

## Article

# A Landscape-Clustering Zoning Strategy to Map Multi-Crops in Fragmented Cropland Regions Using Sentinel-2 and Sentinel-1 Imagery with Feature Selection

Guanru Fang <sup>1</sup>, Chen Wang <sup>1</sup>, Taifeng Dong <sup>2</sup>, Ziming Wang <sup>1</sup>, Cheng Cai <sup>1</sup>, Jiaqi Chen <sup>1</sup>, Mengyu Liu <sup>1</sup> and Huanxue Zhang <sup>1,\*</sup>

<sup>1</sup> College of Geography and Environment, Shandong Normal University, Jinan 250300, China; 2021020800@stu.sdn.edu.cn (G.F.)

<sup>2</sup> Ottawa Research and Development Centre, Agriculture and Agri-Food Canada, Ottawa, ON K1A 0C6, Canada

\* Correspondence: zhanghuanxue@sdnu.edu.cn

**Abstract:** Crop mapping using remote sensing is a reliable and efficient approach to obtaining timely and accurate crop information. Previous studies predominantly focused on large-scale regions characterized by simple cropping structures. However, in complex agricultural regions, such as China's Huang-Huai-Hai region, the high crop diversity and fragmented cropland in localized areas present significant challenges for accurate crop mapping. To address these challenges, this study introduces a landscape-clustering zoning strategy utilizing multi-temporal Sentinel-1 and Sentinel-2 imagery. First, crop heterogeneity zones (CHZs) are delineated using landscape metrics that capture crop diversity and cropland fragmentation. Subsequently, four types of features (spectral, phenological, textural and radar features) are combined in various configurations to create different classification schemes. These schemes are then optimized for each CHZ using a random forest classifier. The results demonstrate that the landscape-clustering zoning strategy achieves an overall accuracy of 93.52% and a kappa coefficient of 92.67%, outperforming the no-zoning method by 2.9% and 3.82%, respectively. Furthermore, the crop mapping results from this strategy closely align with agricultural statistics at the county level, with an  $R^2$  value of 0.9006. In comparison with other traditional zoning strategies, such as topographic zoning and administrative unit zoning, the proposed strategy proves to be superior. These findings suggest that the landscape-clustering zoning strategy offers a robust reference method for crop mapping in complex agricultural landscapes.



Academic Editor: Simone Priori

Received: 16 December 2024

Revised: 2 January 2025

Accepted: 14 January 2025

Published: 16 January 2025

**Citation:** Fang, G.; Wang, C.; Dong, T.; Wang, Z.; Cai, C.; Chen, J.; Liu, M.; Zhang, H. A Landscape-Clustering Zoning Strategy to Map Multi-Crops in Fragmented Cropland Regions Using Sentinel-2 and Sentinel-1 Imagery with Feature Selection. *Agriculture* **2025**, *15*, 186. <https://doi.org/10.3390/agriculture15020186>

**Copyright:** © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** crop mapping; landscape heterogeneity; feature selection; crop heterogeneity zone; complex agricultural regions

## 1. Introduction

China is one of the largest agricultural producers in the world, with vast areas under cultivation and a wide variety of food and cash crops, resulting in a highly diverse planting structure. Acquiring timely and accurate information on crop cultivation is therefore essential for ensuring food security, monitoring crop production, and supporting policy-making efforts [1,2]. Traditional methods such as sample statistics and manual surveys are labor-intensive and resource-demanding. By contrast, remote sensing offers the ability to perform frequent monitoring over short periods and to capture crop dynamics on a large scale [3,4]. With advancements in classification algorithms and the increasing availability

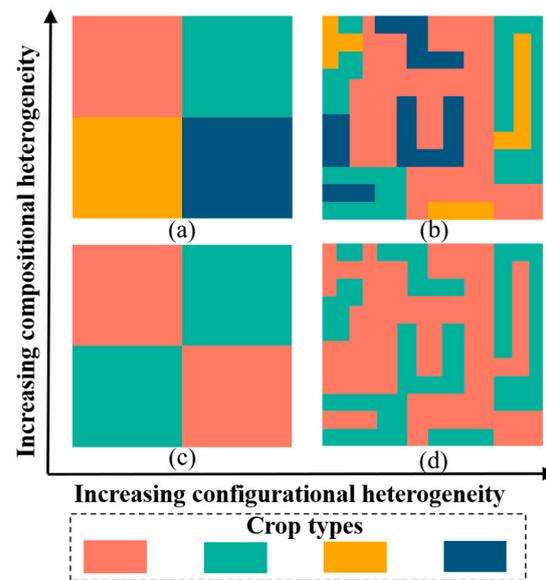
of open-source satellite data (e.g., Landsat and Sentinel), remote sensing has emerged as a widely adopted and effective method for monitoring crop conditions and estimating yields [5,6].

Currently, although numerous relatively mature remote sensing-based crop products have been developed in several developed countries and provincial regions in China—such as the Cropland Data Layer (CDL) in the USA and cotton mapping in Xinjiang, China [7,8]—these efforts primarily target areas with large-scale automated cultivation and relatively simple cropping structures. By contrast, regions characterized by complex agricultural landscapes, such as the central region of the Huang-Huai-Hai regions of China, present challenges due to their highly diversified crops and fragmented cropland patches. The crop conditions and natural environments in these regions are often insufficiently represented in large-scale studies, making their findings less universally applicable. Moreover, research has shown that crop diversity and cropland fragmentation significantly influence the accuracy of crop mapping [9–11]. For example, Chen et al. achieved a higher overall accuracy (OA) in areas with lower crop diversity and evenness in crop classification using phenology metrics derived from MODIS data [12]. Han et al. reported an 11% disparity in OA for rapeseed identification between fragmented and non-fragmented areas using features derived from Landsat 8 and Sentinel-1 satellites [13]. These findings highlight the need to refine crop mapping models to improve their applicability in regions with diverse crops and fragmented croplands.

