Reviewer #1

We thank the reviewer for a thoughtful and thorough review of our manuscript. The responses to suggestions and comments are shown in **blue** text. We have highlighted the revised sections and corresponding references in **red** text. The page (P) and line (L) numbers indicated in the response refer to the revised manuscript. The item-by-item responses to all comments are listed below.

Suggestions and comments:

Point 1. Line 174-175: "A spatially stratified sampling strategy based on quadrilateral grids was adopted to mitigate the effects of spatial autocorrelation." Could the authors specify the spatially stratified sampling strategy?

Response: Thank you for your valuable comment. In this study, we adopted a spatially stratified sampling strategy based on quadrilateral grids to effectively mitigate the influence of spatial autocorrelation among samples. Taking Xinyang City in Henan Province as an example (Fig. R1-1), we first divided the study area into 20km*20km quadrilateral grids on the Google Earth Engine platform and examined the monthly Sentinel-2 composite imagery from October to the following June. During the interpretation process, all land cover types were classified into six categories: winter wheat, built-up, water, trees, bare land, and other crops. Based on the proportional coverage of each land cover type within a grid, no more than 10 sample points were randomly selected per grid, thereby avoiding excessive spatial clustering of samples. Specifically, for grids containing wheat fields, 1–8 wheat samples were randomly selected in proportion to the wheat coverage within the grid, along with additional non-wheat samples to maintain category balance. For grids without wheat fields, 1–2 non-wheat samples were randomly selected for each category, including built-up land, trees, water, and bare land. This approach ensured both the spatially uniform distribution of samples and the representativeness of diverse land cover types.

P6, L180-182 in the revised manuscript and Supplementary data: As shown in Fig.R1-1, we take Xinyang City in Henan Province as an example to demonstrate the spatial hierarchical sampling strategy based on quadrilateral grids. We first divided the study area into a series of regular quadrilateral grids within the Google Earth Engine platform and examined the monthly Sentinel-2 composite imagery from October to the following June. During the interpretation process, all land cover types were classified into six categories: winter wheat, built-up, water, trees, bare land, and other crops. Based on the proportional coverage of each land cover type within a grid, no more than 10 sample points were randomly selected per grid, thereby avoiding excessive spatial clustering of samples. Specifically, for grids containing wheat fields, 1–8 wheat samples were randomly selected in proportion to the wheat coverage within the grid, along with additional non-wheat samples to maintain category balance. For grids without wheat fields, 1–2 non-wheat samples were randomly selected for each category, including built-up land, trees, water bodies, and bare land. This approach ensured both the spatially uniform distribution of samples and the representativeness of diverse land cover types.

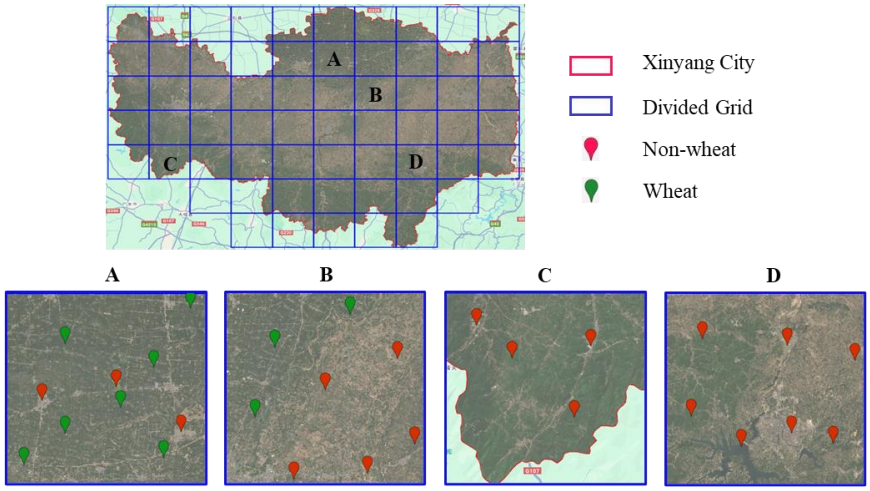

Figure R1-1: Spatial stratified sampling strategy in Xinyang City.

Point 2. Line 177: "more than 50,000 valid sample points were collected annually" How much of the sample points are from the field survey? Also, how did the authors distinguish winter wheat fields from other crop types based on visual inspection? Some examples can be provided.

