Response to Referee 1

In this work, the authors present a long-term rice distribution dataset covering South and Southeast Asia at a five-year interval. The paddy flooding signals and vegetation indices are used to identify multi-paddy fields. The overall concept of the study is sound, and the manuscript is clearly written. However, several issues require further clarification and justification. My detailed comments are as follows:

Response to Reviewer

We sincerely thank the reviewer for their careful reading of our manuscript and for the constructive and insightful comments. These suggestions have been extremely helpful in improving the clarity, methodological transparency, and presentation quality of our work. Below, we provide point-by-point responses to each comment. All changes have been incorporated into the revised manuscript.

All revisions and responses are highlighted in blue font in the response letter. Modifications in the main manuscript are indicated using italicized text enclosed in quotation marks.

1. Issues related to figure presentation. It is recommended that Figure S1 be moved to the main text, as both readers interested in the methodology and potential data users would benefit from a clear visualization of the study area. In addition, representative remote-sensing image examples of rice fields should be included to better illustrate the spectral and phenological characteristics used for rice identification. Furthermore, Figure S3 lacks a legend, making it difficult for readers unfamiliar with rice phenology to distinguish between the two curves shown.

Response:

We thank the reviewer for these valuable suggestions regarding figure presentation and visualization.

First, following your recommendation, we have moved Figure S1 from the Supplementary Materials to the main text and now present it as **Figure 1**, providing a clearer and more immediate overview of the study area for both methodological readers and potential data users.

Second, to better illustrate the spectral and phenological characteristics used for rice identification, we have added representative remote-sensing examples from typical rice systems in South and Southeast Asia. Specifically, we now include time-series profiles and corresponding imagery examples for:

- the rice–wheat rotation system in northern India,
- single-season rice systems in Bangladesh, and
- multi-season rice systems in the Mekong Delta.

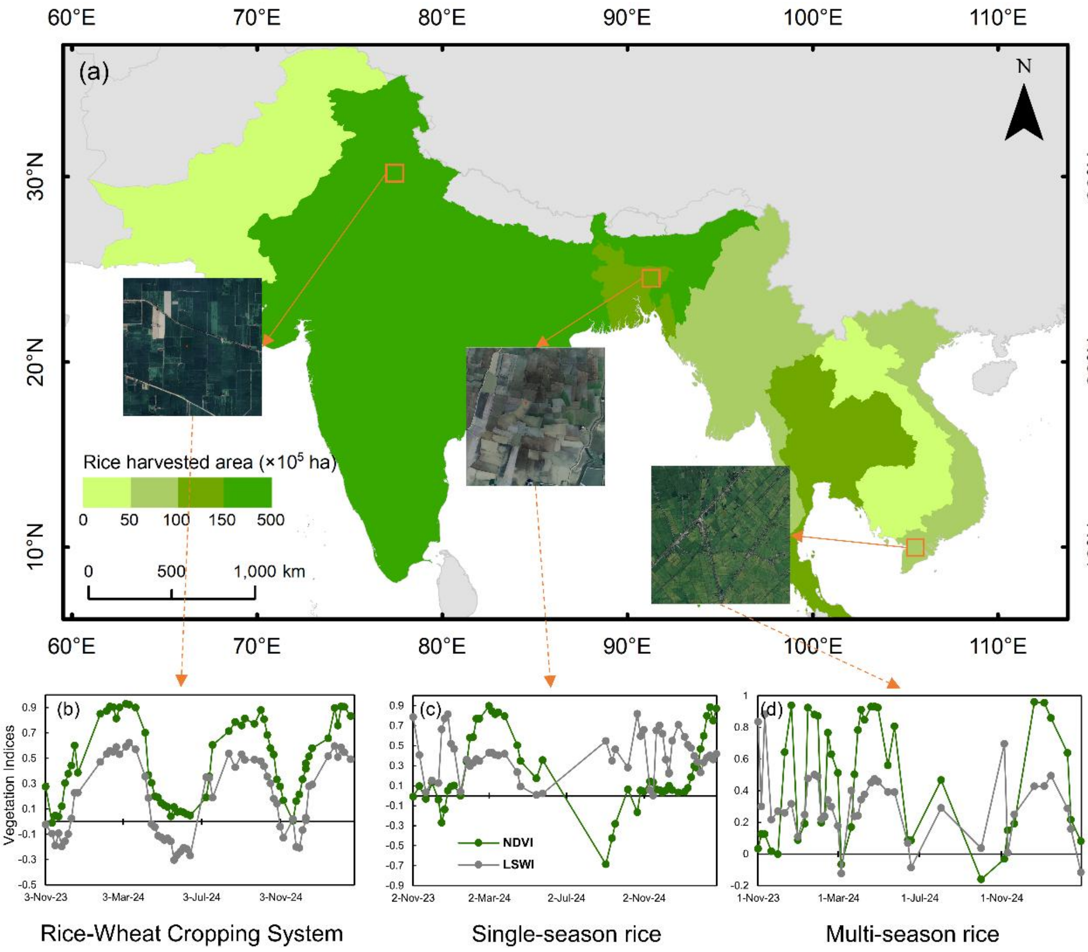


Figure 1. (a) Spatial distribution of rice harvested area in South and Southeast Asia. Insets show example rice fields at the locations marked by orange squares. (b–d) Representative NDVI (green) and LSWI (gray) time series for (b) rice-wheat cropping system, (c) single-

season rice, and (d) multi-season rice. The locations of the example pixels are indicated by arrows on the map.

Accordingly, the main text has been revised to better align with the updated figure, as follows (Lines 94-98): *"Rice is the dominant crop in this region and exhibits the highest cropping intensity worldwide. The region also includes diverse rice production systems, such as the rice–winter wheat rotation system in northern India, flood-associated single-season rice cultivation in Bangladesh, and typical multi-season rice systems in the Mekong River Basin."*

In addition, we have added a legend to **Figure S2** to clearly distinguish the LSWI and EVI curves, ensuring that readers unfamiliar with rice phenology can easily interpret the figure.

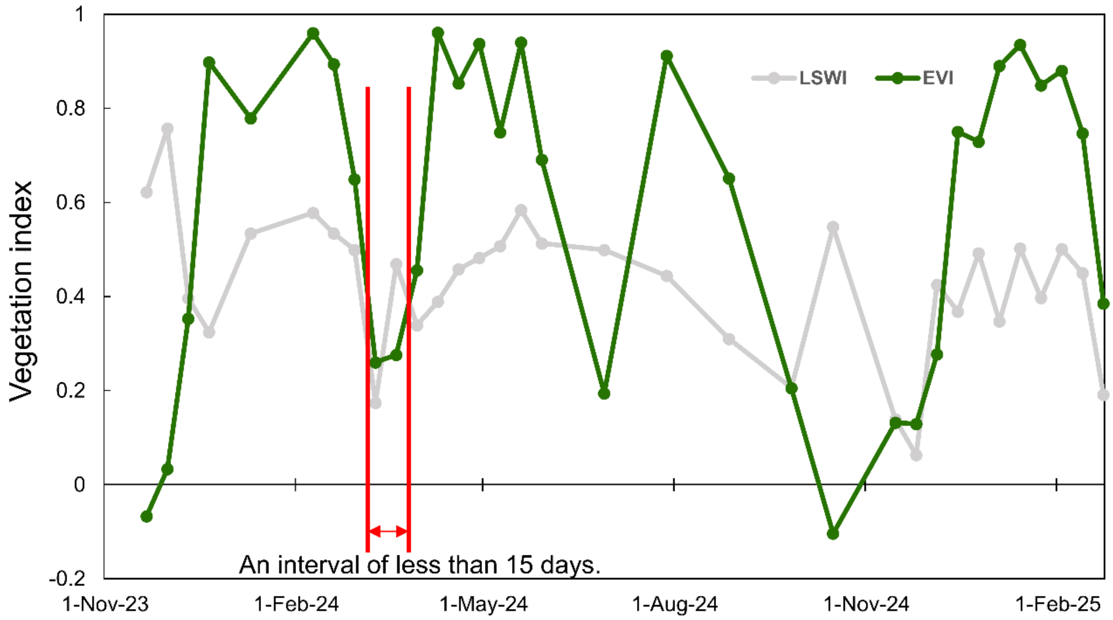


Figure S2. Time series of Land Surface Water Index (LSWI) and Normalized Difference Vegetation Index (NDVI) for a representative rice pixel in the Mekong Delta, Vietnam, from November 2023 to February 2025 with a red vertical line indicating an interval of less than 15 days.