To address the challenges posed by spatial heterogeneity, zoning strategies that divide entire regions into homogeneous subzones have been widely explored and assessed [14–16]. Some studies have divided the study area into a set of ortho-hexagonal units or administrative units [17,18]; however, these studies do not take into account the spatial heterogeneity between units and need large sample sizes. Some studies have employed agro-ecological or climatic factors to delineate subzones in large-scale regions [19,20]. However, in localized complex agricultural regions, these strategies are often difficult to refine due to the similarity of cropping environments across the area. In addition, some studies have utilized environmental variables, such as digital elevation models (DEMs) and temperature, for partitioning [21,22]. For instance, Ren et al. used DEM data to partition a study area with significant topographical variation into subzones with relatively homogeneous spatial distributions of crops. Sentinel-1/2 imagery was then applied within each subzone to map crops, achieving an overall accuracy (OA) of approximately 90% [23]. Similarly, Donmez et al. employed ERA-5 temperature data to delineate multi-crop areas, achieving an OA of about 92% by accounting for the strong correlation between crop phenology and temperature in the study area [24]. These strategies integrated key environmental influences to delineate subzones tailored to the study area's characteristics. However, they face limitations in complex agricultural regions with high crop diversity and fragmented cropland because such areas are influenced by multiple environmental factors.

In landscape ecology, crop diversity and cropland fragmentation in agricultural areas correspond to the compositional and configurational heterogeneity of crop heterogeneity [25] (as shown in Figure 1). Crop heterogeneity can be effectively quantified using landscape metrics, which serve as analytical tools for describing the spatial characteristics of complex agricultural landscapes [26,27]. Significant correlations have also been identified between agricultural landscape characteristics, as quantified by landscape metrics, and the accuracy of crop classification. For instance, Zhang et al. quantified fragmented cropland using the Splitting Index (SPLIT) and demonstrated that incorporating spectral and textural features derived from Sentinel-2 data effectively mitigated reductions in crop mapping accuracy caused by cropland fragmentation [11]. Mondal and Jeganathan extracted agricultural land using MODIS imagery and a spectral similarity index calculated by Euclidean

distance. Their study revealed that increases in the Number of Cropland Patches (NP) and decreases in the Mean Patch Size (MPS) of cropland both contributed to higher omission errors in crop classification. These findings underscore the utility of landscape metrics in providing critical reference information for crop mapping [9]. However, due to limitations in data sources, the above studies calculated landscape metrics based on pre-acquired mapping results, focusing on exploring the effects of crop heterogeneity on crop mapping accuracy. Few studies directly integrate ancillary data to quantify crop heterogeneity and improve crop mapping approaches. This gap highlights the need for further research to expand the application of landscape metrics in complex agricultural regions.



**Figure 1.** Illustration of crop heterogeneity. (a–d) represent landscapes, and different colors indicate different crop types within the landscape. (a) to (b) and (c) to (d) indicate increases in the fragmentation of cropland (configurational heterogeneity). (c) to (a) and (d) to (b) indicate increases in crop diversity (compositional heterogeneity).

Given the high crop diversity and cropland fragmentation in local areas of the Huang-Huai-Hai regions of China, this study selected several representative cities from the intertwined mountainous plains of central Shandong Province. A landscape-clustering zoning strategy considering crop heterogeneity based on Sentinel-1 and Sentinel-2 satellite data was proposed to address these complexities.

The study focused on three specific processes:

Quantitatively describe complex agricultural regions using landscape metrics based on crop heterogeneity.

Divide complex planting regions into Crop Heterogeneity Zones (CHZs) and determine optimal classification schemes.

Evaluate the precision of the proposed landscape-clustering zoning strategy.