Response: We sincerely thank the reviewer for the constructive comments. Approximately 2,000–3,000 points were collected from field surveys each year during 2020–2024. These field samples were mainly concentrated in the Huang-Huai-Hai Plain and the middle and lower reaches of the Yangtze River Plain, with a smaller number collected from other provinces. The spatial distribution of these field survey samples is illustrated in Fig.R1-2.

P6, L180-182 in the revised manuscript and Supplementary data: Approximately 2,000–3,000 points were collected from field surveys each year during 2020–2024. These field samples were mainly concentrated in the Huang-Huai-Hai Plain and the middle and lower reaches of the Yangtze River Plain, with a smaller number collected from other provinces. The spatial distribution of these field survey samples is illustrated in Fig.R1-2.

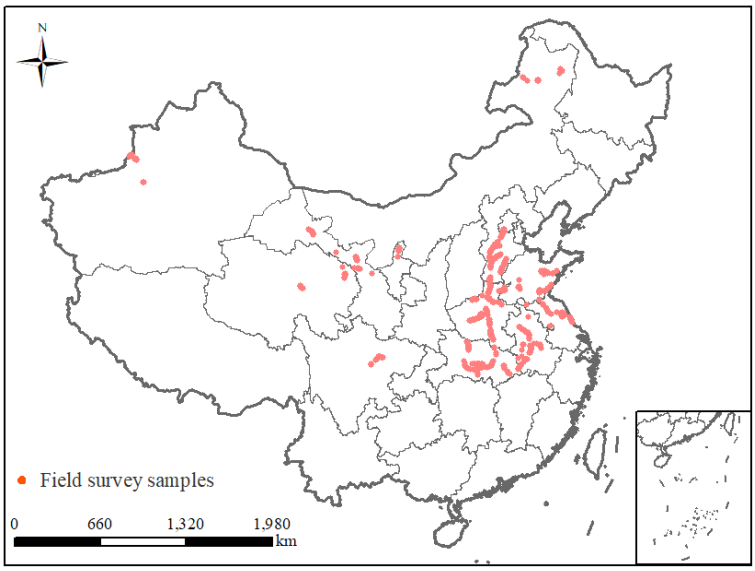

Figure R1-2: Spatial distribution of sample points from field surveys in 2020-2024.

During visual interpretation of remote sensing imagery, winter wheat was distinguished from other crop types based on three main criteria: (1) Temporal dynamics (Fig.R1-3(a)): multi-temporal imagery from October to the following June reveals the complete growth cycle of winter wheat, including sowing, overwintering, regreening, jointing, heading, grain filling, and maturity, while other crops exhibit distinct temporal profiles; (2) Spectral characteristics (Fig.R1-3 (b)): winter wheat demonstrates high vegetation indices during the mid-growing season (around April), appearing as dark green in imagery, whereas rapeseed typically appears yellow and garlic is often light green; (3) Texture and spatial distribution (Fig.R1-3 (b)): winter wheat fields are generally regular in shape and spatially aggregated, in contrast to other land cover types, which often appear fragmented or irregular.

P6, L180-182 in the revised manuscript and Supplementary data: In the process of visual interpretation of remote sensing images, wheat is distinguished from other crop types mainly based on three criteria, taking winter wheat as an example: (1) Temporal dynamics (Fig.R1-3(a)): multi-temporal imagery from October to the following June reveals the complete growth cycle of winter wheat, including sowing, overwintering, regreening, jointing, heading, grain filling, and maturity, while other crops exhibit distinct temporal profiles; (2) Spectral characteristics (Fig.R1-3 (b)): winter wheat demonstrates high vegetation indices during the mid-growing season (around April), appearing as dark green in imagery, whereas rapeseed typically appears yellow and garlic is often light green; (3) Texture and spatial distribution (Fig.R1-3 (b)): winter wheat fields are generally regular in shape and spatially aggregated, in contrast to other land cover types, which often appear fragmented or irregular.