2. Section 2 states that the data source is based on five-year median composite imagery. However, rice planting patterns in South and Southeast Asia may exhibit considerable

interannual variability, including crop rotation and alternating cultivation of different crops. It is unclear whether compositing over five years could obscure such variability and potentially affect the reliability of rice mapping. Please provide quantitative or statistical evidence to demonstrate that the five-year compositing strategy does not significantly distort the spatial distribution of rice. Without such justification, the robustness of the mapping methodology and the reliability of the resulting dataset remain questionable.

Response:

We greatly appreciate the reviewer's careful attention to the five-year compositing strategy, as this is indeed a critical aspect affecting the robustness of the final dataset.

Our decision to adopt a five-year median composite approach was primarily motivated by data availability constraints in tropical monsoon regions, particularly during the 1990s. Although annual rice mapping would be ideal, the number of valid Landsat observations within a single year—especially under persistent cloud cover across South and Southeast Asia—is often insufficient to reliably capture rice flooding signals and key phenological stages.

To explicitly address this concern, we have added a new analysis comparing rice mapping results derived from single-year imagery and multi-year composited imagery for the mid-1990s. As shown in Figure 9, rice maps generated using individual years (e.g., 1993, 1994, 1995, 1996, and 1997 separately) suffer from severe omission errors due to missing observations. In some regions, rice fields are barely detected in single-year results. In contrast, the five-year composite substantially improves spatial completeness and consistency, reducing uncertainty caused by observational gaps.

The main text has been revised accordingly, as follows (Lines 369-376): *"Another aspect that merits consideration is the effect of multi-year image compositing. As shown in the comparison with NESEA-Rice10 (Fig. 5), although both studies employ flooding-signal-based approaches, our results identify a larger number of rice pixels in some regions, mainly due to the increased availability of valid observations.*

This effect is particularly evident for the 1990s. As illustrated in Fig. 8, rice maps derived from single-year imagery show pronounced spatial instability, especially for 1995, owing to *severe limitations in observation frequency. In contrast, the five-year composite produces a more spatially coherent and reliable rice distribution. Under such data-limited conditions, the uncertainty associated with single-year mapping exceeds the potential smoothing effect introduced by multi-year compositing."*

Furthermore, as illustrated by the Landsat imagery shown in **Figure 9**, the spatial extent of flooding signals observed in August during 1993–1997 is consistent with the rice distribution captured in our corresponding rice product, providing additional qualitative support for the validity of the composite-based results.

Finally, we note that multi-year compositing has been widely adopted in other long-term land-cover and rice-mapping studies to address historical data limitations. For example, the **GLC_FCS30** global land-cover product employs five-year compositing to improve classification stability in early periods with sparse observations. Similarly, long-term rice mapping studies in Japan published in *ISPRS Journal of Photogrammetry and Remote Sensing* have adopted multi-year aggregation strategies to reconstruct historical paddy rice dynamics under limited data availability. These precedents support our conclusion that, under severe observational constraints, multi-year compositing introduces less uncertainty than single-year mapping.

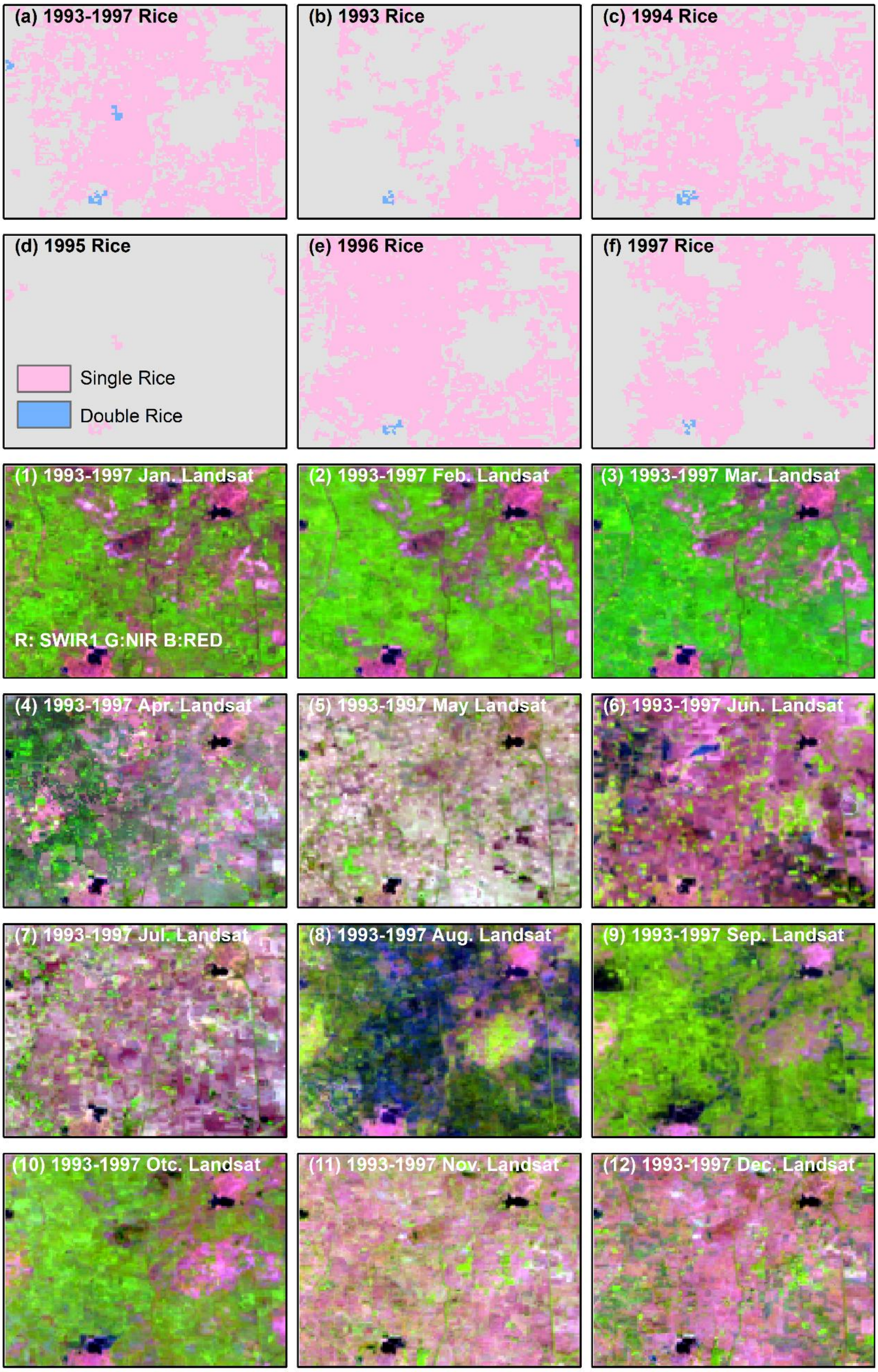


Figure 9. Comparison of rice maps derived from single-year imagery and five-year composite

imagery in the mid-1990s.

3. It is recommended that the process of collecting and labeling the sample point validation data be re-described and placed in Section 2. Describing point-based and statistic-based validation data in one section can improve readability.

Response:

We agree with the reviewer that reorganizing the validation description improves the overall readability of the manuscript.

In the revised version, we have moved and consolidated the description of sample collection, labeling, and validation procedures into **Section 2 (Materials and Methods)**. Specifically, point-based validation, comparisons with existing rice products, and validation against FAO statistical data are now described together in a unified subsection titled **"Accuracy assessment and comparisons"**.

The main text has been revised accordingly, as follows (Lines 189-218):

*"**2.5 Accuracy assessment and comparisons***

Validation samples. Validation samples for 2024 were primarily derived through visual interpretation using multiple complementary data sources. First, false-color composites of Sentinel-2 imagery (R/G/B = SWIR1, NIR, RED) were generated to represent different rice growth stages……

Based on these two sets of composites, rice sample points were manually labeled and further cross-validated using ultra–high-resolution Google Earth imagery and the Global Geo-Referenced Field Photo Library…..

Rice products and statistical data. To evaluate the reliability of our mapping results, we conducted spatial comparisons with several established rice mapping products, including: ……. "

4. When accessing the interactive map through Google Earth Engine (GEE), noticeable mosaic

like artifacts appear in many mapped areas, and the boundaries between rice and cropland classes show poor spatial consistency. Please explain the potential causes of these mapping quality issues. In addition, I believe that a published dataset should not only demonstrate high accuracy in statistical metrics or sample-based validation, but should also provide sufficient spatial consistency and reliability to support downstream applications and reuse by the research community.