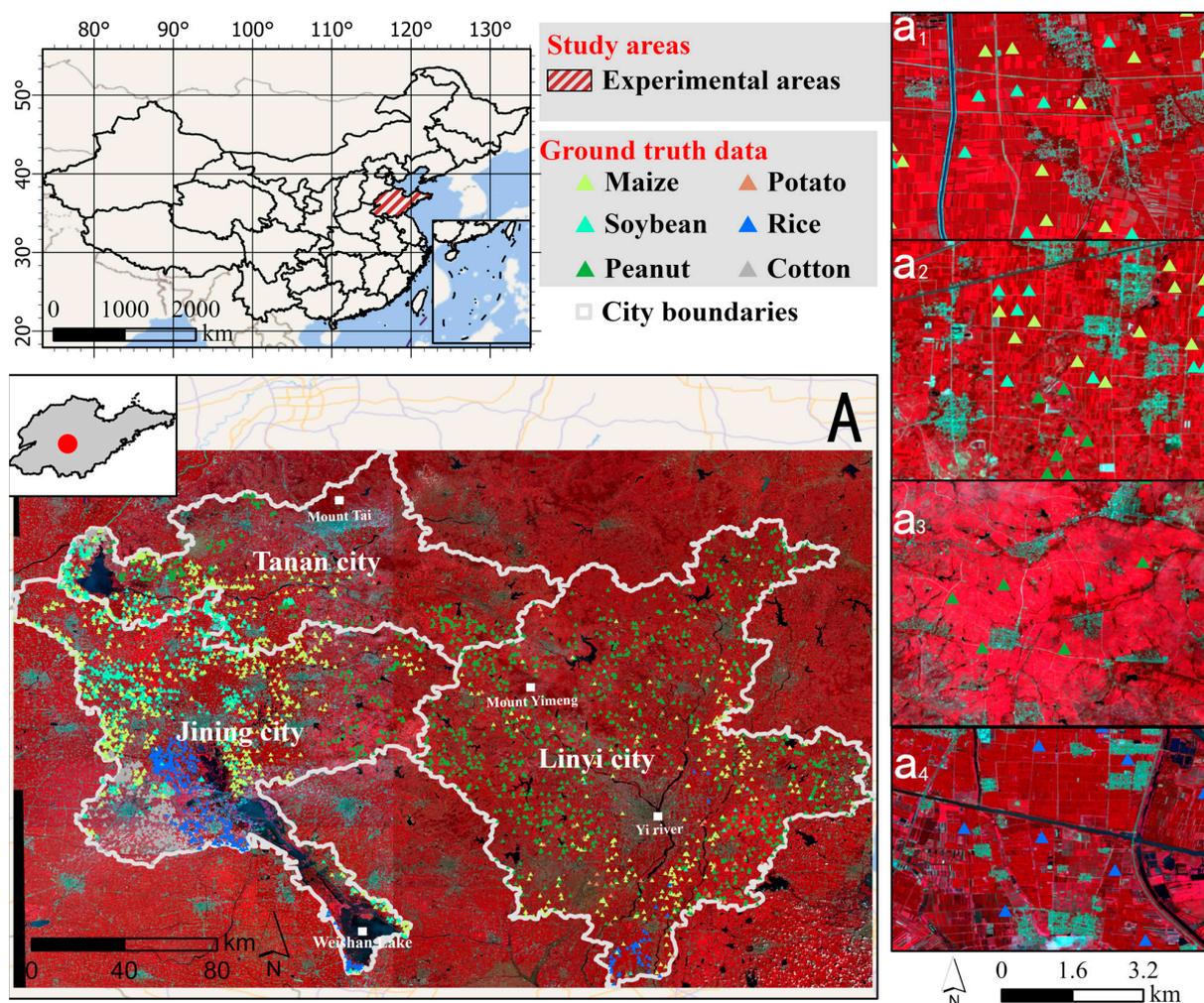
## 2. Materials

### 2.1. Study Areas

Shandong Province is located in the eastern part of the Huang-Huai-Hai Plain ( $34^{\circ}22'–38^{\circ}23' N$ ,  $114^{\circ}47'–122^{\circ}42' E$ ) and features a warm temperate monsoon climate. Summers are rainy, while winters are relatively dry. The province's topography consists of central and eastern mountainous areas, plains in the northwest and southwest, and coastal hills, creating significant topographic variation. Shandong is also characterized by complex

water systems, including the Yellow River as the primary system and the Huaihe and Haihe River systems located in the southern and northern parts of the province.

In this study, Jining City, Tai'an City, and Linyi City, located in central Shandong, were selected as the study areas (Figure 2A). The climate in these areas exhibits a similar distribution of rainfall and heat throughout the year, with a mean annual temperature of 16 °C and a mean annual precipitation of 701 mm. The terrain in this region includes plains in the west and east, as well as mountainous and hilly areas in the center, with elevations ranging from 32 m to 1545 m. The plains in the western and eastern parts are rich in water resources, including Weishan Lake and the Yi River, while the central region is dominated by Mount Tai and Mount Yimeng, surrounded by hills. Due to the influence of various natural environments, the study area is characterized by fragmentation of arable land and crop diversification. The study area encompasses 14,227.9 km<sup>2</sup> of arable land, primarily cultivating corn, soybean, rice, cotton, peanuts, and potatoes. Maize and soybean are widely distributed in the plains, while rice and cotton, which require abundant water, are concentrated in areas with extensive water systems. Peanuts and potatoes are grown in hilly areas because of their drought tolerance. Most crops have a growing period from mid-May to mid-October. Potatoes have a shorter growing season, lasting only from mid-August to mid-October (Table 1).



**Figure 2.** The geo-location of the study area with survey samples of different crops. The four panels (a1–a4) present representative examples, illustrating details of crop distribution. All images were acquired using Sentinel-2 satellites in August 2019 and are displayed as false color composites (Bands: B8, B4, and B3).

**Table 1.** Crop phenological stages in the study areas.

| Crop Types | Apr. |   |   | May |   |   | Jun. |   |   | Jul. |   |   | Aug. |   |   | Sept. |   |   | Oct. |   |   |  |
|------------|------|---|---|-----|---|---|------|---|---|------|---|---|------|---|---|-------|---|---|------|---|---|--|
|            | E    | M | L | E   | M | L | E    | M | L | E    | M | L | E    | M | L | E     | M | L | E    | M | L |  |
| Maize      |      |   |   |     |   |   |      |   |   |      |   |   |      |   |   |       |   |   |      |   |   |  |
| Soybean    |      |   |   |     |   |   |      |   |   |      |   |   |      |   |   |       |   |   |      |   |   |  |
| Peanut     |      |   |   |     |   |   |      |   |   |      |   |   |      |   |   |       |   |   |      |   |   |  |
| Cotton     |      |   |   |     |   |   |      |   |   |      |   |   |      |   |   |       |   |   |      |   |   |  |
| Rice       |      |   |   |     |   |   |      |   |   |      |   |   |      |   |   |       |   |   |      |   |   |  |
| Potato     |      |   |   |     |   |   |      |   |   |      |   |   |      |   |   |       |   |   |      |   |   |  |

Note: Each month is split into three phases: early (E), middle (M), and late (L). Pink regions denote the sowing and seeding/flooding stages, green regions signify the reviving and tillering/flowering stages, and brown regions represent the maturity and harvest stages.