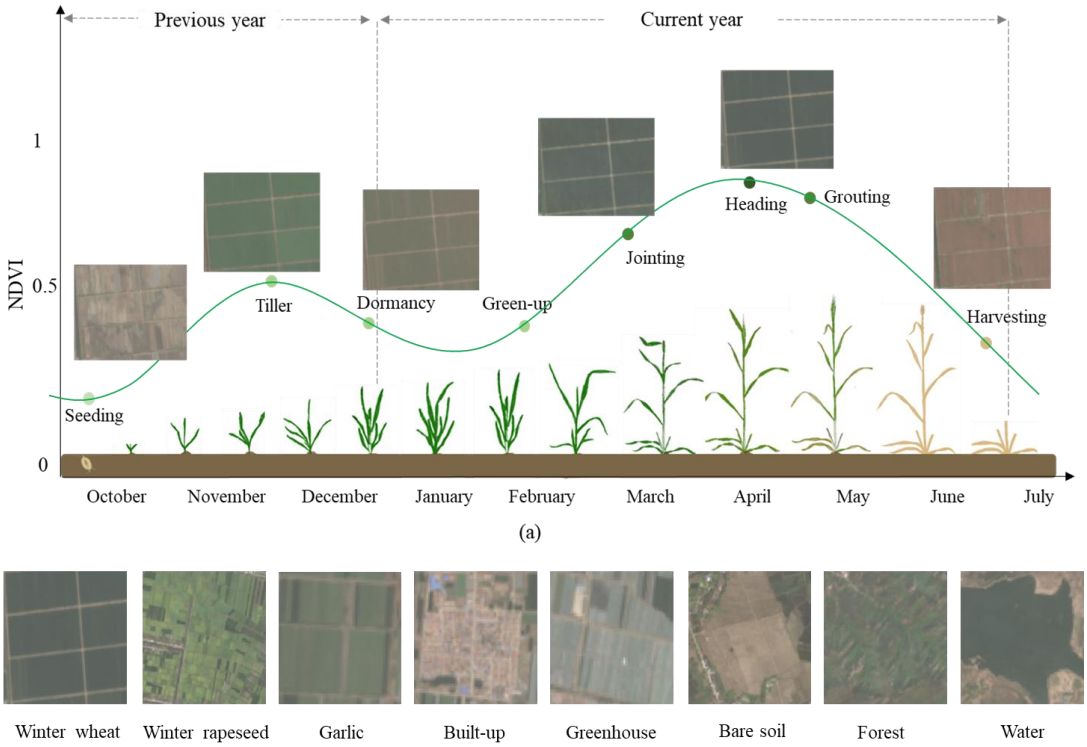

Figure R1-3: Distinguishing winter wheat from other land-cover types based on visual interpretation of remote sensing Images. (a) Phenological spectrum curve of winter wheat and remote sensing image characteristics at different growth stages. (b) Spectral texture characteristics of winter wheat and other land-cover types in April.

Point 3. Line 234-235: "all non-wheat pixels (Section 3.1) were classified into two types: non-wheat winter crops vs. non-winter crops and non-wheat spring crops vs. non-spring crops, according to their respective growth stages": I think there are four types?

Response: We sincerely thank the reviewer for this valuable comment and for pointing out the potential ambiguity in our wording. Our intention was not to classify all non-wheat pixels into four categories in a single step, but rather to conduct two separate binary classifications according to the respective growth stages. Specifically, one binary classification distinguishes between non-wheat winter crops and non-winter crops, and the other distinguishes between non-wheat spring crops and non-spring crops. These two classifications were carried out independently to match the phenological differences between winter and spring crops.

P12, L277-280: Specifically, for the winter growing season (October–June), the remaining non-wheat pixels were classified into winter crops (non-wheat) and non-winter crops using a binary classifier. Similarly, for the spring growing season (April–August), another binary classifier was applied to the remaining non-wheat pixels to separate spring crops (non-wheat) from non-spring crops.

Point 4. Line 238: What is the definition of the "spectral separability indices (SI)"?

Response: We thank the reviewer for this comment. The separability index (SI) is used to assess the sensitivity of the spectral separability of two classes under certain conditions, determined by the ratio of inter-class and intra-class variability. The higher the value, the better the separation between the two classes in the specified condition.

P12, L284-286: The SI is used to assess the sensitivity of the spectral separability of two classes under certain conditions, determined by the ratio of inter-class and intra-class variability (Somers and Asner, 2013). The higher the value, the better the separation between the two classes in the specified condition.

Somers, B. and Asner, G. P.: Multi-temporal hyperspectral mixture analysis and feature selection for invasive species mapping in rainforests, Remote Sens. Environ., 136, 14-27, https://doi.org/10.1016/j.rse.2013.04.006, 2013.

Point 5. Section 3.2: what would the accuracy be if you do not do the Selection of provincial feature set?

