscript require minor revision to improve the overall quality and interpretability. Detailed suggestions are listed below.

There is a mistake in Line 483. Based on Fig. 12, the most important embedding band for wheat is A21 rather than A22. The RGB composites using band A22 in Fig. 10 and Fig. 11 should also be replaced with A21.
Response: Thank you for your careful check. We have corrected the mistake in Line 483. The most important embedding band for wheat has been revised from A22 to A21 according to Fig 12 (Lines 432, 443 and 445). Accordingly, the RGB composites in Fig 10 and Fig 11 have also been updated by replacing band A22 with A21 (See Lines 448-455).

The proposed mapping framework uses satellite embeddings as input features for supervised classification. Given that satellite embeddings are a relatively recent and emerging data product that has not yet been widely adopted in remote sensing applications, The Discussion could be strengthened by a more detailed discussion of satellite embeddings in the context of remote sensing applications and related recent literature (In Lines 489-491).
Response: Thank you for your suggestion. We agree that a more comprehensive discussion of satellite embeddings in the context of remote sensing applications would strengthen the manuscript. In the revised Discussion (Lines 540–565), we have expanded this part by incorporating recent studies and providing a more structured perspective.

Specifically, we discuss how satellite embeddings have recently emerged from general land-cover and feature representation tasks and are gradually being extended to agricultural applications. Recent studies in land-cover mapping show that satellite embeddings outperform traditional spectral indices, improving overall accuracy by ~5% and Kappa by ~3% (Khan and Ahmad, 2025; Mantey et al., 2025). These improvements are attributed to their ability to capture high-level semantic information and integrate multi-sensor signals, thereby enhancing the discrimination of complex land-cover types and improving mapping stability.

Within this context, satellite embeddings have also shown strong performance in cropland mapping, with contributions more than twice those of traditional remote sensing features (Cai et al., 2026). Their non-linear response patterns suggest that embeddings encode richer and more complex information, although this may reduce their direct interpretability.

From a biophysical perspective, recent studies indicate that satellite embeddings preserve key vegetation-related information, including pigment, structural, and phenological signals, enabling accurate modeling of crop traits (e.g., R^2 up to ~0.9) while maintaining efficient representation (Alam and Simic Milas, 2025). Notably, despite being provided at annual temporal resolution, embeddings can still capture growing-season signals, suggesting their ability to retain phenological information relevant to crop dynamics.

At a more application-oriented level, AEF-based models generally achieve strong performance across a range of agricultural tasks and are competitive with purpose-built remote sensing models in applications including yield prediction ($R^2 \approx 0.7-0.8$ for major crops), as well as tillage and cover crop mapping (Ma et al., 2025).

However, several limitations have also been identified. In particular, AEF-based models may exhibit reduced spatial transferability due to region-specific feature representations, as well as limited interpretability arising from the lack of explicit physical meaning in embedding dimensions. In addition, their annual temporal resolution constrains sensitivity to within-season dynamics.

Nevertheless, these limitations do not undermine their effectiveness for large-scale crop mapping, where satellite embeddings provide a compact, information-rich, and scalable representation, as demonstrated in this study.

Reference:

Alam, M. M. T. and Simic Milas, A.: Dimensionality optimized machine learning retrieval of canopy chlorophyll, nitrogen, and phosphorus from google satellite embeddings, Smart Agricultural Technology, 12, 101601, https://doi.org/10.1016/j.atech.2025.101601, 2025.

Cai, Y., Li, B., Liu, X., Jiang, X., Zhu, Y., Luo, S., Qin, Y., Xie, S., Ye, J., Shen, H., Guo, Z., Liu, X., and Zeng, Z.: Annual 10-m high-resolution cropland maps for Southeast Asia since 2019 using AlphaEarth embeddings, Earth Syst. Sci. Data Discuss., 2026, 1-30, 10.5194/essd-2026-78, 2026.

Khan, H. and Ahmad, A.: Evaluating AlphaEarth Foundation Embeddings for Pixel- and Object-Based Land Cover Classification in Google Earth Engine, 10.20944/preprints202511.2172.v1, 2025.

Ma, Y., Shen, Y., Swatantran, A., and Lobell, D. B. J. a. p. a.: Harvesting AlphaEarth: Benchmarking the Geospatial Foundation Model for Agricultural Downstream Tasks, 2025.

Mantey, S., Attipoe, I., and Alhassan, A.: A Comparative Analysis of Satellite Imagery for Land Cover Classification: Evaluating Google Satellite Embeddings, Sentinel-2, And Landsat-8 Data with XGBoost, 2025.

In Line 221, “Multi-cropping” is used, whereas “Multi-cropped” appears in the legend of Fig. 2. Please ensure consistent terminology throughout the manuscript.
Response: Thank you for your suggestion. We apologize for the inconsistency in terminology. We have standardized the term to “Multi-cropping” throughout the manuscript, including Line 221 and the legend of Figure 2 (Lines 215-220). All related occurrences have been carefully checked and revised to ensure consistency (Figure 5, Lines 350-355; Figure 8,Lines 408-410; Figure 10, Lines 448-450; Figure 11, Lines 450-452).

In the Table 1, NorthEast should be revised to Northeast. Sourthern should be Southern. In Line 184, check the punctuation in this sentence, as there appears to be an extra comma at the end. Also, check the manuscript for inconsistent capitalization within sentences. (e.g. In Line 529).
Response: Thank you for your careful review. We have corrected “NorthEast” to “Northeast” and “Sourthern” to “Southern” in Table 1. The extra comma in Line 184 has also been removed. In addition, we have carefully checked the entire manuscript and corrected inconsistent capitalization within sentences (e.g., Line 529) to ensure consistency throughout.

This study produces a national-scale annual mapping product of cropping patterns in China for 2018–2021, which is valuable given the complexity of cropping systems and the limited availability of similar products. Nevertheless, it would be helpful for the Discussion to consider the potential extension of the proposed framework to broader spatial and temporal contexts, including cross-year transferability and within-season mapping, particularly given that the satellite embeddings are provided as annual products, which may limit their flexibility for in-season applications.
Response: Thank you for this constructive comment. We agree that the potential extension of the proposed framework to broader spatial and temporal contexts is an important aspect. We clarify that the current study is based on year-specific model training using locally derived samples, and does not explicitly address cross-year or cross-region transfer.

Recent studies indicate that, while satellite embeddings achieve strong performance under within-region settings, their spatial transferability can be limited due to geographic shifts in the feature space, which may reduce generalization performance in cross-region applications, particularly for tasks such as yield prediction (Ma et al., 2025).

In addition, we note that the current generation of satellite embeddings is typically provided at annual temporal resolution, which limits their ability to capture intra-seasonal crop dynamics and constrains their applicability for in-season mapping.

Despite these limitations, the contribution of this study lies in providing a more accurate and consistent baseline mapping product of cropping patterns at the national scale. Historical crop mapping products have been widely used to support subsequent prediction and in-season mapping. For example, in the United States, the long-term and consistently updated CDL dataset has enabled the development of automated national-scale in-season crop mapping and temporal transfer frameworks (Zhang et al., 2021; Li et al., 2024; Zhang et al., 2025). Similarly, the mapping results generated in this study can serve as a valuable reference for future work on spatiotemporal transfer and in-season crop mapping.

In the revised Discussion (Lines 578–587), we have expanded this part by incorporating additional relevant studies and providing further discussion.

Reference:

Li, H., Di, L., Zhang, C., Lin, L., Guo, L., Yu, E. G., and Yang, Z.: Automated In-Season Crop-Type Data Layer Mapping Without Ground Truth for the Conterminous United States Based on Multisource Satellite Imagery, IEEE Transactions on Geoscience and Remote Sensing, 62, 1-14, 10.1109/TGRS.2024.3361895, 2024.

Ma, Y., Shen, Y., Swatantran, A., and Lobell, D. B. J. a. p. a.: Harvesting AlphaEarth: Benchmarking the Geospatial Foundation Model for Agricultural Downstream Tasks, 2025.

Zhang, C., Di, L., Hao, P., Yang, Z., Lin, L., Zhao, H., and Guo, L.: Rapid in-season mapping of corn and soybeans using machine-learned trusted pixels from Cropland Data Layer, International Journal of Applied Earth Observation and Geoinformation, 102, 102374, https://doi.org/10.1016/j.jag.2021.102374, 2021.