### 2.2. Image Data

The Sentinel-2 satellite system, comprising A and B sensors, offers a revisit cycle of 5 days. Its imagery consists of 13 spectral bands, covering wavelengths from visible to shortwave infrared, with spatial resolutions ranging from 10 to 60 m. In this study, Sentinel-2 Level-2A surface reflectance (SR) images were acquired and pre-processed using Google Earth Engine (GEE), and ten bands were used with a resolution of 10 m and 20 m for crop mapping. Considering the growing seasons of the main crops in the study area, images from 10 May to 10 October 2019 were selected. To minimize the cloud contamination effects, only images with less than 25% cloud cover were retained, and cloud removal was performed using the QA60 band, which contains cloud mask information. To address gaps caused by clouds and maintain consistent temporal resolution [28], monthly composite images were generated by averaging available image values for each month.

In this study, Sentinel-1 Ground Range Detected (GRD) images were used to supplement crop cultivation information. Sentinel-1 images in Interferometric Wide (IW) mode were selected due to their ability to maintain a revisit performance of 6 days and provide extensive land surface coverage with spatial resolutions ranging from 5 to 40 m. This mode includes two polar-orbiting satellites equipped with C-band Synthetic Aperture Radar (SAR) instruments operating in dual polarization (VV and VH). All Sentinel-1 images obtained from the GEE had undergone pre-processing steps such as thermal noise removal, radiometric calibration, and terrain correction. To further reduce image speckle noise, a refined Lee filter with a 7 × 7 window was applied in this study. Finally, the images were composited into monthly averages to ensure temporal consistency with the Sentinel-2 imagery.

### 2.3. Classification Features

To reduce the impact of spectral similarity on crop classification caused by the diverse crop types in the study areas, multiple feature types derived from multi-temporal Sentinel-2 and Sentinel-1 images were utilized (Table 2). Spectral features, which include spectral bands and indices, provide information on crop reflectance at short times during the growing season. These features are closely related to crop growth traits such as vegetation canopy and leaf chlorophyll [29]. Phenological features capture the temporal dynamics of crops over the entire growing period, such as the timing of growth onset and cessation. To represent these dynamics, a double logistic regression function [30] was applied to fit the Enhanced Vegetation Index (EVI) time series imagery, effectively modeling changes in crop growth patterns (Figure 3). Radar features [31], including backscattering coefficients and their combinations, are sensitive to structural attributes like vegetation height and density. These features supplement the spectral data by providing additional insight into crop morphology. Texture features describe the spatial distribution of crops by analyzing the gray-scale relationships between image pixels. In this study, texture features based

on the Gray-Level Co-occurrence Matrix (GLCM) were employed to enhance the spatial representation of crop distributions [32].

**Table 2.** Overview of the feature types utilized in the study.

| Feature Types                           | Features  | Description   |
|---|---|---|
| Spectral features<br>(1) Spectral bands | Sentinel-2 bands  | $B2, B3, B4, B5, B6, B7, B8, B8A, B11,$ and $B12$   |
| Spectral features                       | Enhanced Vegetation Index (EVI)   | $2.5 \times (B8 - B4) / (B8 + 6 \times B4 - 7.5 \times B2 + 1)$   |
| (2) Spectral indices                    | MERIS Terrestrial Chlorophyll Index (MTCI)  | $(B6 - B5) / (B5 - B4)$   |
|   | Chlorophyll Absorption in Reflectance Index (CARI)  | $(B5/B2) - 1$   |
|   | Land Surface Water Index (LSWI)   | $(B8 - B11) / (B8 + B11)$   |
|   | Wide Dynamic Range Vegetation Index (WDRVI)   | $(0.2 \times B8 - B4) / (0.2 \times B8 + B4)$   |
|   | Green Normalized Difference Vegetation Index (GNDVI)  | $(B8 - B3) / (B8 + B3)$   |
|   | Sentinel-2 Red Edge Position Index (S2REP)  | $705 + 35 \times (((B4 + B7)/2) - B5) / (B6 - B5)$  |
|   | Renormalized Difference Vegetation Index (RDVI)   | $(B8 - B4) / ((B8 + B4)^{0.5})$   |
| Phenological features                   | SOS   | The date in the green-up phase when the rate of increase in the first derivative of EVI reaches its peak. |
|   | Integral  | The cumulative EVI across the crop season.  |
|   | EOS   | The senescence phase date when the first derivative of EVI shows the steepest decline.                    |
|   | LOS   | The time span between the season's end (EOS) and start (SOS).   |
|   | BL  | The lowest EVI observed during the crop season.   |
|   | MOS   | The date of the maximum EVI value during the crop growing season.   |
|   | value_SOS   | The EVI value at the start of the season (SOS).   |
|   | value_EOS   | The EVI value at the end of the season (EOS).   |
|   | value_MOS   | The EVI value at the peak of the season (MOS).  |
|   | SA  | EVI values throughout the crop season.  |
| Radar features                          | Backscattering coefficient and their combinations   | $VV, VH, VV - VH, VV + VH, (VH - VV) / (VV + VH), VV/VH$  |
| Textural features                       | Contrast (CON), Variance (VAR), Homogeneity (IDM), Correlation (CORR), Entropy (ENT), and Angular Second Moment (ASM) | Calculated from the Red Edge (B5), NIR (B8), and SWIR (B11) based on GLCM with a $3 \times 3$ window.     |