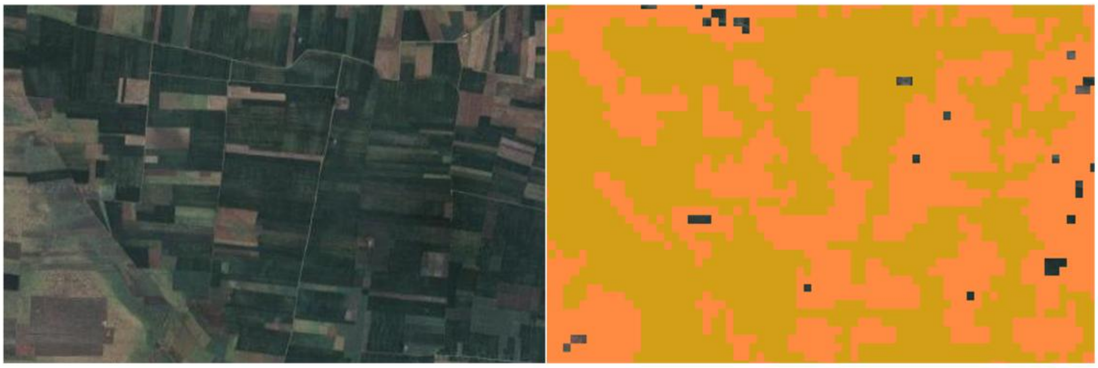


Figure C1 Imagery in 2024 and corresponding mapping results

Response:

We sincerely thank the reviewer for carefully examining the interactive GEE map and for raising concerns regarding spatial consistency, which we fully agree is essential for a published dataset.

First, we would like to clarify that the different colors displayed in the interactive map represent **rice cropping intensity** (i.e., the number of rice seasons per year), rather than different crop types. Thus, some visually abrupt color changes may reflect rice pixel-level differences in cropping intensity rather than classification noise.

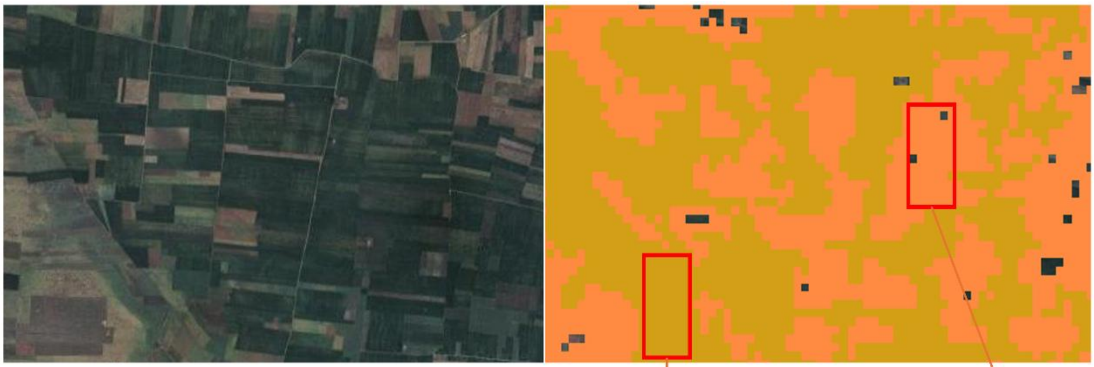


Figure C1 Imagery in 2024 and corresponding mapping results

That said, we do not exclude the possibility of poor spatial consistency. Because the reviewer did not provide the exact geographic coordinates of Figure C1, we summarize the most likely causes as follows.

1. **Uneven observation frequency.**

Despite multi-year compositing, some pixels still have relatively few valid observations due to persistent cloud cover. As our algorithm strictly identifies pixels that exhibit flooding signals during crop growth cycles over multiple years, missing observations in certain seasons may lead to partial omission or uncertainty in cropping intensity estimation.

2. **Spatial resolution and mixed-pixel effects.**

We provide an example from a rice-growing area in Kamphaeng Phet Province, Thailand. While our results show imperfect alignment with field boundaries when compared with ultra–high-resolution imagery, this limitation is inherent to 30-m Landsat data. In many parts of Southeast Asia, agricultural fields are narrow and fragmented, making it difficult for 30-m pixels to precisely delineate parcel boundaries. Mixed-pixel effects at field edges are therefore unavoidable and cannot be fully resolved given the spatial resolution of historical Landsat imagery.

To further address this concern, we have added a comparative visualization in GEE that displays our product alongside other widely used rice datasets (e.g., NESEA-Rice10 and Open-SEA-Rice-10). This comparison demonstrates that, aside from expected differences arising from sensor type (optical vs. SAR) and rice system definitions (paddy vs. rainfed), our product exhibits comparable spatial consistency at both regional and local scales.

https://zz-cloud-storage-and-computing.projects.earthengine.app/view/comparea

5. The qualitative comparison results presented in Figures 4, 5, and S2 do not include corresponding remote-sensing images, making it difficult to visually assess the reliability of the mapping outcomes. Moreover, the figures mainly show large-scale regional patterns, while finer spatial details are not adequately displayed. It is recommended to include zoomed-in comparisons with reference imagery to better demonstrate the spatial accuracy and consistency of the rice mapping results.

Response:

We appreciate this important suggestion and agree that the qualitative comparison figures can be substantially improved by incorporating reference imagery and finer spatial details.

In the revised manuscript, we have comprehensively redesigned the comparison figures as follows:

1. **Figure 4** has been updated to include both regional-scale comparisons and field-scale zoom-ins. The rightmost panels now include corresponding Sentinel-1 and Landsat imagery using appropriate false-color composites to visually validate the rice mapping results.