Zhang, H. K., Shen, Y., Zhang, X., Li, J., Yang, Z., Xu, Y., Zhang, C., Di, L., and Roy, D. P.: Robust and timely within-season conterminous United States crop type mapping using Landsat Sentinel-2 time series and the transformer architecture, Remote Sensing of Environment, 329, 114950, https://doi.org/10.1016/j.rse.2025.114950, 2025.

Reviewer #2:

The manuscript integrated the embedding datasets and multiple cropland dataset to generate the cropping patterns by developing the data-driven framework. The work is novel and useful for understanding the dynamics of cropping patterns and guiding agricultural practices. There were two major comments: The impacts of spatiotemporal autocorrelation on accuracy should be considered when using RF. Since the study used multiple datasets, crop-specific area comparisons between the present study and other datasets would provide a more comprehensive view for guiding data application.
Response: Thank you for your suggestion. In the following text, we respectively conducted supplementary analysis and responses to these two major comments.

Minor comments:

Lines 20, what is the ADM-2 statistics, please state the full name of dataset at the first mention.
Response: Thank you for your suggestion. We have revised the text to provide the full name of ADM-2 (second-level administrative units) at its first occurrence in the abstract (See Line 20).

Lines 20, what are the units of RMSE and MAE?
Response: Thank you for your suggestion. We have now specified the units of RMSE and MAE in Line 21 (e.g., $\times 10^4$ ha) to improve clarity. We have also carefully reviewed the manuscript and ensured that the units are clearly stated wherever RMSE and MAE are reported to maintain consistency throughout the paper (See Lines 414-416,652).

Lines 29, please add common before and growing ….. To make the usage consistent throughout the manuscript.
Response: Thank you for your suggestion. We have added a comma before "and" in Line 29 and ensured consistent use of the serial (Oxford) comma throughout the manuscript.

Lines 50, you described the necessity of China mapping from the technic perspective. That is good. But adding some socio-economic background of China could help strengthen the context.
Lines 65, Please also mention its impacts on social impacts.
Response: Thank you for these helpful suggestions. We have revised the Introduction to incorporate additional socio-economic context of China, including its large population, limited arable land resources, and its reliance on intensive agricultural production to sustain global food supply (See Lines 50-53). We have also strengthened the description of the broader social

implications by revising the concluding sentence (Line 68-69), highlighting the relevance of this work for food security assessments and decision-making under complex cropping systems.

Lines 130, “Embeddings’ general”. Please check this throughout the manuscript to make their usage consistent.
Response: Thank you for your suggestion. We have revised the phrase to “the general features derived from Google Satellite Embeddings” and standardized the terminology throughout the manuscript for consistency (See Lines 133-134).

Lines 170, what is the threshold “t”? does “t” equal to 0.8?
Response: Thank you for your suggestion. The threshold t equals 0.8 in this study. We have clarified this explicitly in Line 184 to avoid ambiguity.

Line 175, how do you process the OCTC when it is less than the estimated cropping intensity? Does the “unconfused area” mean that OCTC is equal to the intensity? Please specify these categories.
Response: Thank you for your suggestion. In our framework, pixels where the number of overlapping crop types (Overlapping Crop Type Counts, OCTC) is less than or equal to the estimated cropping intensity are treated as *consistent* (i.e., “unconfused”), provided that OCTC is greater than 0.

As described in the manuscript, pixels where OCTC exceeds the estimated cropping intensity are classified as *Confused* areas, while pixels with a cropping intensity greater than 0 but no coverage from existing crop mapping products (i.e., OCTC = 0) are classified as *Unlabeled*.

Therefore, the “unconfused” area does not require OCTC to be exactly equal to the cropping intensity; rather, it includes all pixels where OCTC is less than or equal to the estimated cropping intensity and greater than 0. The corresponding description has been revised in the manuscript to clarify these category definitions (See Lines 195-203).

Lines 165-178: Please re-organize these paragraphs to make the method clearer. If CMCI less than 0.8, then the relative grid is assigned as low-consistency. Meanwhile, if the OCTC is greater than the estimated cropping intensity, then is it “confused” area? Please plot a figure to link consistency and confusion. I am a little confused about this.
Response: Thank you for your suggestion. We agree that the description of the classification rules could be clearer. We have reorganized Lines 179–203 to improve the logical flow and explicitly distinguish between the two criteria.

In the revised text, low-consistency is first identified within the same crop type based on CMCI (Crop Map Consistency Index, CMCI), reflecting inconsistencies among products of the same crop type. Subsequently, confused and unlabeled areas are determined based on the integration of multiple product types, using the relationship between OCTC (Overlapping Crop Type Counts, OCTC) and the estimated cropping intensity (e.g., OCTC > cropping intensity for confused areas, and OCTC = 0 for unlabeled areas). This revision clarifies both the sequential procedure and the distinction in their analytical scope.

To further improve clarity, we have added a schematic figure illustrating the overall workflow of the analysis, as shown in Figure S2 in the Supplementary Information (also provided below).

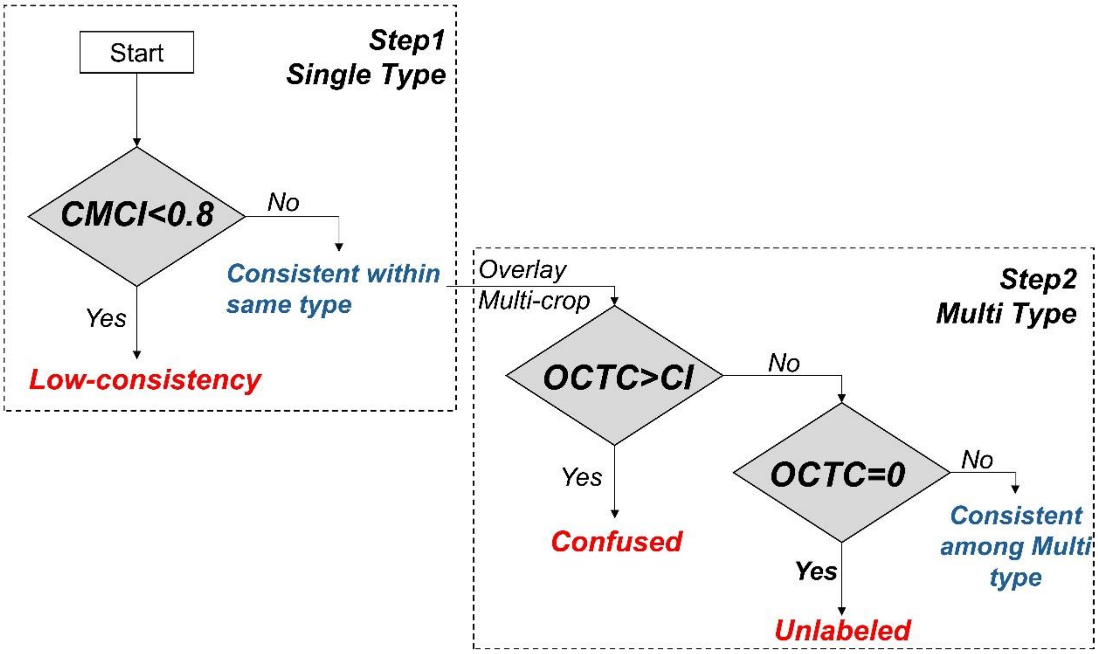


Figure S2. Workflow of Consistency Analysis

Lines 184: delete one "."

Response: Thank you for pointing this out. The extra period has been removed (see Lines 210).

Lines 205-207: Please test the model performance by testing various combination and ntree and mtry. Not sure 500 is the best value.

Response: Thank you for your valuable suggestion. We conducted additional experiments to evaluate model performance under different combinations of ntree and mtry. To further assess the effect of parameter tuning, we computed the relative differences in OA and Kappa with respect to the baseline configuration (mtry = 8 and ntree = 500) for the year 2020.