Response: We sincerely thank the reviewer for this constructive question. To address it, we conducted a detailed evaluation using Henan Province from 2018 to 2023 as an example. The results in Table R1-1 indicate that, without the selection of provincial feature sets, the overall accuracy ranged from 0.960 to 0.980. After applying feature selection, the accuracy improved to 0.974–0.988, with annual improvements ranging from 0.003 to 0.018. This demonstrates that, although the model already achieved relatively high accuracy without feature selection, the use of provincial feature selection further enhanced its discriminative capacity. To better illustrate this improvement, six representative locations were selected for comparison in Fig.R1-4, where it is evident that feature selection enabled more precise spatial identification of winter wheat, thereby increasing the reliability and robustness of the results.

P26-27, L498-504: Taking Henan Province as an example, we calculated the variation of wheat

precision before and after feature selection (Table R1-1). Without provincial feature selection, the overall precision ranged from 0.960 to 0.980; after feature selection, the overall precision increased to 0.974 to 0.988. Although the model already achieved relatively high accuracy without feature selection, the use of provincial feature selection further enhanced its discriminative capacity. To better illustrate this improvement, six representative locations were selected for comparison in Fig. R1-4, where it is evident that feature selection enabled more precise spatial identification of winter wheat, thereby increasing the reliability and robustness of the results.

Table R1-1: Comparison of wheat recognition accuracy in Henan Province before and after feature selection.

	2018	2019	2020	2021	2022	2023
No feature selection	0.976	0.974	0.979	0.980	0.964	0.960
Feature selection	0.987	0.979	0.982	0.988	0.982	0.974

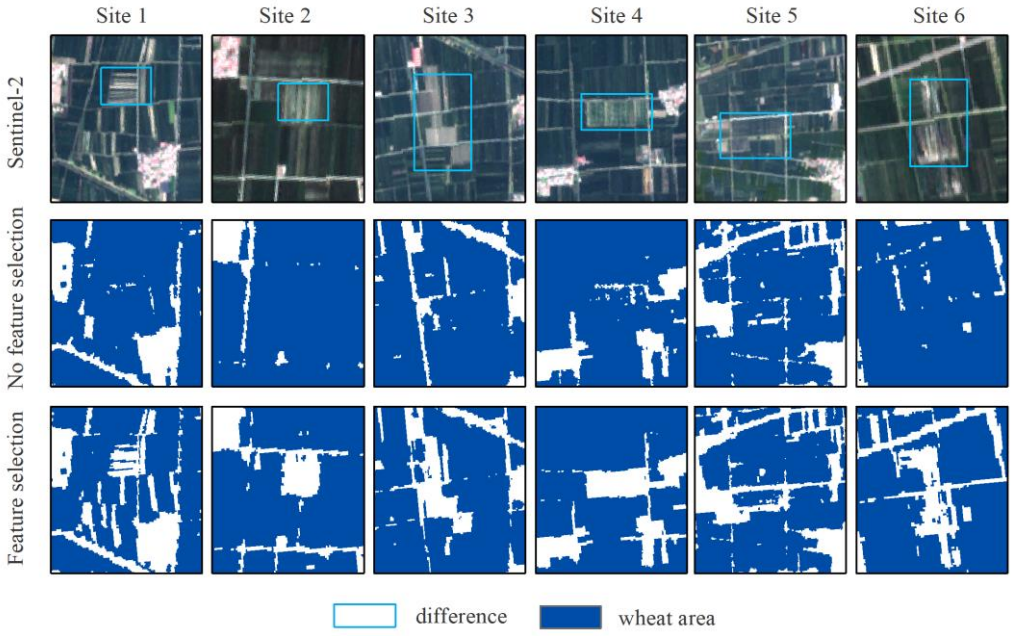

Figure R1-4: Comparison of wheat remote sensing recognition regions before and after feature selection.

Point 6. Line 256: It is not clear why (Yang et al. 2023) is cited here.

Response: We sincerely thank the reviewer for pointing out this ambiguity. Our original intention in citing Yang et al. (2023) was to reference a study that adopted a similar parameter setting for the classifier, specifically using 100 decision trees in the Random Forest algorithm while keeping other parameters at their default values. However, we acknowledge that the current wording may not clearly convey this purpose, which could lead to confusion. In order to improve the clarity, we modified the corresponding content and verified the influence of RF parameters on accuracy:

P14, L301-304: The classifier was implemented with 100 decision trees, there was no more significant difference in accuracy starting with 100 trees and continuing until 200 trees, as shown in Fig. R1-5. The remaining parameters were maintained at their default values, following the approach adopted in recent remote sensing studies (Yang et al., 2023).