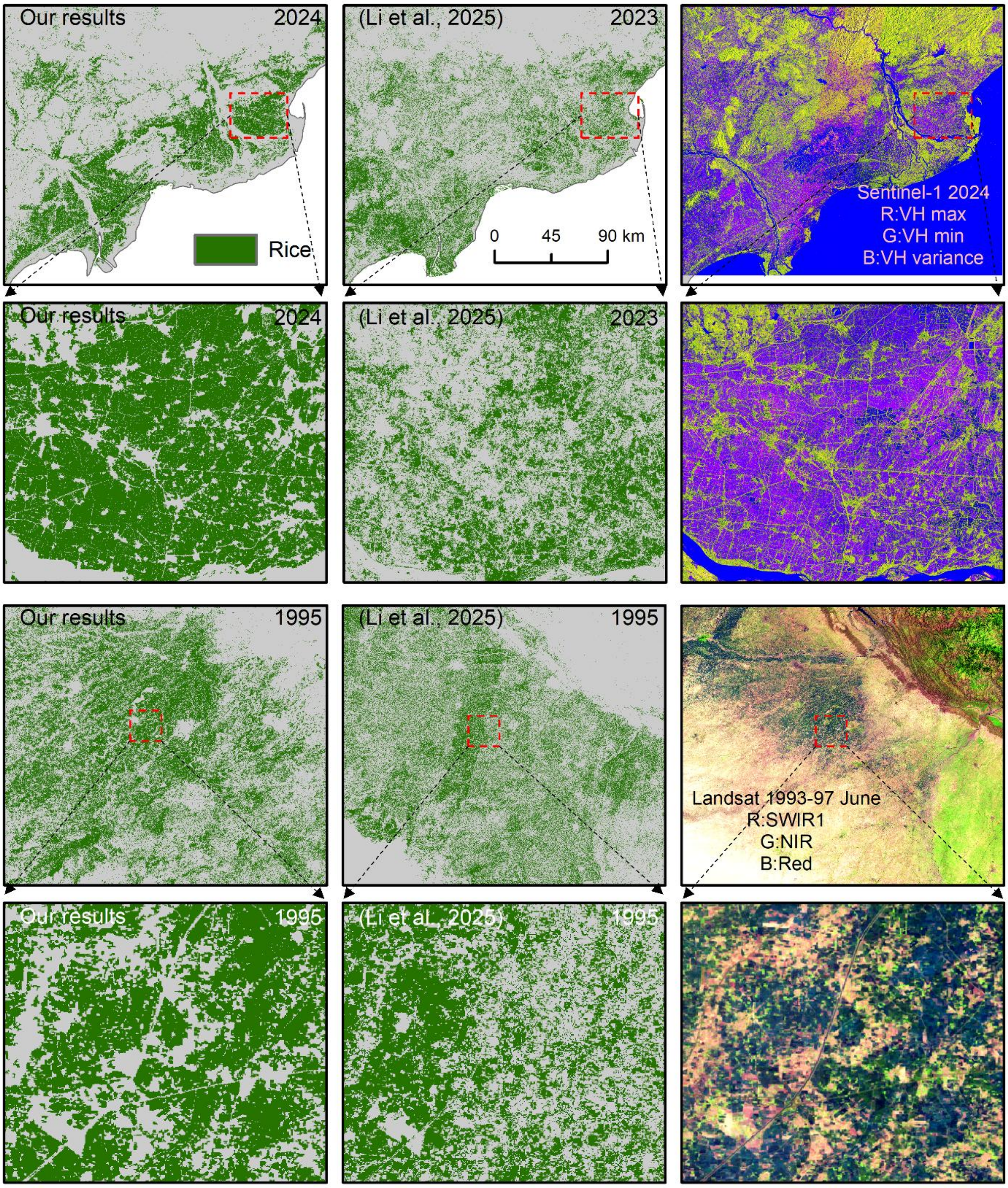
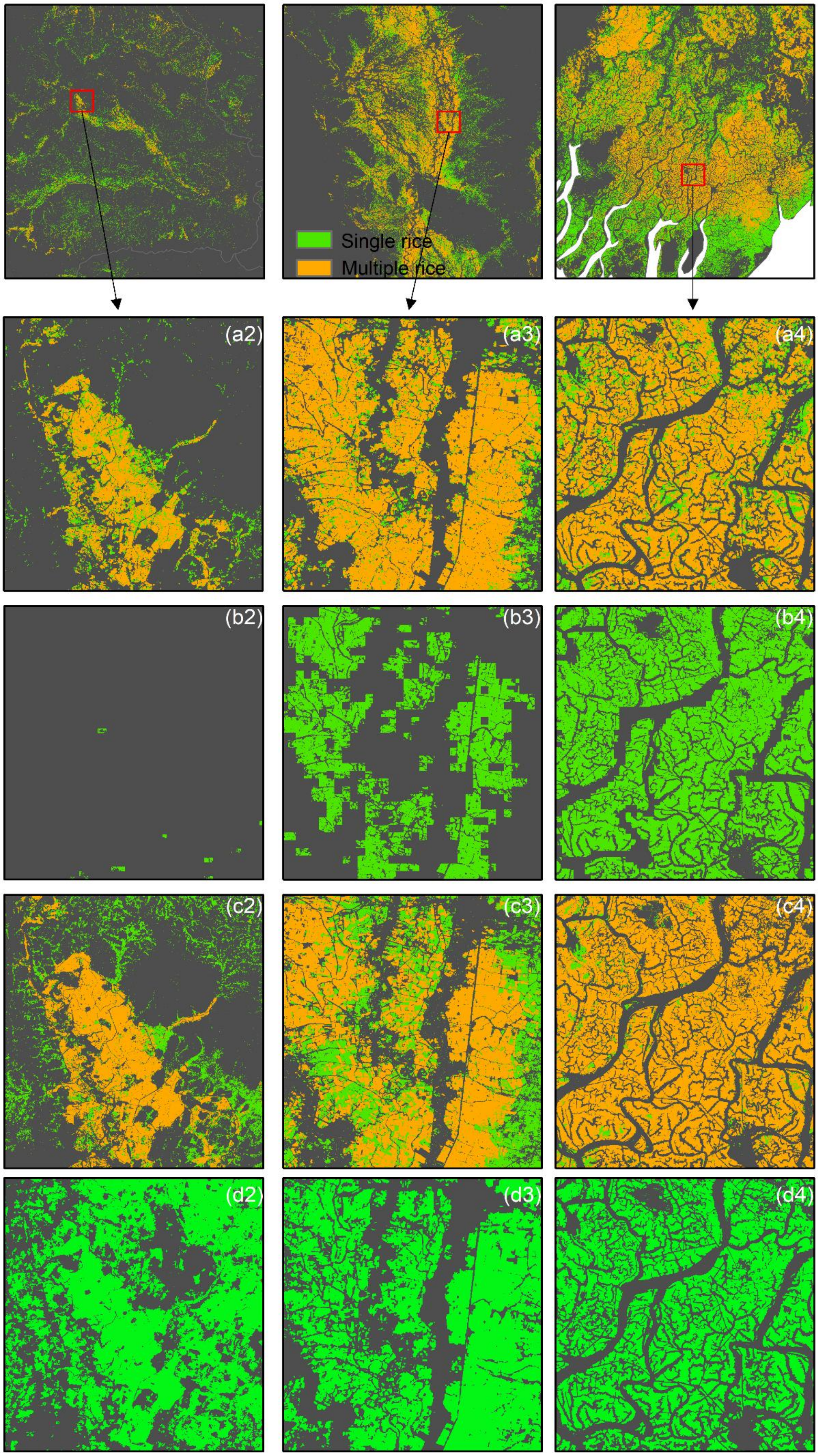


Figure 5. Comparison of rice distribution maps: Our Results compared to Li et al. (2025) for South Asia (1995, 2024)

2. **Figure 5** presents a comparative analysis of single- and multi-season rice field distribution in Southeast Asia. To avoid excessive visual complexity, we have added a new supplementary figure focusing specifically on field-level details. This figure demonstrates strong spatial agreement between our results and Open-SEA-Rice-10 at the parcel scale, particularly for multi-season rice systems, while also highlighting omission errors in some single-year optical products (e.g., NESEA-Rice10) caused by data availability limitations.



Single rice
Multiple rice
(a2)
(a3)
(a4)
(b2)
(b3)
(b4)
(c2)
(c3)
(c4)
(d2)
(d3)
(d4)

Figure S5. Field-scale comparison of rice distribution maps corresponding to Figure 5. Panel labels are consistent with those in Figure 5.

3. For **Figure S2**, in addition to the original country-scale comparison for Vietnam, we now include zoomed-in comparisons over the Mekong Delta and the Red River Delta, together with corresponding Sentinel-1 SAR imagery. These comparisons show strong consistency between our mapped rice areas and radar-derived flooding signals.

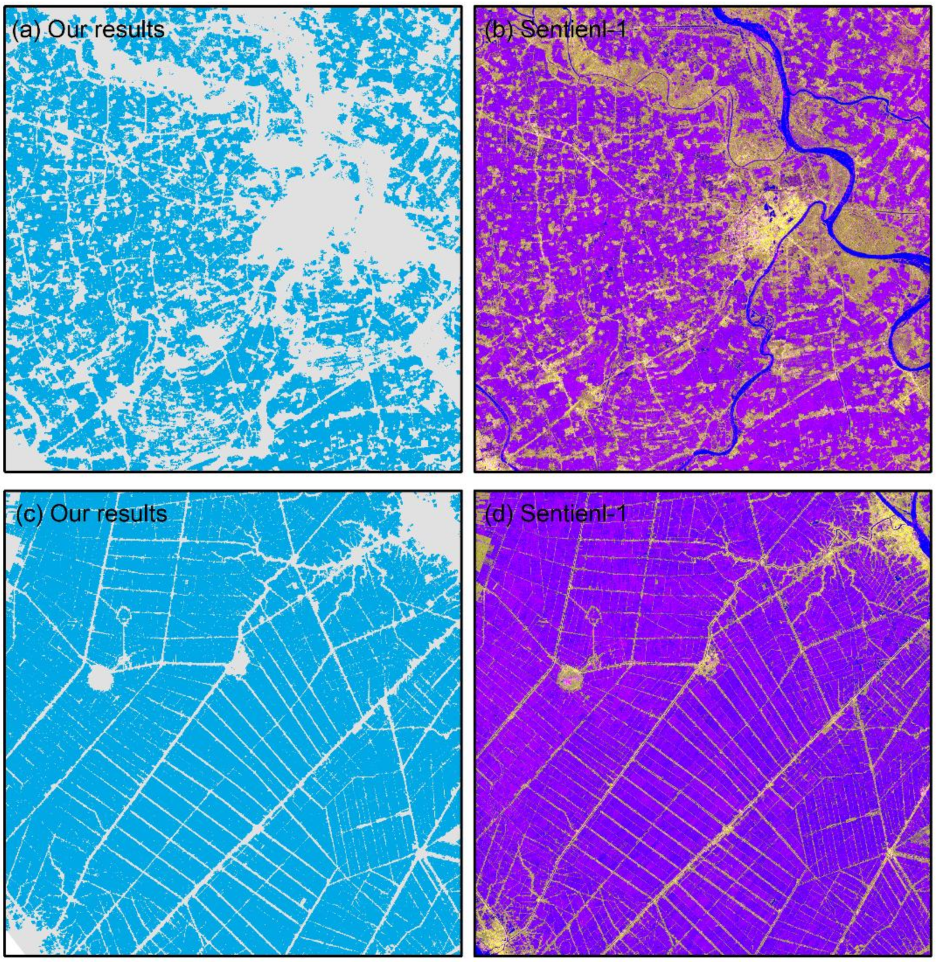


Figure S3. Spatial comparison of rice distribution maps and reference imagery over selected regions in Vietnam, highlighting consistency between optical- and SAR-based rice detection.