The results show that the impact of parameter variations is generally limited. Most differences in OA are within 0.01, and differences in Kappa are similarly small, indicating that model performance is relatively insensitive to parameter tuning. For mtry, the median performance with mtry = 6 is slightly lower than the baseline, while mtry = 12 shows only marginal improvement, suggesting that the default setting (mtry = 8, the square root of the number of input features (64)) is appropriate. For ntree, performance increases initially but stabilizes around 500 trees, with differences in OA below 0.005 and in Kappa below 0.05 (Figure S4).

Based on these results, we retained the baseline configuration, as it provides stable performance without unnecessary computational cost. The corresponding analysis and results have been added to the revised supplementary materials (See Section S2.2 and Figure S4).

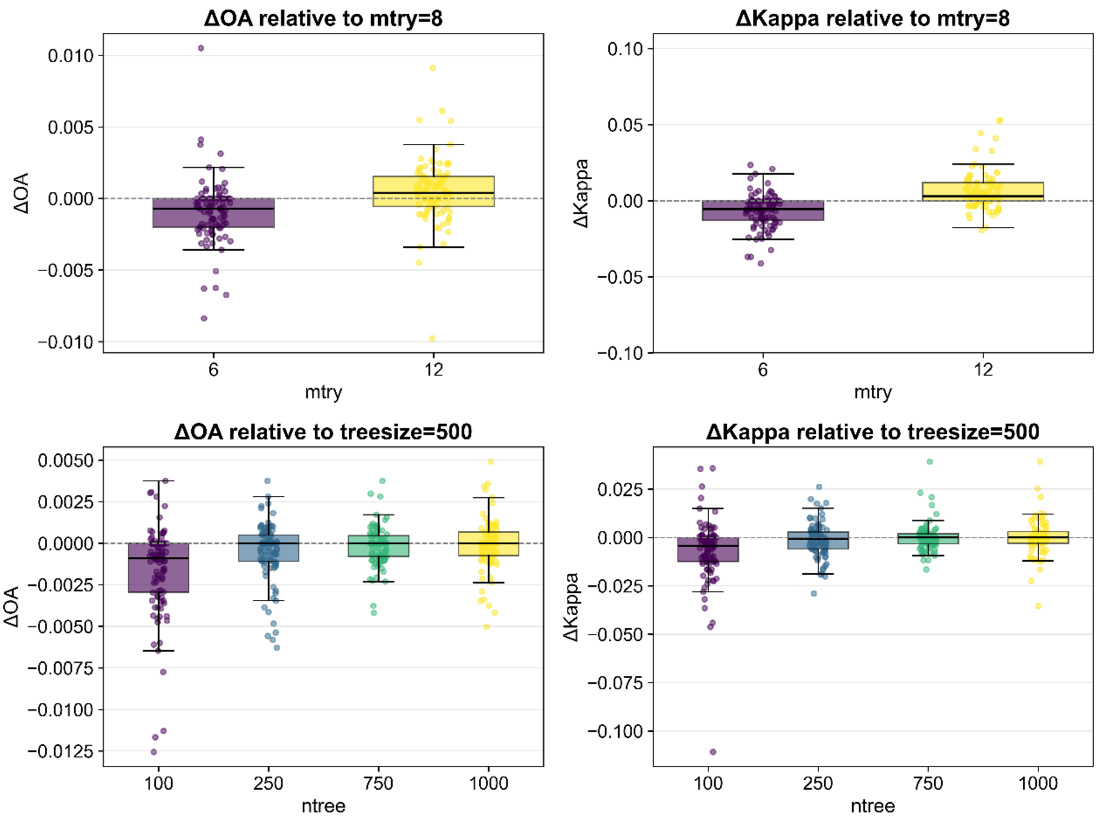


Figure S4. Relative differences in OA and Kappa compared with the baseline parameter settings.

Lines 207: “To capture regional heterogeneity, the RF classification was partitioned using the Agro-Natural Regionalization of China, a climatic-geographic framework dividing the country into 38 sub-regions by temperature and precipitation” what does “the RF classification was partitioned”? do you mean train RF separately in each sub-regions? Also, please consider the influences of spatiotemporal autocorrelation on the accuracy evaluation.

Response: Thank you for your helpful comments.

First, we apologize for the ambiguity in the expression “the RF classification was partitioned”. What we meant is that independent RF models were trained within each of the 38 agro-natural sub-regions defined by the Agro-Natural Regionalization of China. In other words, samples were not pooled nationwide; instead, each sub-region was modeled separately to better capture regional heterogeneity in climate conditions and cropping systems. We have revised the wording in the manuscript to clarify this point.

Second, we acknowledge that sample-level random splitting within the same region and year may lead to optimistic accuracy estimates due to spatial autocorrelation (i.e., spatial leakage between nearby samples). To quantify this effect, we conducted an additional spatially explicit sensitivity analysis. Specifically, we assigned each sample to a hexagonal grid cell (as shown in Figure 2 in Main Text) and performed group-wise splitting at the hex level, ensuring that all samples within the same hex cell were assigned exclusively to either the training or testing set. This hex-based split reduces spatial leakage compared to random splitting.

We report the performance gap between random splitting and hex-group splitting (ΔOA and Δ kappa) as an empirical estimate of the inflation in accuracy induced by spatial autocorrelation (Figure S5). The results show that while accuracy metrics decrease under the more conservative hex-based validation, this mainly reflects the limited spatial transferability of classifiers trained on embedding features rather than deficiencies in the proposed mapping framework. It is important to note that the primary contribution of this study lies in developing a multi-product integration and reclassification framework and generating corresponding mapping products, for which the original sample-based validation strategy using random splitting provides a more representative assessment of mapping accuracy within the target regions and years. This is further supported by the fact that the training and validation samples were selected using a stratified random sampling scheme within consistent areas across multiple products, ensuring adequate spatial coverage without evident regional gaps. These additional results are now provided in the revised supplementary materials (See Section S2.3).

Finally, regarding temporal autocorrelation, we clarify that the proposed framework does not involve cross-year transfer, but is designed to generate consistent crop maps using year-specific samples and imagery within each year. Therefore, the current evaluation focuses on within-year classification performance, which is aligned with the objective of this study. We acknowledge that the temporal transferability of embedding-based models remains an important open question. Recent studies suggest that embedding features may exhibit limitations in generalization across different spatial or temporal contexts. Despite these limitations, this study provides a more accurate and consistent national-scale baseline dataset that can support future spatiotemporal transfer and in-season crop mapping studies.

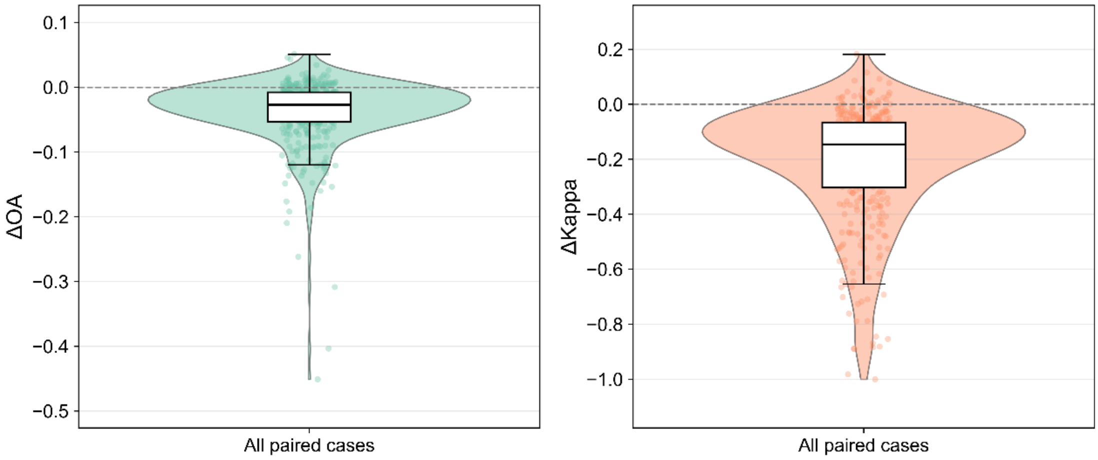


Figure S5. Relative differences in OA and Kappa under random and spatial (hex-based) splitting

Lines 229: is the “adm-1 level” the same as ADM-1 level? Please make all terminology consistent.