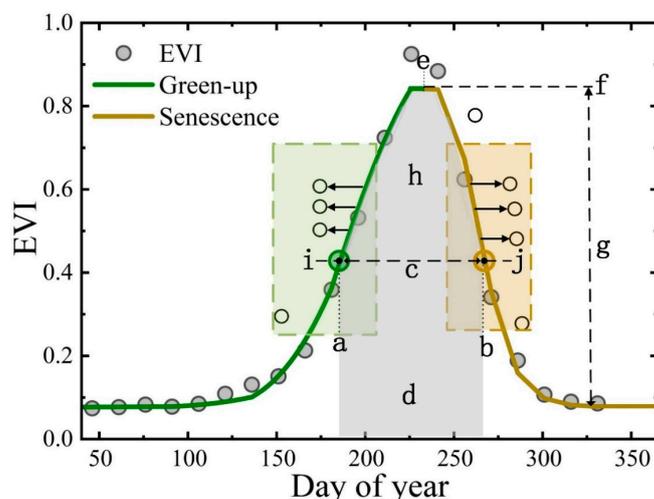
Note: Spectral features, radar features, and textural features are extracted using images from May to October.

The initial features from different growth stages might have strong intercorrelations, making them ineffective for crop identification. Here, a data reduction approach based on the JM distance (Equations (1) and (2)) was used to compress the initial features and select the optimal features [33].

$$JM = 2 \left( 1 - e^{-B_{ij}} \right) \quad (1)$$

$$B_{ij} = \frac{1}{8} (m_i - m_j) \frac{2}{\sigma_i^2 + \sigma_j^2} + \frac{1}{2} \ln \left( \frac{\sigma_i^2 + \sigma_j^2}{2\sigma_i\sigma_j} \right) \quad (2)$$

where  $B$  denotes the Bhattacharyya distance;  $JM$  represents the JM distance between classes  $i$  and  $j$ ;  $m_i$  and  $m_j$  are the means of the feature vectors for the samples; and  $\sigma_i$  and  $\sigma_j$  are the standard deviations of the feature vectors for the samples.



**Figure 3.** Phenological features fitted by double logistic regression function. (a) SOS, (b) EOS, (c) LOS, (d) BL, (e) MOS, (f) value\_MOS, (g) SA, (h) Integral, (i) Value\_SOS, and (j) Value\_EOS, with detailed descriptions in Table 2.

In the study, the JM distance was calculated using 30 crop samples between any two crop types (i.e., maize, soybean, peanut, cotton, and rice). These 30 crop samples were collected from regions with varying levels of crop diversification and cropland fragmentation across the entire study area. This approach ensured that the selected features were applicable to all CHZs. Features with a JM distance  $> 1$  were selected as candidate features for crop classification [34].

#### 2.4. Auxiliary Data

##### 2.4.1. Ground Truth Data

Ground truth data were collected during the 2019 cropping season in the study area. The survey recorded detailed information on representative crop types, including maize, soybean, peanut, cotton, rice, and other crops. Geo-locations for part of the samples were recorded using a handheld GPS with a positional accuracy of 5 m. Additionally, other sample locations were identified through visual interpretation of multi-temporal Sentinel-2 images (Figure 2). The dataset comprised 990 maize samples (232 from the handheld GPS), 627 soybean samples (185 from the handheld GPS), 748 peanut samples (262 from the handheld GPS), 363 cotton samples (102 from the handheld GPS), 251 rice samples (156 from the handheld GPS), and 321 potato samples (172 from the handheld GPS). To ensure robust analysis, all collected samples were randomly split into 70% training data and 30% validation data.

##### 2.4.2. Agricultural Statistical Data

County-level crop statistics for 2019 were obtained from local municipal statistical yearbooks (Figure 4). These statistics were used to calculate county-level compositional heterogeneity metrics (i.e., crop diversity). Additionally, the statistical crop areas were compared with the monitored areas derived from the crop mapping results to assess the performance of the landscape-clustering zoning strategy.