Response to Referee 2

This manuscript presents a 30 m resolution dataset of paddy rice spatial distribution and cropping intensity across South and mainland Southeast Asia for the years 1995, 2005, 2015, and 2024, derived from multi-year composites of Landsat and Sentinel-2 imagery using a refined phenology-based approach. The manuscript is generally well-structured, and the authors have compared their results with several existing rice mapping products as well as FAO statistics. However, there are several methodological and validation concerns that need to be addressed.

Response to Reviewer

We sincerely thank the reviewer for their careful reading of our manuscript and for the constructive and insightful comments. These suggestions have been extremely helpful in improving the validation framework and methodological robustness of our work. Below, we provide point-by-point responses to each comment. All changes have been incorporated into the revised manuscript.

All revisions and responses are highlighted in blue font in the response letter. Modifications in the main manuscript are indicated using italicized text enclosed in quotation marks.

1.The manuscript reports validation using 23,396 samples across four decades. However, only the 2024 samples were manually interpreted, and earlier-year labels were transferred using a sample migration algorithm. While this algorithm is reasonable for constructing historical training data, the migrated samples may not be considered fully independent validation data, as they mainly represent long-term stable pixels and may not capture areas of actual land-use change, such as newly cultivated or abandoned fields. As a result, using them to assess accuracy in earlier years could overestimate algorithm performance.

Response:

Thank you for raising this important concern. We agree that validation samples derived from the sample migration algorithm may introduce potential bias, as they mainly represent long-term stable pixels and may not adequately capture areas experiencing land-use change. Consequently, relying solely on migrated samples for accuracy assessment could lead to an optimistic estimation of classification performance in earlier years.

To address this issue, we additionally generated an independent validation dataset through manual visual interpretation for three representative years (1995, 2005, and 2015). The validation samples were interpreted using false-color composites of Landsat imagery to ensure reliable identification of rice and non-rice areas. Through this process, a total of 2,371 independent validation samples were collected, including 1,330 rice samples and 1,041 non-rice samples.

These samples were used exclusively for independent accuracy assessment of the corresponding years. The confusion matrices have been provided in the Appendix for transparency.

Table S1. Confusion matrices for the independent validation of rice mapping results in 1995, 2005, and 2015.

Year	Overall Accuracy	User Accuracy	Producer Accuracy	F1 Score
2015	0.8018	0.8785	0.7774	0.8249
2005	0.8748	0.8966	0.8000	0.8456
1995	0.7890	0.9675	0.6980	0.8109

The main text has been revised accordingly, as follows (Lines 207-210): *"To ensure the reliability of the validation results, an additional set of 2,371 validation samples was independently generated through manual visual interpretation for the years 1995, 2005, and 2015. These samples were used solely for independent accuracy assessment, and the corresponding validation results are presented in the Appendix (Table S1)."*

2.The dataset is based on half-monthly NDVI/LSWI/EVI composites and multi-year (3-5 year) compositing to represent individual reference years. While this strategy likely improves data availability and reduces noise, it may also smooth short-duration cropping signals or interannual variability, particularly in regions with triple cropping or short-cycle varieties. Providing additional assessment of temporal aggregation effects would improve confidence in the annual representativeness of the product.

Response:

We thank the reviewer for this insightful comment regarding the potential effects of temporal aggregation on cropping signals and interannual variability.

The multi-year compositing strategy was adopted mainly to address the extremely limited number of valid observations in earlier years, especially during the 1990s. In tropical South and Southeast Asia, persistent cloud cover often results in very sparse valid observations within a single year, which makes reliable detection of the flooding signal difficult. Aggregating observations across several years substantially increases observation density and improves the probability of capturing flooding events.

To evaluate the potential influence of this strategy, we conducted an additional comparison between single-year mapping results and multi-year composite results for the mid-1990s. As shown in Fig. 9, rice maps derived from individual years (1993–1997) exhibit strong spatial instability and severe omission errors due to missing observations. In contrast, the multi-year composite produces a more spatially coherent and complete rice distribution. Under such data-limited conditions, the uncertainty associated with single-year mapping exceeds the potential smoothing effects introduced by multi-year compositing.

Importantly, the flooding signals observed in Landsat imagery during the transplanting season (Fig. 9 (8)) are spatially consistent with the rice distribution captured in our product, providing additional qualitative support for the validity of the composite-based approach.

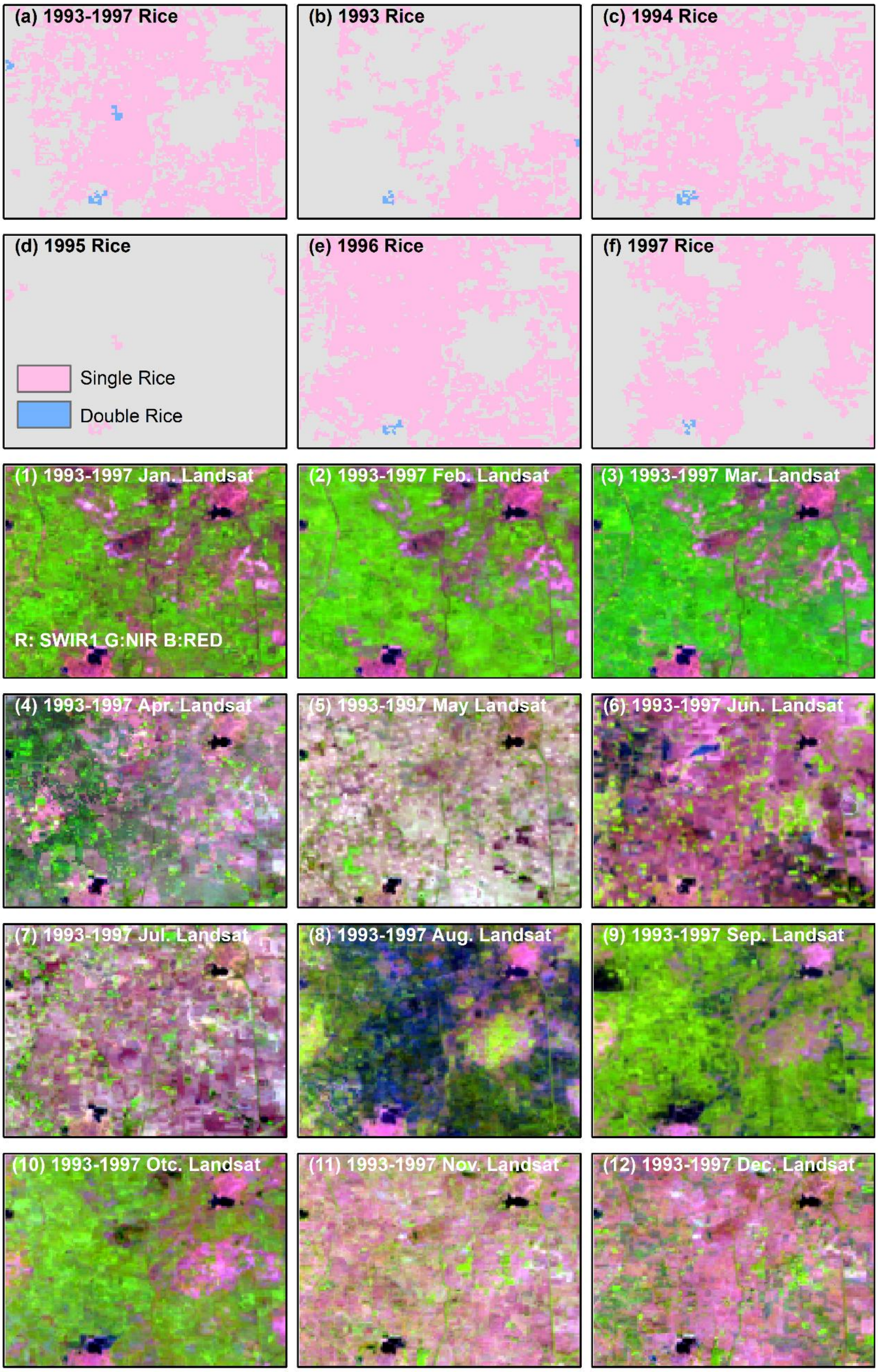
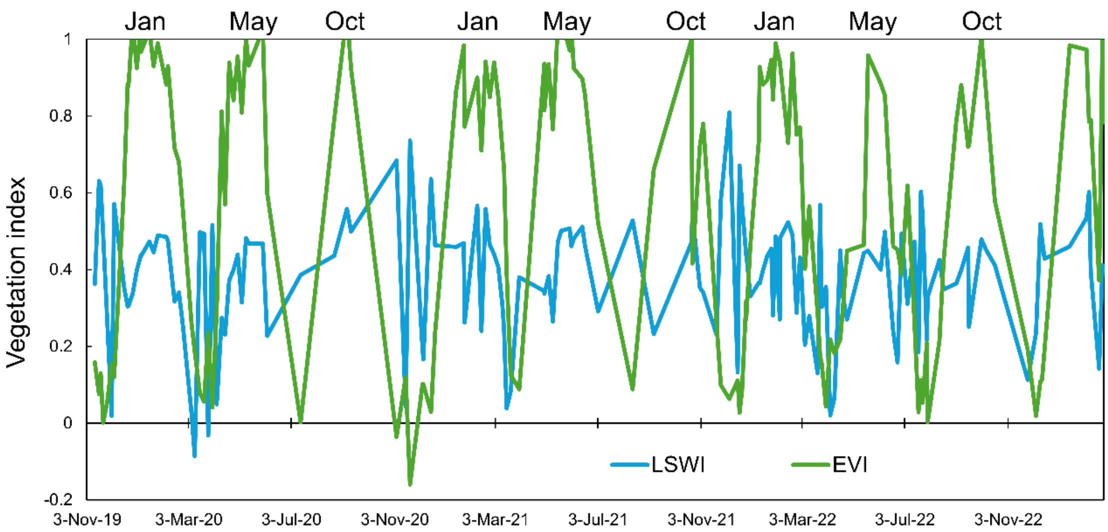