Response: Thank you for your careful reading. We apologize for the inconsistency in terminology. We have now standardized the expression to “ADM-1 level” throughout the manuscript for consistency. All related occurrences have been carefully checked and revised accordingly (e.g. see Line 256).

Lines 248-254: this section should be moved to the Method.

Response: Thank you for your helpful suggestion. We agree that the content in Lines 248–254 is methodological rather than result-oriented. Accordingly, we have moved this section to the Methods (Section 2.1) and integrated it into the relevant methodological description (See Lines 185-203).

Lines 255-256: please cite a figure to show the province boundary and the Hu Huanyong line.

Response: Thank you for your suggestion. We agree that a visual reference would improve clarity. Instead of adding a new figure, we have updated Figure 3 by incorporating the provincial boundaries and the Hu Huanyong Line (see Lines 293-294). The revised figure now provides the necessary geographic context for the subsequent analysis.

Fig 4(f): it seems that most areas have the error less than 0.25. Might use a more detailed legend below 0.25 to show the spatial heterogeneity.

Response: Thank you for your suggestion. We have revised the legend design in Fig. 4 as a whole, introducing more detailed intervals to better highlight the spatial heterogeneity (See Lines 328-329), the corresponding descriptions in the manuscript have also been updated accordingly (See Lines 300-326).

Lines 315-316: confused sentence. The area of cotton is small. Although 98% of cotton remains stable, it can't represent the stability of all cropping systems. Please revise this.

Response: Thank you for the helpful comment. We have revised this sentence to improve clarity and avoid potential misinterpretation. The revised text now specifies that the high stability refers specifically to the cotton cropping system and is presented in a comparative context across different crop types, rather than implying the stability of overall cropping systems (See Lines 362-363).

Lines 332: "there was a high degree of alternation from wheat–soybean to wheat–maize systems, accounting for 48% of the wheat–soybean area,". What is the driver for this transition? Market or policy change?

Response: Thank you for this insightful comment. We agree that the drivers of this transition warrant further explanation. Previous studies, mainly conducted in Northeast China, indicate that soybean-support policies (e.g., the soybean producer subsidy) have had limited effectiveness in expanding soybean planting areas (Di et al., 2023), largely due to lower economic returns of soybean compared to maize, as well as constraints related to resources and farmers' planting preferences (Di et al., 2023; Liu et al., 2019; Peng et al., 2022).

Although these findings are region-specific, they suggest that similar economic and behavioral factors may also contribute to the observed transition from wheat–soybean to wheat–maize systems in the North China Plain. We therefore interpret these factors as possible drivers, while acknowledging that further region-specific analysis would be needed to fully disentangle the underlying mechanisms.

Reference:

Di, Y., You, N., Dong, J., Liao, X., Song, K., and Fu, P.: Recent soybean subsidy policy did not revitalize but stabilize the soybean planting areas in Northeast China, European Journal of Agronomy, 147, 126841, https://doi.org/10.1016/j.eja.2023.126841, 2023.

Liu, S., Zhang, P., Marley, B., and Liu, W.: The Factors Affecting Farmers' Soybean Planting Behavior in Heilongjiang Province, China, 10.3390/agriculture9090188, 2019.

Peng, L., Zhou, X., Tan, W., Liu, J., and Wang, Y.: Analysis of dispersed farmers' willingness to grow grain and main influential factors based on the structural equation model, Journal of Rural Studies, 93, 375-385, https://doi.org/10.1016/j.jrurstud.2020.01.001, 2022.

In Section 3.4 : Please report the validation for each crop type. Although crop-specific validation might be poor, it is import for the data application, especially for the crop-specific application. AMD reported the harvested area, but the map area reflected the planted area. Therefore, the map area might be higher than AMD area, which is reasonable. In addition, please conduct a crop area comparison between your map results and the used crop products, crop by crop. This could give the reader a more comprehensive view of the current dataset.

Response: Thank you for your helpful suggestion. We agree that crop-specific validation is essential for practical applications, particularly for crop–specific analyses. Crop-specific accuracy has already been reported separately in the original manuscript, including F1-score, omission error and commission error in Supplementary Material (see Section S2) and is also discussed in the main text (see Section 3.2).

In addition, following the reviewer’s suggestion, we conducted crop-by-crop comparisons between our mapping results, the input crop products, and crop-specific statistics. The results, presented in Figures S6–S12 (also provided below), focus on the top five provinces for each crop in terms of planting area.

Overall, our mapping results show good consistency with both the input products and statistical data across most provinces, particularly for major staple crops such as rice, wheat, and maize. For other crops (e.g., soybean, sugarcane, and rapeseed), larger discrepancies are observed among different products and between products and statistics. In some cases, our framework helps reconcile these differences, while in others it remains more consistent with specific input products. This is because the proposed framework performs multi-product integration and reclassification at the pixel level under consistent cropland extent and cropping intensity constraints, rather than aggregating results at the provincial scale (e.g., by averaging or majority voting).

It should also be noted that this study includes only seven single-cropping types and five major crop sequences, while all other cropping patterns are grouped into an “other multi-cropping” category. As a result, some cropping systems involving these crops are partially included in this aggregated class, which may lead to systematic underestimation of crop-specific areas in our results. This effect is evident for rice, where both our mapping results and similar products tend to be lower than statistical data, and is also observed for other crops such as rapeseed. These additional results are now provided in the revised supplementary materials (See Section S2.4).

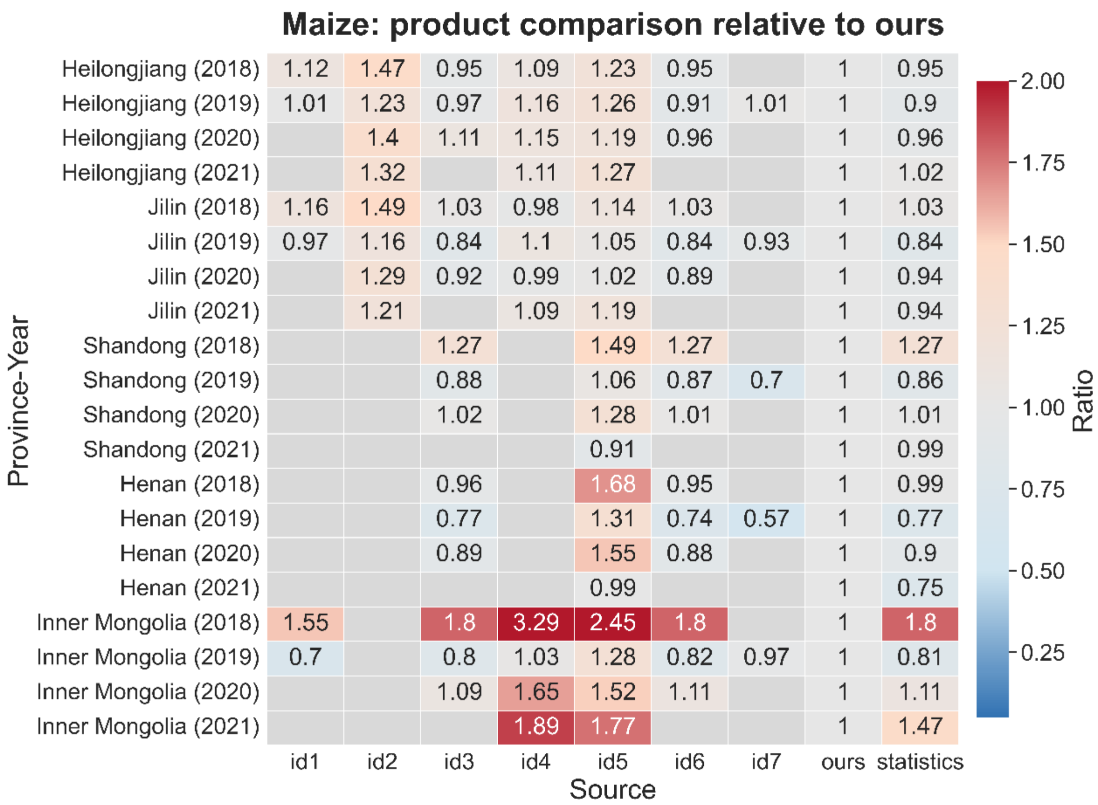


Figure S6. Comparison of maize area among our mapping results, input products, and statistics.