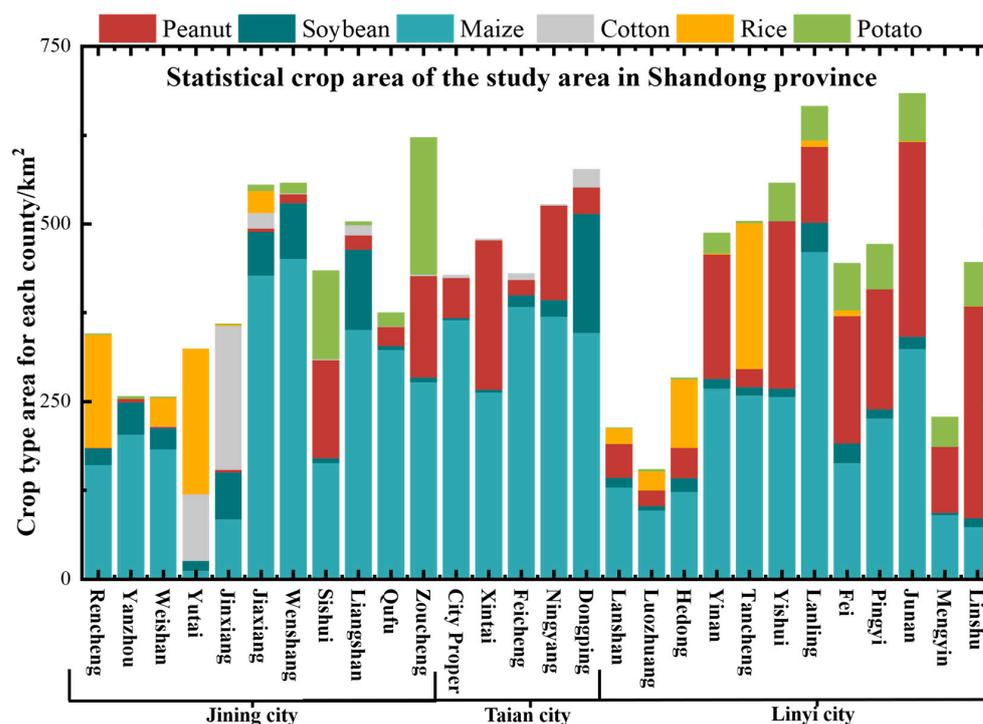


Figure 4. Statistical crop area in the study area for each county.

#### 2.4.3. Land Use Data

The ESA WorldCover 2020 product at 10 m spatial resolution has 11 land use categories [35]. The distribution of cropland in the product (coded as 40) was used to remove non-crop areas.

### 3. Methods

The workflow of this study is illustrated in Figure 5. First, CHZs, representing relatively homogeneous areas, were clustered using comprehensive landscape metrics. Second, diverse feature combinations were paired with a Random Forest (RF) classifier within each CHZ, and the optimal feature combination for each CHZ was selected. Finally, the proposed landscape-clustering zoning strategy was employed to map multiple crop types in fragmented areas, and its performance was evaluated.

#### 3.1. The Landscape-Clustering Zoning Strategy

This section mainly describes the landscape-clustering process. Comprehensive landscape metrics and CHZs were obtained in Sections 3.1.1 and 3.1.2, respectively. The optimal feature combination in each CHZ was selected in Section 3.1.3.

##### 3.1.1. Obtaining Comprehensive Landscape Metrics

###### Cropland Extraction

To accurately capture features of cropland fragmentation, a hierarchical extraction method was employed to identify cropland patches. The process utilized Sentinel-2 August mean composite images (Figure 6) and a series of filtering rules. First, water bodies were removed using the NDWI (Normalized Difference Water Index) with an appropriate threshold (Equation (3)). Next, artificial objects such as bare ground, buildings, and road networks were excluded based on the NDBI (Normalized Difference Built-up Index) (Equation (4)). Finally, trees were filtered out using the B8 (NIR) band to produce preliminary results. To ensure the exclusion of non-cultivated image elements, the preliminary results were refined by clipping them with the cultivated map derived from the ESA WorldCover 2020 product.

This process produced a final cropland mask, which was subsequently used to calculate the configuration heterogeneity metrics and exclude non-cropland objects in the classification.