Figure 9. Comparison of rice maps derived from single-year imagery and five-year composite

imagery in the mid-1990s.

In addition, we extracted multi-year rice samples located in the **Mekong River Basin** using Sentinel-2 imagery to further examine the temporal stability of rice phenology. The time-series curves of vegetation indices from **November 2019 to March 2023** (see figure below) indicate that the phenological stages of rice cultivation remain relatively consistent across years. Even in regions with the highest cropping intensity, the timing of key phenological phases shows only minor interannual variation. Therefore, the multi-year compositing strategy represents a reasonable trade-off between data availability and mapping accuracy under conditions of limited historical observations. The sample location used for this analysis is Longitude: 105.391; Latitude: 10.299.



Finally, we note that multi-year compositing has been widely adopted in long-term land-cover and rice mapping studies to address historical data limitations. For example, the GLC_FCS30 global land-cover dataset uses five-year compositing in early periods to improve classification stability (Zhang et al., 2024). Similar strategies have also been applied in long-term rice mapping studies (Carrasco et al., 2022) in Japan published in *ISPRS Journal of Photogrammetry and Remote Sensing*. These precedents suggest that, under severe observational constraints, multi-year compositing introduces less uncertainty than single-year mapping.

The main text has been revised accordingly, as follows (Lines 369-376): *"Another aspect that merits consideration is the effect of multi-year image compositing. As shown in the* *comparison with NESEA-Rice10 (Fig. 5), although both studies employ flooding-signal-based approaches, our results identify a larger number of rice pixels in some regions, mainly due to the increased availability of valid observations.*

This effect is particularly evident for the 1990s. As illustrated in Fig. 8, rice maps derived from single-year imagery show pronounced spatial instability, especially for 1995, owing to severe limitations in observation frequency. In contrast, the five-year composite produces a more spatially coherent and reliable rice distribution. Under such data-limited conditions, the uncertainty associated with single-year mapping exceeds the potential smoothing effect introduced by multi-year compositing."

3. The manuscript introduces the "enhanced peak detection" and "false peak elimination" methods. It is suggested to include quantitative comparisons with existing phenology-based approaches.

Response:

Thank you for this valuable suggestion. We agree that demonstrating the effectiveness of the enhanced peak detection and false peak elimination steps is important for clarifying the advantages of our phenology-based framework.

To address this comment, we conducted an additional ablation experiment that directly compares the RiceAtlas phenological calendar (as a baseline phenology-based approach using fixed transplanting windows) with our proposed crop peak detection method. Specifically, the rice flooding detection rule (LSWI + 0.05 ≥ EVI) was applied separately within each of the two-time windows (the fixed transplanting windows derived from RiceAtlas and the dynamically detected crop peaks from our method to monitor rice cropping intensity.

The results indicate that fixed phenological windows may lead to misclassification in flood-prone environments. During seasonal flooding events, LSWI can temporarily exceed EVI in non-rice areas, triggering false flooding signals when predefined transplanting windows are used. By contrast, our method first identifies valid crop peaks from the vegetation index time series and then restricts flooding detection to the corresponding growth cycles, effectively reducing such false detections. The comparison (Fig. 10) shows that the RiceAtlas-based approach overestimates multi-season rice in flood-affected regions, whereas our peak-detection-based framework produces more realistic cropping intensity patterns.

A corresponding figure and discussion have been added to the revised manuscript (Section 4.2).

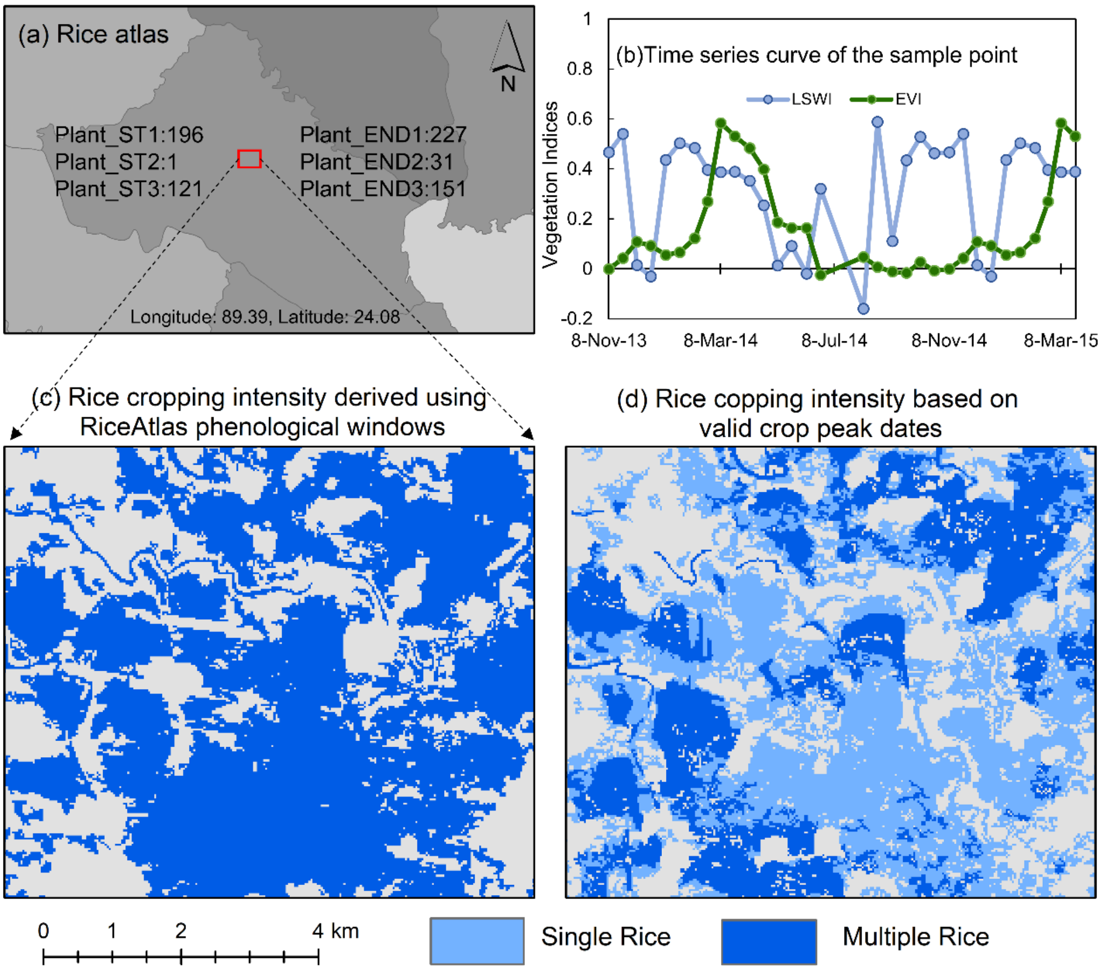


Fig. 10. Time-series comparison between the RiceAtlas-based approach and the proposed peak-detection framework at an example location affected by seasonal flooding.