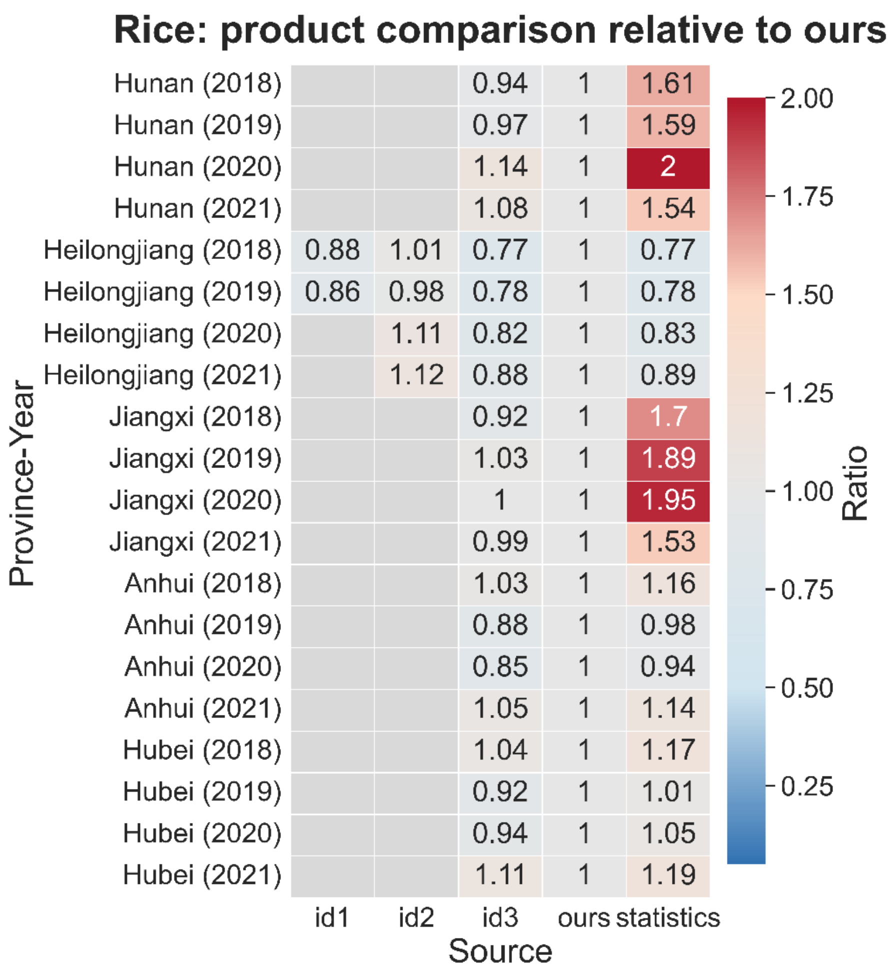
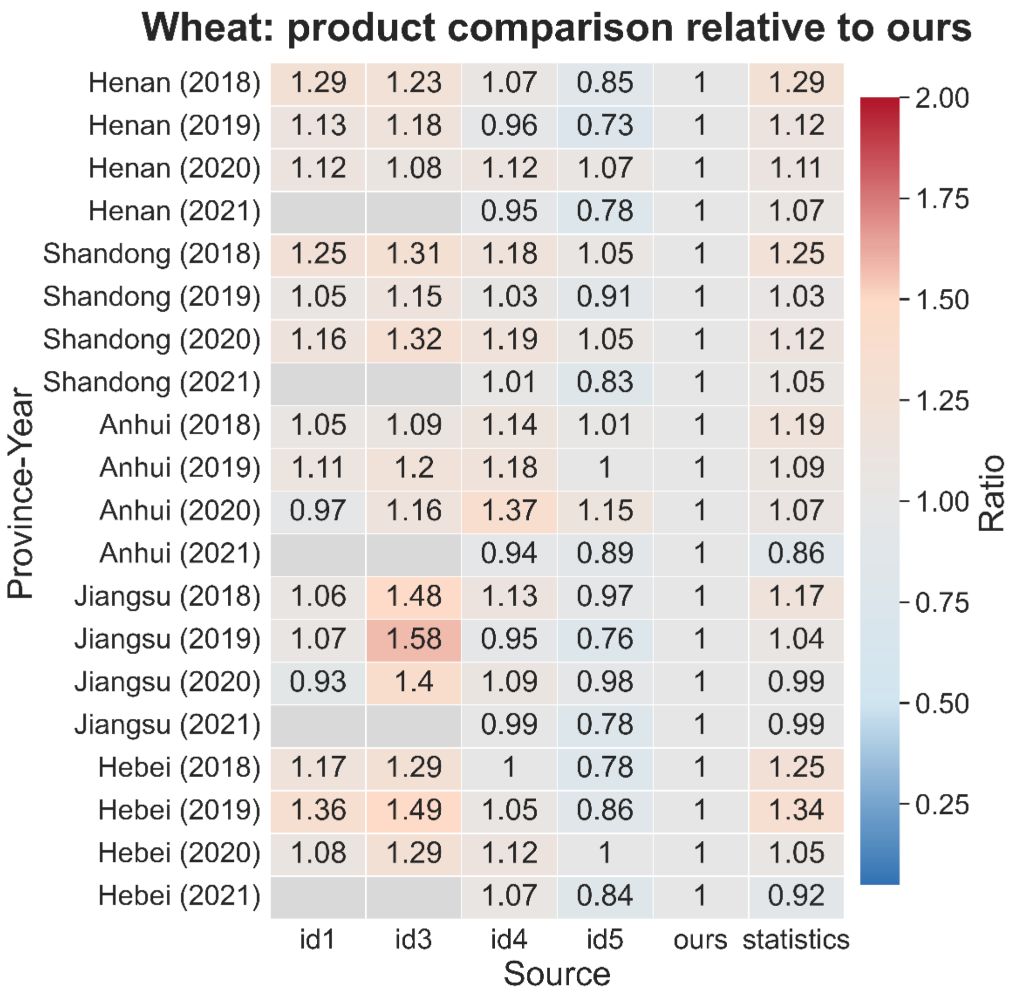
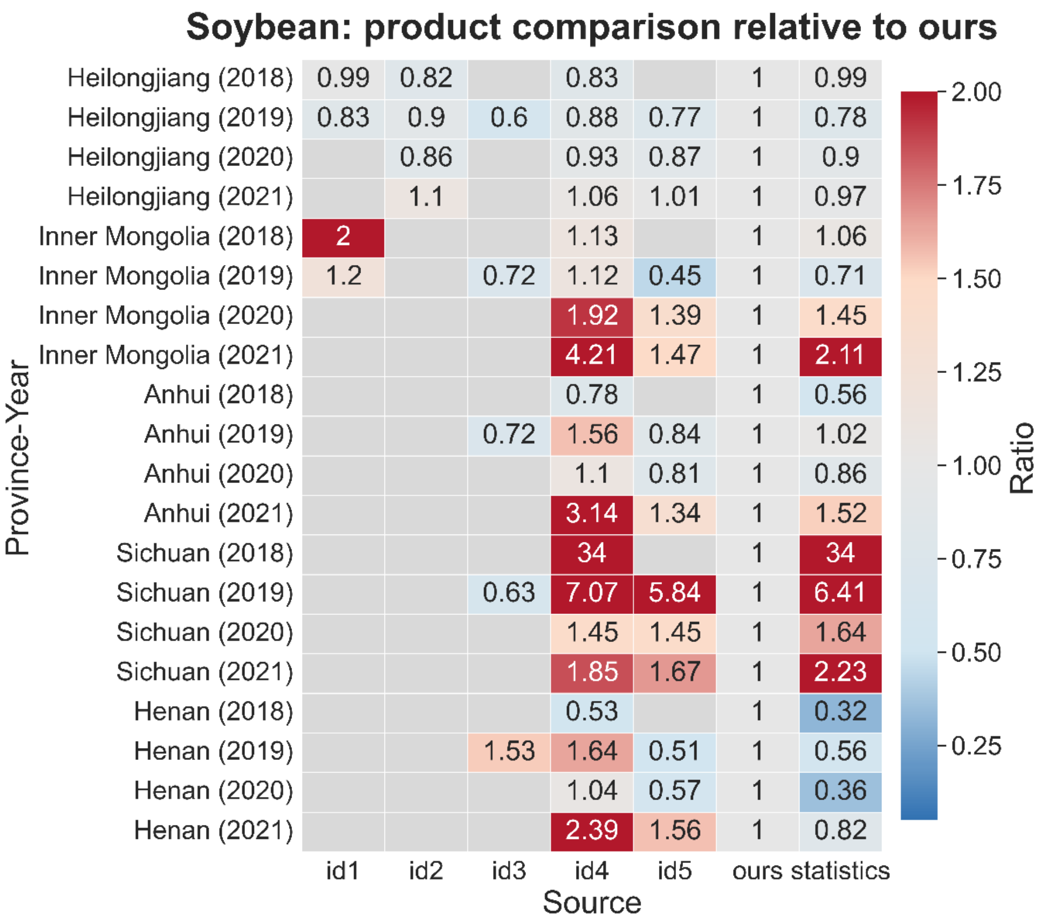