$$NDWI = (B3 + B8)/(B3 - B8) \tag{3}$$

$$NDBI = (B11 + B8)/(B11 - B8) \tag{4}$$

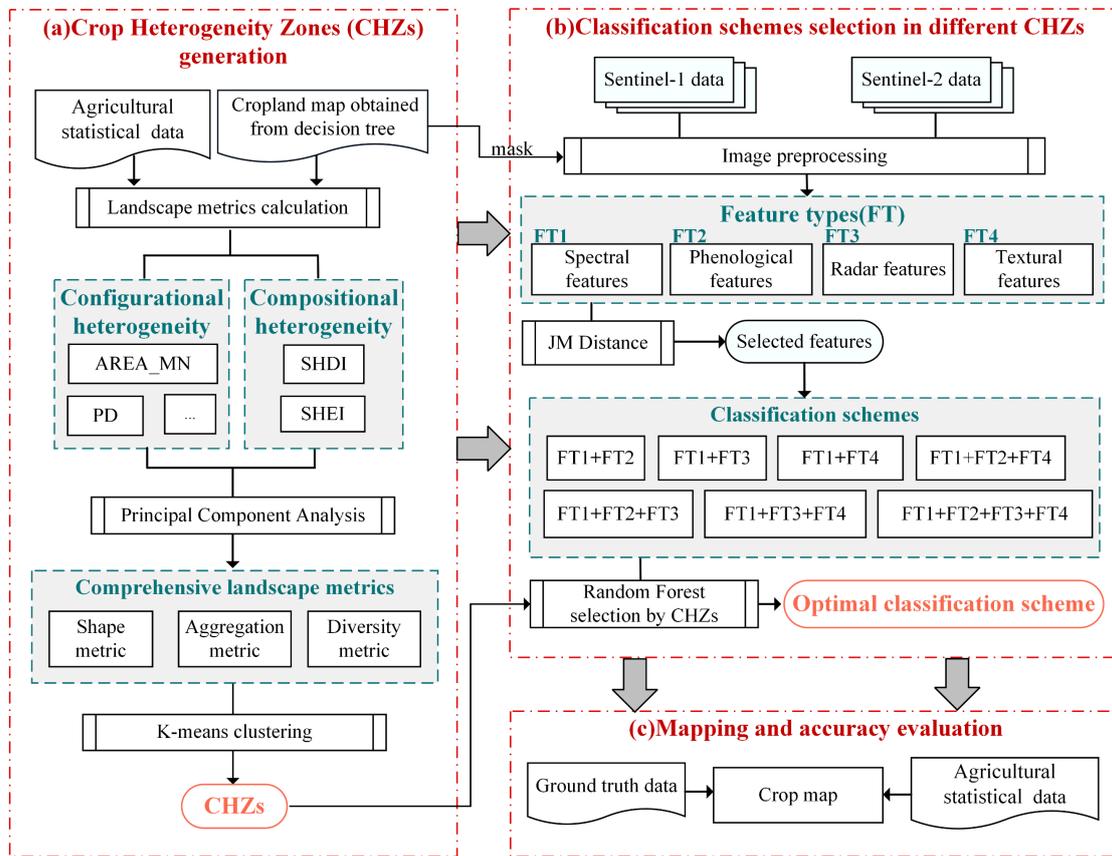


Figure 5. Overall framework of the study.

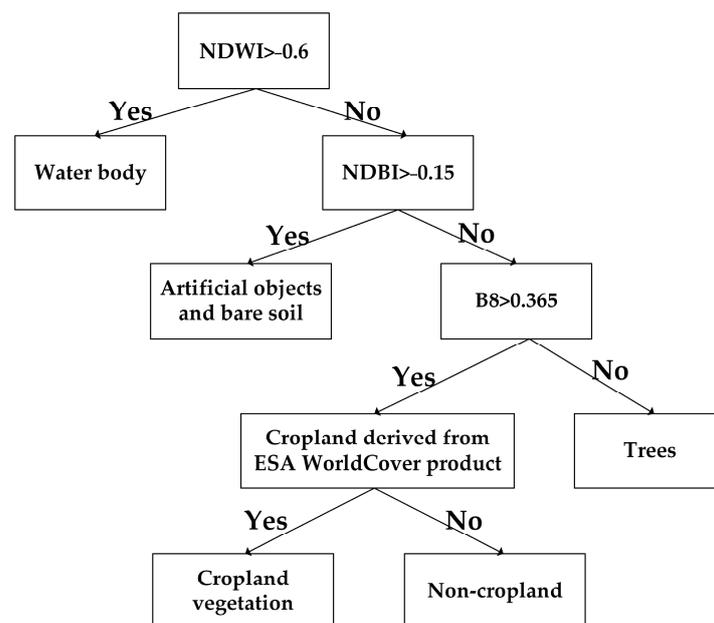


Figure 6. Decision tree for cropland information extraction.

### Crop Heterogeneity Metrics Calculation

Six landscape metrics expressing crop heterogeneity were selected for this study because they were widely used in previous research [9,12]. These metrics were categorized into two groups based on crop heterogeneity: configurational heterogeneity, which represents cropland fragmentation, and compositional heterogeneity, which reflects crop diversity. The landscape metrics were derived from various dimensions, including the aggregation and shape of cultivated patches and the diversity of crops (Table 3). All metrics were calculated at the county scale. Configurational heterogeneity metrics were derived from the cultivated patches described in Section Cropland Extraction, while compositional heterogeneity metrics were calculated using agricultural statistics detailed in Section 2.4.2.

**Table 3.** Selection of crop heterogeneity metrics.

| Type                          | Subtype             | Detail Metrics  | Description                            |
|-------------------------------|---------------------|---|--|
| Configurational heterogeneity | Aggregation metrics | Patch density (PD)<br>Mean patch size (AREA_MN)   | The aggregation of fine cropland       |
|                               | Shape metrics       | Area-weighted mean shape metric (SHAPE_AM)<br>Area-weighted mean fractal dimension metric (FRAC_AM) | The complexity of shape of arable land |
| Compositional heterogeneity   | Diversity metrics   | Shannon’s Diversity index (SHDI)<br>Shannon’s Evenness index (SHEI)                                 | Crop diversity and evenness            |

### Principal Component Analysis

To reduce redundancy between landscape metrics, Principal Component Analysis (PCA) was applied to condense the initial metrics into comprehensive landscape metrics. The process involved three main steps. First, the six initial landscape metrics were standardized using Z-score normalization. Second, the component matrix was rotated using the varimax rotation method to calculate the eigenvalues of the components and the loadings of the landscape metrics. Eigenvalues represent the amount of information contained in the corresponding principal components, with higher values indicating greater significance. Using these eigenvalues, the variance contribution rate was calculated to determine the proportion of information each principal component contributed to the total variance of all components. Loadings revealed the correlation between individual landscape metrics and their associated principal components. Finally, components with eigenvalues greater than one were retained as comprehensive landscape metrics [36]. The attributes of these comprehensive metrics were then interpreted based on the loadings.

Each composite component explained multiple landscape indices and could be expressed as values in the pattern matrix (Equation (5)), as follows:

$$F_i = a_{ie} * x_{ie} + a_{if} * x_{if} + \dots + a_{in} * x_{in} \tag{5}$$

where  $a_{ie}$ ,  $a_{if}$ , and  $a_{in}$  refer to the coefficients corresponding to the chosen landscape metrics  $x_{ie}$ ,  $x_{if}$ , and  $x_{in}$  in factor  $F_i$ , respectively.