The main text has been revised accordingly, as follows (Lines 380-392): *"Example: limitation of fixed phenology windows. Another advantage of the proposed framework is the use of dynamically detected crop peaks rather than predefined phenological calendars. To illustrate this point, we conducted a comparison using transplanting windows derived from the RiceAtlas phenology database (hereafter referred to as the RiceAtlas approach).*

As shown in Fig. 10, the RiceAtlas approach relies on fixed planting windows (e.g., DOY

196–227 for the first season in the example location). However, during the monsoon period, temporary flooding events may cause LSWI to exceed EVI even in non-rice areas. When flooding detection is applied within these predefined windows, such transient signals can be incorrectly interpreted as rice transplanting events.

In contrast, the proposed method first identifies valid crop peaks from the vegetation index time series and then restricts flooding detection to phenologically consistent crop growth cycles. This strategy effectively reduces the influence of seasonal flooding signals. In the example shown in Fig. 10, the RiceAtlas-based approach identifies an additional rice season, leading to an overestimation of cropping intensity, whereas our method avoids this flood-induced overestimation."

We would like to clarify that the "enhanced peak detection" and "false peak elimination" procedures used in this study were adopted from our previous work (Yang et al., 2023), where the land surface phenology detection framework was developed and quantitatively compared with other phenology detection approaches. Therefore, the focus of the present study is not to compare different phenology algorithms, but to apply this validated framework to improve paddy rice mapping and cropping intensity monitoring.

4. Key parameters, including Whittaker smoothing (λ = 300), minimum season length (120 days), and many thresholds, are applied uniformly across all regions, years, and management regimes. Such uniform choices may fail to capture short-duration varieties or respond appropriately to differences in irrigation practices and cropping calendars. These parameters raise concerns regarding the method's transferability across regions, years, and management conditions.

Response:

We thank the reviewer for the careful examination of the parameter settings used in this study. We agree that demonstrating the generality and rationality of these thresholds is important. Therefore, we provide additional explanations for all empirical thresholds adopted in this study.

Whittaker smoothing (λ = 300): To address the concern regarding the choice of the smoothing parameter λ in the Whittaker smoothing algorithm, we have conducted a sensitivity analysis using λ values of 100, 300, 500, 1000, and 2000.

As shown in the newly added Figure S7 (Appendix), although the degree of smoothness varies with λ, the detected peak date remains consistently stable at 26 February across all tested values. Differences in λ are more likely to affect the identification of certain phenological transition dates, such as the start of season (SOS) and the end of season (EOS). However, our study does not rely on extracting precise phenological stages. Instead, we only identify the relative peak date within each rice growth cycle and construct a fixed temporal window around this peak for subsequent classification.

Because the timing of the maximum NDVI value (the key feature we use) shows no meaningful shift within the tested range, minor changes in curve shape caused by different λ values do not alter the position of the peak or the resulting temporal window. Therefore, as long as λ falls within a reasonable range (e.g., 300–500), the smoothing process does not substantially influence the final accuracy of the rice mapping results. We ultimately adopted λ = 300 in the main analysis.

The comparison figure and the corresponding explanation have been added to the Appendix (Appendix S7) for transparency.

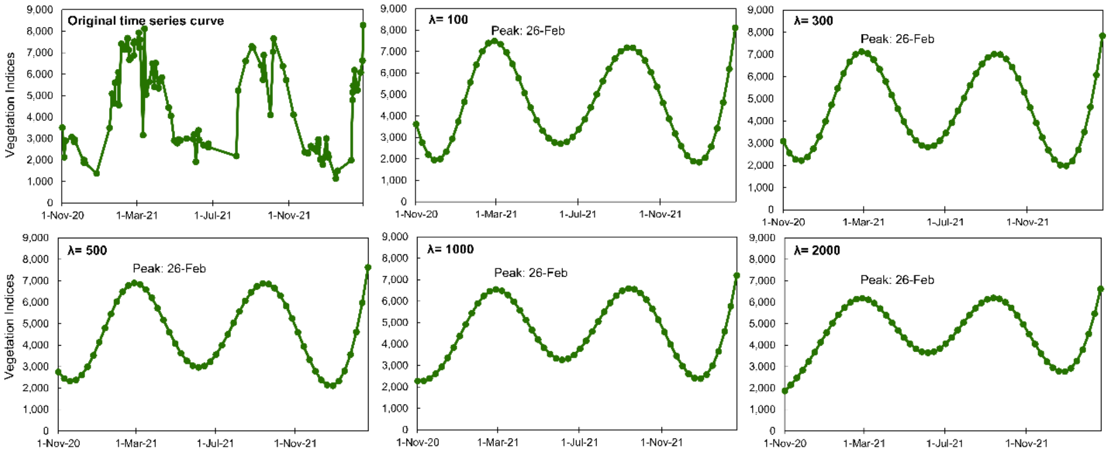


Figure S7. Whittaker smoothing sensitivity analysis on NDVI time series curve (λ = 100, 300, 500, 1000, 2000). Note: In the intermediate processing step, the original NDVI values were multiplied by 1000 (i.e., the plotted vegetation indices represent NDVI × 1000).

The main text has been revised accordingly, as follows (Lines 141-143): *"Subsequently,* *the NDVI time series was smoothed using the Whittaker smoothing algorithm with a smoothing parameter λ = 300 (see the Appendix for details on the selection of the λ threshold) to generate a smoothed time series."*

Appendix has been revised accordingly, as follows: *"To evaluate the sensitivity of the Whittaker smoothing algorithm to the smoothing parameter λ, a comparative analysis was conducted using λ values of 100, 300, 500, 1000, and 2000. As shown in the figure, although the degree of smoothness varies noticeably with λ, the detected peak date remains consistently stable at 26 February across all tested values.*

Differences in λ are more likely to affect the identification of certain phenological metrics, such as the start of season (SOS) and the end of season (EOS). However, this study does not rely on extracting precise phenological transition dates. Instead, we only identify the relative peak date within each rice growth cycle and then construct a fixed temporal window around this peak for subsequent classification.

Because the timing of the maximum NDVI value (the key feature we use) shows no meaningful shift within the tested range, minor changes in curve shape caused by different λ values do not alter the position of the peak or the resulting temporal window. Therefore, as long as λ falls within a reasonable range (e.g., 300–500), the smoothing process does not substantially influence the final accuracy of the rice mapping results. The value λ = 300 was ultimately adopted in this study."

120days: We appreciate the reviewer's constructive comment regarding the 120-day threshold. This value was deliberately chosen to retain only biologically plausible and complete rice growth cycles in South and mainland Southeast Asia, following the trough-based cycle decomposition method adapted from Yang et al. (2023).

Recent phenology-based rice mapping studies in mainland Southeast Asia explicitly adopt "around 120 days" as the representative length of a full rice growth cycle for algorithm design. For example, Ginting et al. (2025) state: "*this study's phenology approach requires observing the entire rice growth cycle (around 120 days)*" and further note that "*the length of one rice season is approximately 4 months for irrigated rice fields … while both sugarcane and cassava have a longer season of around 10 months.*"

We acknowledge that short-duration rice varieties (<120 days) can occur in certain regions and management systems. However, in satellite-derived vegetation index time series, the observable phenological cycle (including the full development and decline of canopy signals) often spans approximately four months. Consequently, shorter intervals detected in the time series are more frequently associated with noise, partial growth signals, or overlapping cropping cycles in multi-cropping systems. The 120-day threshold therefore serves as a conservative quality-control constraint to reduce false peaks while capturing the majority of realistic rice seasons.

Nevertheless, we recognize that parameter optimization may further improve transferability across regions and management regimes. Future work will explore region-specific or adaptive parameterization to better accommodate shorter growth cycles where they occur.

Other empirical thresholds: We appreciate the reviewer's constructive comment regarding the empirical thresholds. The thresholds adopted in this study (LSWI < 0.05, EVI ≥ 0.35, and LSWI + 0.05 ≥ EVI) have been repeatedly validated as reasonable and effective in numerous previous remote sensing studies for rice paddy identification and crop discrimination. These thresholds, originally proposed and extensively tested in the literature, are widely recognized as robust for distinguishing flooded rice fields from other vegetation and non-vegetated surfaces.

To address the reviewer's concern and enhance transparency, we have now added several key references in the revised manuscript that explicitly support the rationale behind these thresholds, drawing from prior validation studies in the literature. We believe these well-established expert-derived values remain appropriate for our framework. The main text has been revised accordingly, as follows (Lines 167-173): *"Subsequently, flood signal detection was performed on the imagery data within this time window by applying the rule LSWI + 0.05 ≥ EVI (Xiao et al., 2005; Xiao et al., 2006; Dong and Xiao, 2016). Notably, the LSWI and EVI used here differ from those employed in crop peak identification, as these vegetation indices (VIs) do not require smoothing. Additionally, to minimize interference from factors such as* *precipitation and soil background value, we applied the rule that, when a flood signal is detected, EVI must be ≤ 0.35 and LSWI must be > 0.05 (Zhao et al., 2025). Pixels meeting these conditions were classified as rice (pixel value = 1).”*

5. The manuscript constructs custom composites of Landsat and Sentinel-2 imagery instead of using the Harmonized Landsat Sentinel-2 (HLS) product. It is suggested to clarify whether this choice affects paddy rice spatial distribution and cropping intensity detection.