Figure S7. Comparison of rice area among our mapping results, input products, and statistics. The reported area of product *id3* corresponds to the combined area of single- and double-season rice (i.e., *id3* + *id4*).



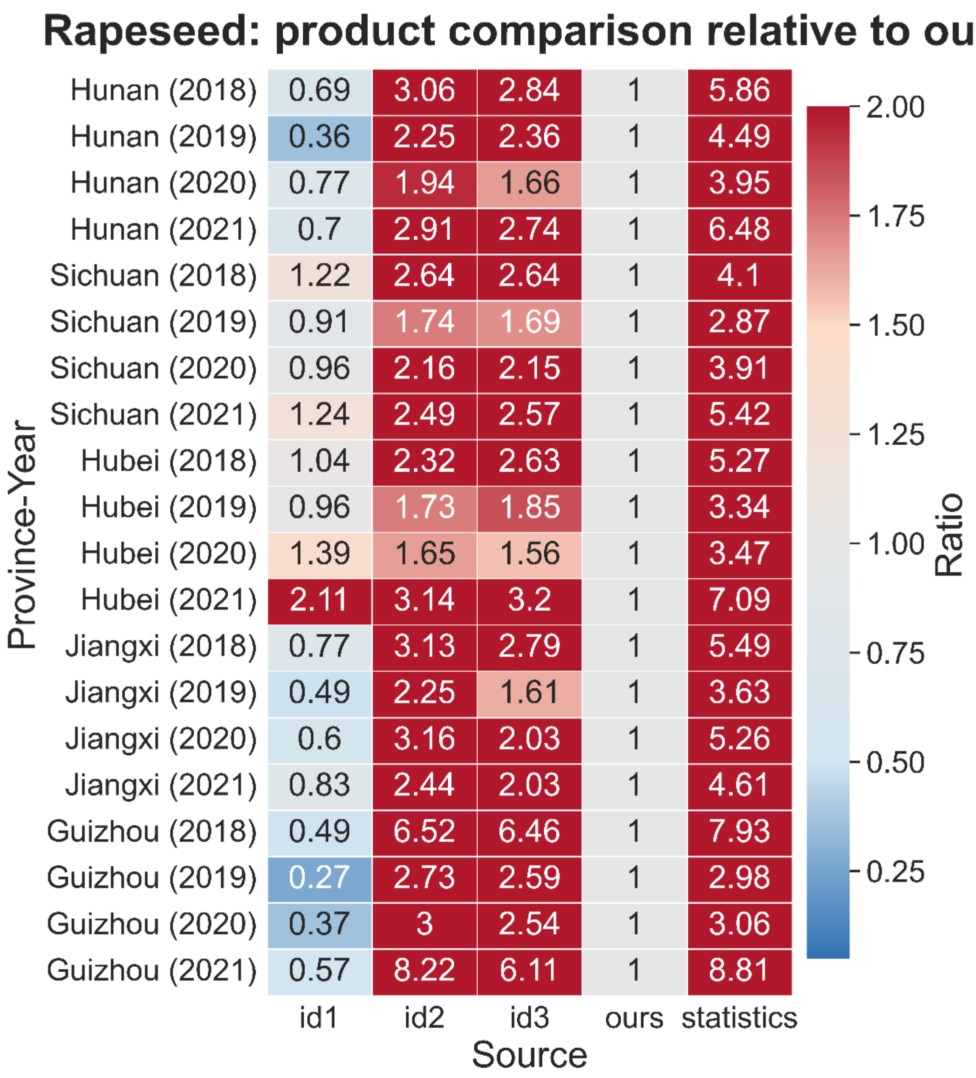
Province-Year	id1	id3	id4	id5	ours	statistics
Henan (2018)	1.29	1.23	1.07	0.85	1	1.29
Henan (2019)	1.13	1.18	0.96	0.73	1	1.12
Henan (2020)	1.12	1.08	1.12	1.07	1	1.11
Henan (2021)			0.95	0.78	1	1.07
Shandong (2018)	1.25	1.31	1.18	1.05	1	1.25
Shandong (2019)	1.05	1.15	1.03	0.91	1	1.03
Shandong (2020)	1.16	1.32	1.19	1.05	1	1.12
Shandong (2021)			1.01	0.83	1	1.05
Anhui (2018)	1.05	1.09	1.14	1.01	1	1.19
Anhui (2019)	1.11	1.2	1.18	1	1	1.09
Anhui (2020)	0.97	1.16	1.37	1.15	1	1.07
Anhui (2021)			0.94	0.89	1	0.86
Jiangsu (2018)	1.06	1.48	1.13	0.97	1	1.17
Jiangsu (2019)	1.07	1.58	0.95	0.76	1	1.04
Jiangsu (2020)	0.93	1.4	1.09	0.98	1	0.99
Jiangsu (2021)			0.99	0.78	1	0.99
Hebei (2018)	1.17	1.29	1	0.78	1	1.25
Hebei (2019)	1.36	1.49	1.05	0.86	1	1.34
Hebei (2020)	1.08	1.29	1.12	1	1	1.05
Hebei (2021)			1.07	0.84	1	0.92

Figure S8. Comparison of wheat area among our mapping results, input products, and statistics.



Province-Year	id1	id2	id3	id4	id5	ours	statistics
Heilongjiang (2018)	0.99	0.82		0.83		1	0.99
Heilongjiang (2019)	0.83	0.9	0.6	0.88	0.77	1	0.78
Heilongjiang (2020)		0.86		0.93	0.87	1	0.9
Heilongjiang (2021)		1.1		1.06	1.01	1	0.97
Inner Mongolia (2018)	2			1.13		1	1.06
Inner Mongolia (2019)	1.2		0.72	1.12	0.45	1	0.71
Inner Mongolia (2020)				1.92	1.39	1	1.45
Inner Mongolia (2021)				4.21	1.47	1	2.11
Anhui (2018)				0.78		1	0.56
Anhui (2019)			0.72	1.56	0.84	1	1.02
Anhui (2020)				1.1	0.81	1	0.86
Anhui (2021)				3.14	1.34	1	1.52
Sichuan (2018)				34		1	34
Sichuan (2019)			0.63	7.07	5.84	1	6.41
Sichuan (2020)				1.45	1.45	1	1.64
Sichuan (2021)				1.85	1.67	1	2.23
Henan (2018)				0.53		1	0.32
Henan (2019)			1.53	1.64	0.51	1	0.56
Henan (2020)				1.04	0.57	1	0.36
Henan (2021)				2.39	1.56	1	0.82

Figure S9. Comparison of soybean area among our mapping results, input products, and statistics.