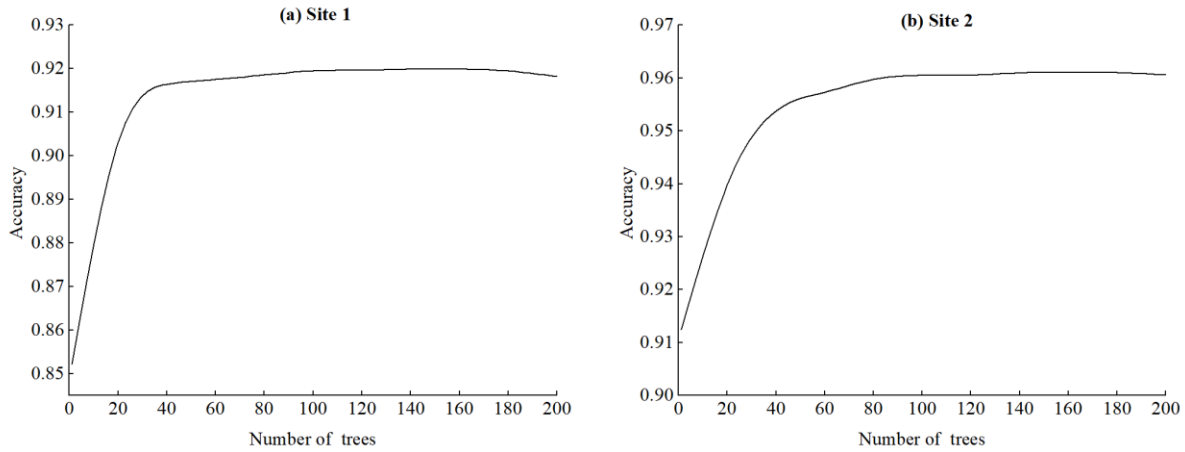

Figure R1-5. Trends in accuracy of RF classifiers under different parameters at two study sites.

Point 7. Line 264-269: it is not clear to me why does the model can map planted winter wheat and harvested winter wheat by changing the time window of the feature set.

Response: We sincerely thank the reviewer for this valuable comment. The approach of using different temporal windows to map planted winter wheat and harvested winter wheat was designed with reference to previous study (Hu et al., 2024), which has demonstrated that crops at different phenological stages exhibit distinct spectral characteristics in remote sensing imagery. By selecting appropriate temporal windows, it is possible to effectively distinguish between planting and post-harvest conditions.

Specifically, winter wheat has a relatively long growth cycle of approximately 8–9 months from sowing to harvest. From early October to early April of the following year, it mainly undergoes sowing, overwintering, regreening, and jointing stages. April represents the peak growth period, during which the canopy is well developed, vegetation coverage is high, and spectral signatures are stable and distinctive, making it the optimal period for identifying the actual planting distribution. From mid- to late April, winter wheat enters the heading stage, followed by grain filling and maturity. During these stages, spectral features can be strongly influenced by adverse meteorological or biological factors, such as drought stress, high-temperature hot dry wind events, pest and disease outbreaks, and flooding. These stressors may cause premature senescence, yield reduction, or even total crop loss. Such abnormal growth events not only directly reduce the final harvested area but also lead to pronounced changes in spectral characteristics in remote sensing imagery, with a shift from the high reflectance typical of healthy vegetation to spectral patterns resembling bare soil or other non-crop surfaces, which in turn results in misclassification or exclusion during harvest-stage mapping. By late June, after harvest, most fields appear as bare soil or stubble-covered surfaces, with spectral properties markedly different from those during the planting period.

As shown in Fig.R1-6, the spectral characteristics of winter wheat at different phenological stages across three representative locations are presented, along with the corresponding differences between planted and harvested areas. This content is also described in Section 4.3 of the manuscript.