Response:

We thank the reviewer for this valuable suggestion. The Harmonized Landsat Sentinel-2 (HLS) product provides high-quality analysis-ready surface reflectance data by harmonizing observations from Landsat and Sentinel-2 sensors. However, HLS is only available from 2013 onward for Landsat (L30) and 2015 onward for Sentinel-2 (S30). Since our dataset spans 1995-2024, and the earlier periods (e.g., 1995 and 2005) rely primarily on Landsat 5 and Landsat 7 imagery, the HLS product cannot support the full temporal coverage required for this study.

To ensure algorithmic consistency across all study years, we constructed a unified preprocessing and compositing workflow for Landsat and Sentinel-2 imagery rather than combining different data products. This approach allows the same processing strategy to be applied throughout the entire time series, thereby minimizing potential inconsistencies in paddy rice mapping and cropping intensity detection.

Nevertheless, we acknowledge the advantages of the HLS product for multi-sensor data integration. In future work, we will consider incorporating HLS datasets to further improve cross-sensor consistency and temporal density where the temporal coverage is suitable.

6. The dataset could be further enhanced by including additional confidence or uncertainty layers, considering potential downstream applications.

Response:

Thank you for this valuable suggestion. Following the reviewer’s recommendation, we generated an observation-based confidence layer to provide users with an estimate of mapping reliability. Specifically, for each pixel we calculated the number of valid optical observations within the 90-day temporal window preceding the detected crop peak, which corresponds to the key period for rice flooding detection. The observation count was then normalized by the total number of compositing intervals to produce a confidence score ranging from 0 to 1. Pixels with more valid observations generally provide more reliable phenological signals for rice identification.

The description of the confidence layer and its calculation has been added to the Methods section (Section 2.4), and the corresponding layer has been included in the released dataset. An example map of the confidence layer has been included in the supplementary materials (Fig. S8).

The main text has been revised accordingly, as follows (Lines 178-185): *"Optical-based rice mapping is often affected by cloud contamination, which reduces the availability of valid observations during critical phenological stages. To support downstream applications and provide an estimate of data reliability, an observation-based confidence layer was generated. For each pixel, the number of valid observations used in the time-series analysis was calculated within the 90-day temporal window preceding the identified crop peak. Pixels with more valid observations generally provide more reliable phenological signals for rice detection; therefore, the observation frequency was used as a proxy for mapping confidence. The observation count was further normalized by the total number of compositing intervals to produce a confidence score ranging from 0 to 1."*

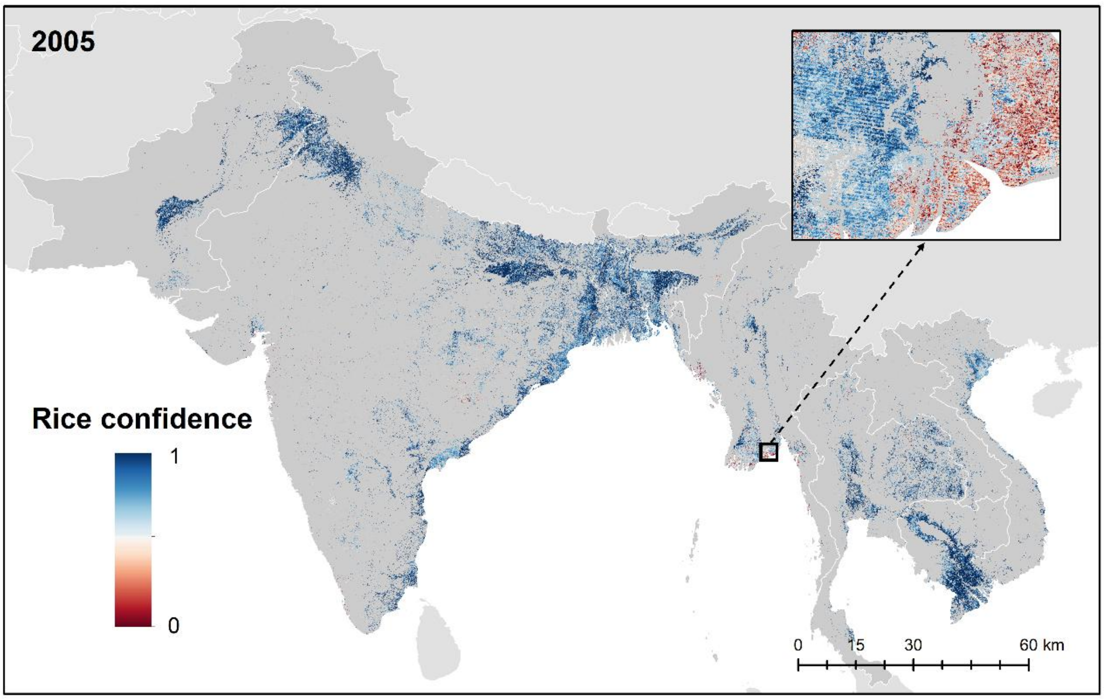


Figure S8. Spatial distribution of rice confidence scores for the year 2005 across South Asia, derived from normalized valid optical observations within the 90-day temporal window preceding crop peak detection. The color scale ranges from 0 (low confidence, red) to 1 (high confidence, blue), with an inset zoom on a selected region for detail.

Ref:

Carrasco, L., Fujita, G., Kito, K., and Miyashita, T.: Historical mapping of rice fields in Japan using phenology and temporally aggregated Landsat images in Google Earth Engine, ISPRS Journal of Photogrammetry and Remote Sensing, 191, 277-289, https://doi.org/10.1016/j.isprsjprs.2022.07.018, 2022.

Dong, J. and Xiao, X.: Evolution of regional to global paddy rice mapping methods: A review, ISPRS Journal of Photogrammetry and Remote Sensing, 119, 214-227, https://doi.org/10.1016/j.isprsjprs.2016.05.010, 2016.

Ginting, F. I., Rudiyanto, R., Fatchurrachman, Mohd Shah, R., Che Soh, N., Eng Giap, S. G., Fiantis, D., Setiawan, B. I., Schiller, S., Davitt, A., and Minasny, B.: High-resolution maps of rice cropping intensity across Southeast Asia, Scientific Data, 12, 1408, 10.1038/s41597-025-05722-1, 2025.

Xiao, X., Boles, S., Frolking, S., Li, C., Babu, J. Y., Salas, W., and Moore, B.: Mapping paddy rice agriculture in South and Southeast Asia using multi-temporal MODIS images, Remote Sensing of Environment, 100, 95-113, https://doi.org/10.1016/j.rse.2005.10.004, 2006.

Xiao, X., Boles, S., Liu, J., Zhuang, D., Frolking, S., Li, C., Salas, W., and Moore, B.: Mapping paddy rice agriculture in southern China using multi-temporal MODIS images, Remote Sensing of Environment, 95, 480-492, https://doi.org/10.1016/j.rse.2004.12.009, 2005.

Yang, J., Dong, J., Liu, L., Zhao, M., Zhang, X., Li, X., Dai, J., Wang, H., Wu, C., You, N., Fang, S., Pang, Y., He, Y., Zhao, G., Xiao, X., and Ge, Q.: A robust and unified land surface phenology algorithm for diverse biomes and growth cycles in China by using harmonized Landsat and Sentinel-2 imagery, ISPRS Journal of Photogrammetry and Remote Sensing, 202, 610-636, https://doi.org/10.1016/j.isprsjprs.2023.07.017, 2023.

Zhang, X., Zhao, T., Xu, H., Liu, W., Wang, J., Chen, X., and Liu, L.: GLC_FCS30D: the first global 30 m land-cover dynamics monitoring product with a fine classification system for the period from 1985 to 2022 generated using dense-time-series Landsat imagery and the continuous change-detection method, Earth Syst. Sci. Data, 16, 1353-1381, 10.5194/essd-16-1353-2024, 2024.

Zhao, Z., Dong, J., Yang, J., Liu, L., You, N., Xiao, X., and Zhang, G.: From rice planting area mapping to rice agricultural system mapping: A holistic remote sensing framework for understanding China's complex rice systems, ISPRS Journal of Photogrammetry and Remote Sensing, 224, 382-397, https://doi.org/10.1016/j.isprsjprs.2025.03.026, 2025.