Province-Year	id1	id2	id3	ours	statistics
Hunan (2018)	0.69	3.06	2.84	1	5.86
Hunan (2019)	0.36	2.25	2.36	1	4.49
Hunan (2020)	0.77	1.94	1.66	1	3.95
Hunan (2021)	0.7	2.91	2.74	1	6.48
Sichuan (2018)	1.22	2.64	2.64	1	4.1
Sichuan (2019)	0.91	1.74	1.69	1	2.87
Sichuan (2020)	0.96	2.16	2.15	1	3.91
Sichuan (2021)	1.24	2.49	2.57	1	5.42
Hubei (2018)	1.04	2.32	2.63	1	5.27
Hubei (2019)	0.96	1.73	1.85	1	3.34
Hubei (2020)	1.39	1.65	1.56	1	3.47
Hubei (2021)	2.11	3.14	3.2	1	7.09
Jiangxi (2018)	0.77	3.13	2.79	1	5.49
Jiangxi (2019)	0.49	2.25	1.61	1	3.63
Jiangxi (2020)	0.6	3.16	2.03	1	5.26
Jiangxi (2021)	0.83	2.44	2.03	1	4.61
Guizhou (2018)	0.49	6.52	6.46	1	7.93
Guizhou (2019)	0.27	2.73	2.59	1	2.98
Guizhou (2020)	0.37	3	2.54	1	3.06
Guizhou (2021)	0.57	8.22	6.11	1	8.81

Figure S10. Comparison of rapeseed area among our mapping results, input products, and statistics.

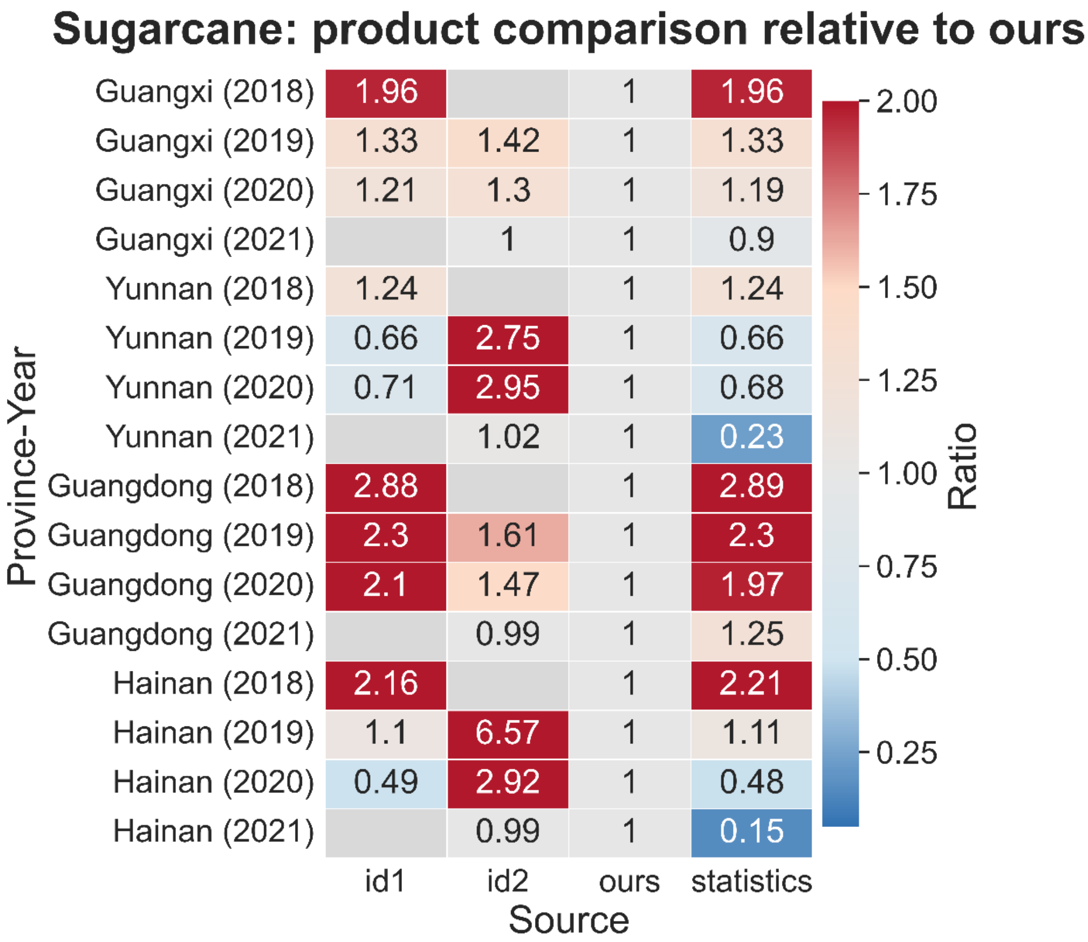


Figure S11. Comparison of sugarcane area among our mapping results, input products, and statistics.

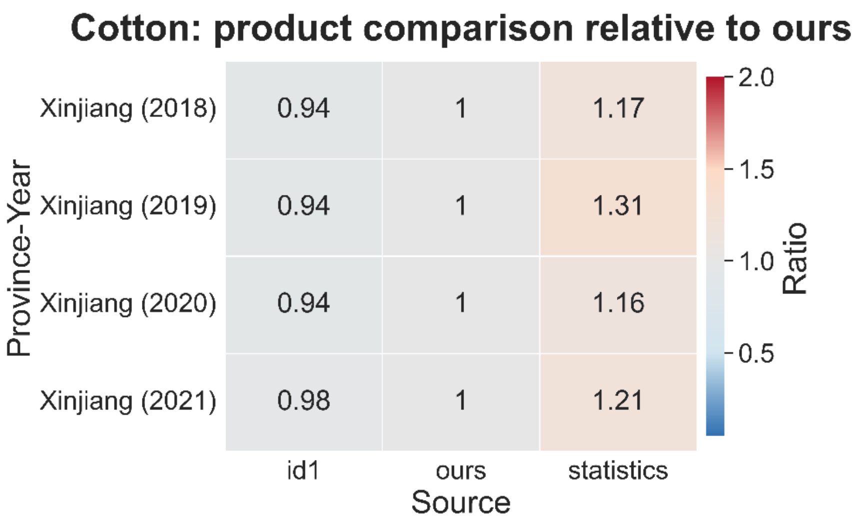


Figure S12. Comparison of cotton area among our mapping results, input products, and statistics.

Lines 480: Please specify the specific information in each band if it is available. For instance, A40 is important for rice, then A40 reflect which kind of information.
Response: Thank you for this insightful comment. We agree that improving the interpretability of embedding features is important. Although individual embedding dimensions do not have explicit physical meanings, recent studies suggest that they encode complex combinations of spectral, structural, and phenological information derived from multi-sensor observations (Alam and Simic Milas, 2025).

To further enhance interpretability, we conducted an additional analysis linking embedding features to raw remote sensing signals. Specifically, instead of using a large set of derived indices, we retained only the original optical (Sentinel-2 bands B2–B12) and SAR features (Sentinel-1 VV and VH) to reduce collinearity and improve physical interpretability. Cloud-contaminated pixels in Sentinel-2 imagery were masked using scene classification and cloud probability thresholds (60%). Both optical and SAR observations were aggregated into 10-day median composites, and missing values were filled using linear interpolation to ensure consistent and continuous time series.

For each crop type, we extracted embedding features and corresponding multi-temporal raw-band observations, and computed correlations between embedding features and raw bands across time, resulting in a time × band correlation matrix (Figure S13-15, also provided below). Signed correlation heatmaps were generated to characterize the direction and strength of these relationships.

To facilitate interpretation, we selected three representative crops in the Northeast China Plain (maize, rice, and soybean) and visualized the most important embedding dimensions identified in their classification models. The temporal axis represents10-day composited time steps throughout the growing season (DOY 90-300).

The results reveal crop-specific associations between embedding features and raw spectral signals. For maize, the most important embedding feature (A08) shows strong negative correlations with shortwave infrared bands (B11–B12) during DOY ~200–270, indicating sensitivity to canopy moisture dynamics during the mid-to-late growing season. For rice, A40 exhibits strong positive correlations with red-edge and SWIR bands (B6–B12) during DOY ~110–130, corresponding to the transplanting and early growth stages, where distinct water–soil conditions enhance separability from upland crops. For soybean, A33 shows strong positive correlations with red-edge bands (B6–B8A) during DOY ~220–270 and with SWIR bands (B11–B12) during DOY ~280–300, reflecting canopy development and subsequent moisture decline during senescence.