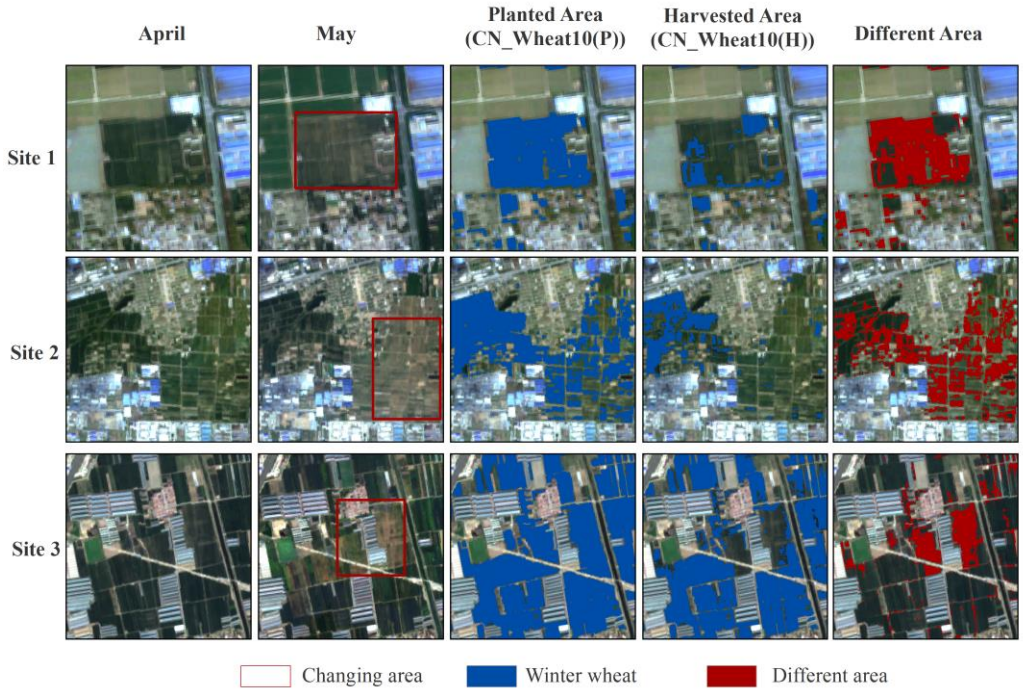

Figure R1-6: Spectral characteristics of winter wheat at different growth stages and differences between planted and harvested area.

Therefore, by changing the time window of the feature set, the model can extract features from different phenological phases: (1) Early October – Early April: captures the actual planted winter wheat distribution. (2) Early October – Late June: encompasses the harvest and post-harvest stages while also reflecting the impacts of adverse meteorological or biological factors on yield and final harvested area, enabling accurate mapping of harvested winter wheat.

The cited reference is as follows:

Hu, J., Zhang, B., Peng, D., Huang, J., Zhang, W., Zhao, B., Li, Y., Cheng, E., Lou, Z., and Liu, S.: Mapping 10-m harvested area in the major winter wheat-producing regions of China from 2018 to 2022, Sci. Data, 11, 1038, https://doi.org/10.1038/s41597-024-03867-z, 2024.

Point 8. Line 269: "The final products include harvested area maps of spring and winter wheat for 15 provinces, as well as planted area maps of winter wheat for 10 provinces." What cause the difference number of available provinces for harvested area maps and planted area maps of winter wheat?

Response: We sincerely thank the reviewer for this question and the opportunity to clarify. The difference in the number of provinces covered by the winter wheat planted area maps (10 provinces) and the harvested area maps (15 provinces) is mainly due to the following factors:

(1) Distribution of major winter wheat production areas: The 10 provinces for which winter wheat planted area maps were generated are located in the Huang–Huai–Hai Plain and the middle–lower reaches of the Yangtze River Plain, representing the core winter wheat-producing regions in China. These provinces have large and contiguous cultivation areas, making them suitable for accurate mapping of planted areas using remote sensing.

(2) Lower wheat cultivation intensity in other provinces: The remaining 5 provinces are mainly in northwestern China, where wheat planting is relatively limited, and fields are often small and scattered. This reduces the feasibility and reliability of extracting winter wheat planted areas at the provincial scale.

(3) Mixed cropping of spring and winter wheat: In some northwestern provinces, spring and winter wheat are cultivated in the same province. In certain areas, the harvest periods of spring and winter wheat are very close, making it difficult to reliably distinguish planted and harvested areas for wheat using remote sensing within a single growing season.

(4) Phenological differences across regions: Wheat growth cycles vary considerably across climatic zones. For example, in eastern China, winter wheat reaches its peak growth stage in April, whereas in northwestern regions, due to colder climates, winter wheat often does not reach its peak growth until May. This temporal difference, combined with overlapping growth stages and spectral characteristics of different wheat types in some regions, increases the difficulty of extracting wheat planted areas at a national scale. However, during the harvest stage, wheat exhibits distinct spectral characteristics compared to other land cover types, and the harvest period is relatively consistent across regions, allowing harvested area maps to be generated for all 15 provinces.