These patterns are consistent with known crop phenology and spectral separability in the region, where SWIR and red-edge bands play key roles in distinguishing crop types at critical growth stages. Compared with interpretations based on high-dimensional derived features, this analysis provides a more direct and physically interpretable understanding of embedding representations. These additional results are now provided in the revised supplementary materials (See Section S2.5).

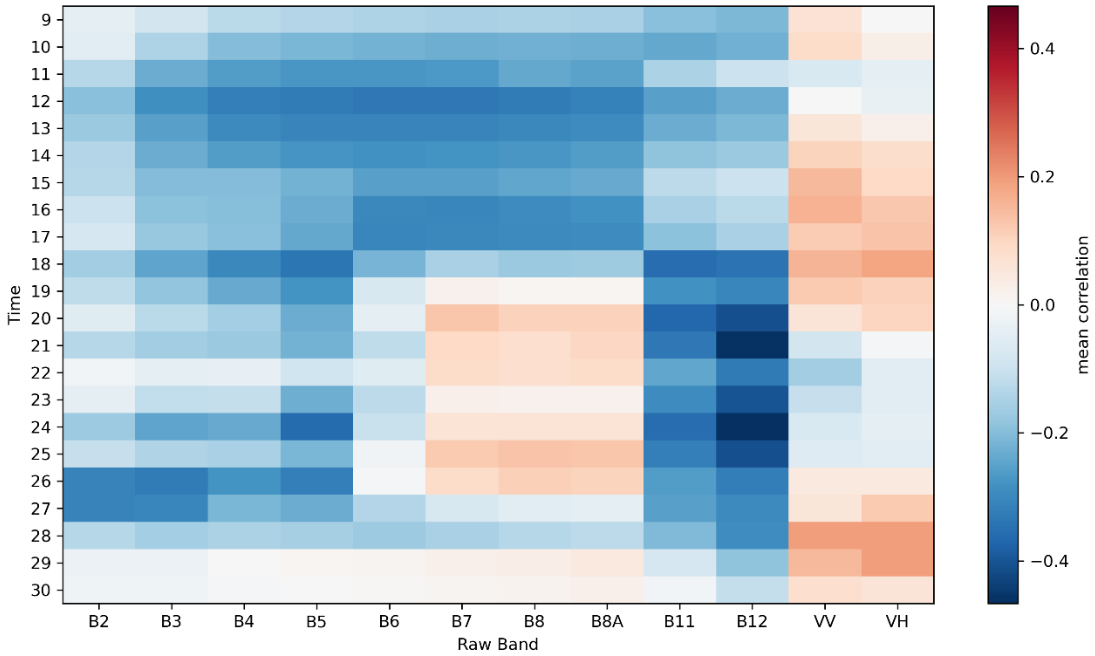


Figure S13. Correlation heatmap between embedding feature A08 and multi-temporal bands (Maize in Northeast China)

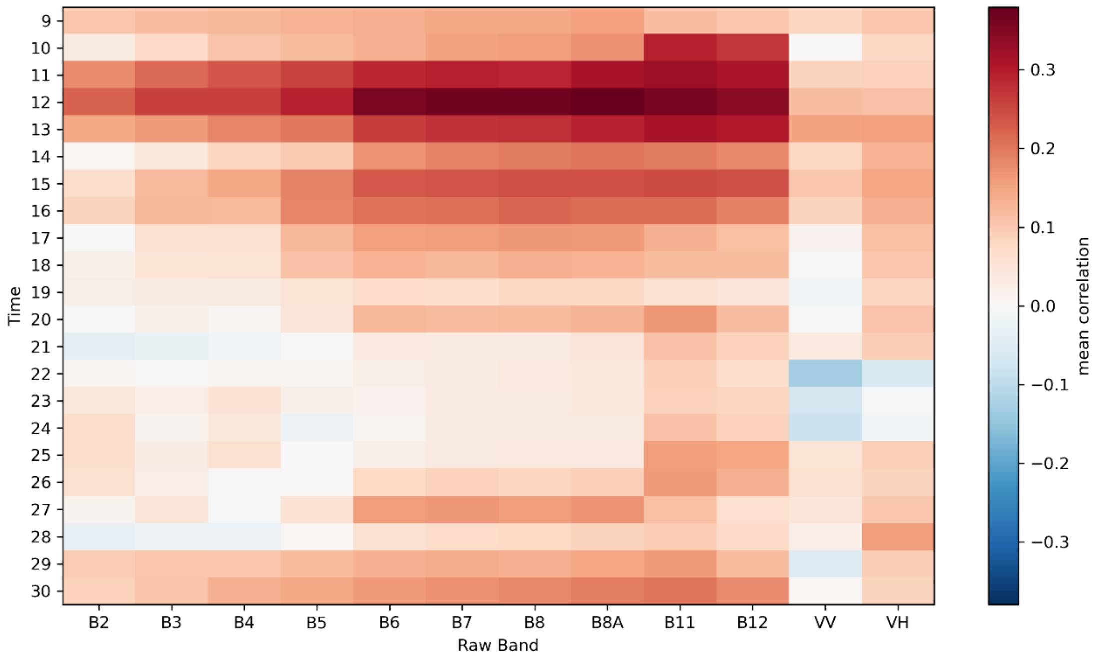


Figure S14. Correlation heatmap between embedding feature A40 and multi-temporal bands (Rice in Northeast China)

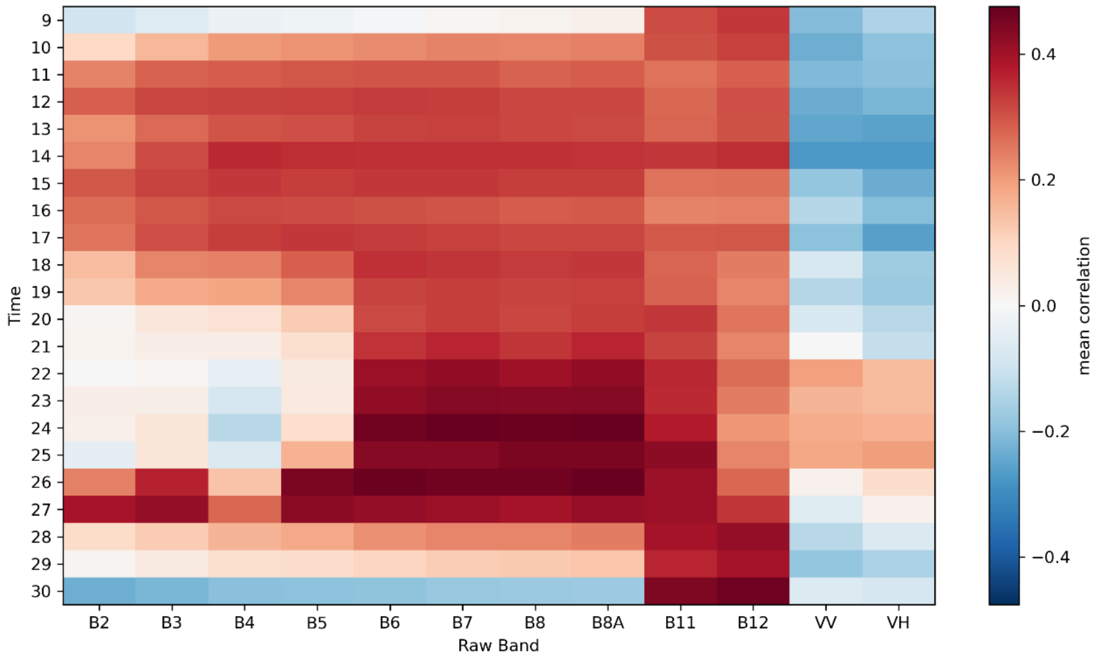


Figure S15. Correlation heatmap between embedding feature A33 and multi-temporal bands (Soybean in Northeast China)

Reference:

Alam, M. M. T. and Simic Milas, A.: Dimensionality optimized machine learning retrieval of canopy chlorophyll, nitrogen, and phosphorus from google satellite embeddings, Smart Agricultural Technology, 12, 101601, https://doi.org/10.1016/j.atech.2025.101601, 2025.