#### 3.1.2. Generating CHZs

K-means is an unsupervised clustering technique that assigns data points to the nearest centroids, minimizing the sum of squared deviations within each cluster [37]. In

this study, the K-means algorithm was applied to group comprehensive landscape metrics into relatively homogeneous areas. To evaluate the effectiveness of the clustering, the silhouette coefficient ( $S(k)$ ) was calculated. This metric combines measures of cohesion and separation to assess cluster quality (Equation (6)). The average silhouette coefficient was computed for numbers of clusters,  $k$ , ranging from 2 to 10. The optimal number of CHZs was set as the  $k$  with the highest average silhouette coefficient, as follows:

$$S(k) = \frac{b(k) - a(k)}{\max\{a(i), b(i)\}} \quad (6)$$

where  $a(k)$  represents the mean distance between vector  $k$  and all other points in its cluster, and  $b(k)$  represents the smallest mean distance between vector  $k$  and points in the nearest neighboring cluster.  $S(k)$  ranges from  $-1$  to  $1$ , with values closer to  $1$  indicating higher clustering quality.

### 3.1.3. Designing and Selecting Classification Schemes

The optimal features for crop classification vary across regions due to differences in crop heterogeneity [11,38]. And selecting appropriate feature combinations can mitigate classification confusions [39]. Conversely, using excessive feature types indiscriminately without considering crop heterogeneity can cause overfitting. Therefore, seven classification schemes were developed with different feature combinations (Table 4) and paired with a Random Forest (RF) classifier. The first three schemes (S1–S3) combined spectral features with one additional feature type, while S4, S5, and S6 combined spectral features with two distinct feature types. The S7 scheme incorporated all four feature types.

**Table 4.** Description of feature combinations.

| Classification Scheme | Feature Combinations   |
|-----------------------|--|
| Scheme 1 (S1)         | Spectral features + Phenological features                                      |
| Scheme 2 (S2)         | Spectral features + Radar backscattering features                              |
| Scheme 3 (S3)         | Spectral features + Textural features  |
| Scheme 4 (S4)         | Spectral features + Phenological features + Radar backscattering features      |
| Scheme 5 (S5)         | Spectral features + Phenological features + Textural features                  |
| Scheme 6 (S6)         | Spectral features + Radar features + Textural features                         |
| Scheme 7 (S7)         | Spectral features + Phenological features + Radar features + Textural features |

The RF classifier is well-suited for high-dimensional data and demonstrates reduced susceptibility to overfitting [40]. In this study, two adjustable parameters—number of trees and number of variables per split—were configured. For the number of trees, values were tested incrementally from 50 to 500 in steps of 10, and the parameter was set at the point where the overall accuracy (OA) no longer improved. The number of variables per split was set to the square root of the total number of input variables. Finally, the optimal classification scheme for each CHZ was selected as the one with the highest OA among the seven schemes.

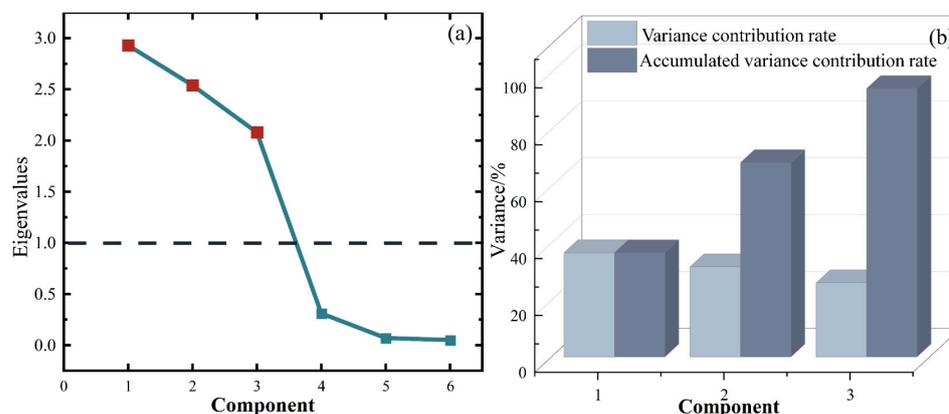
### 3.2. Accuracy Evaluation

We randomly used 70% of the samples for training and the remaining 30% for validation. For the validation, the OA and kappa was used to assess the overall classification performance and consistency, while user's accuracy (UA) and producer's accuracy (PA) were used to evaluate the accuracy of specific crop categories [41]. The F1-score (F1) provides a balanced consideration of both UA and PA, as it is the harmonic mean of UA and PA. Additionally, the mapped crop areas were compared with county-level statistics derived from statistical yearbook data to further validate the classification results.

## 4. Results and Analysis

### 4.1. Acquisition and Analysis of Comprehensive Landscape Metrics

PCA yielded six components from the landscape metrics (Figure 7a). Among these, the first three components, with eigenvalues greater than 1, accounted for 94.36% of the accumulated variance contribution rate (Figure 7b). This result indicates that these three components captured the majority of the information from the original metrics and were retained to analyze in the next step.



**Figure 7.** Variance analysis of PCA with varimax rotation. (a) displays eigenvalues for all components, emphasizing those over 1 with red squares. (b) depicts the variance and cumulative variance of components post-screening.

To determine the characteristics of these three components, the loading coefficients of the landscape metrics were calculated (Figure 8), describing the correlation between each principal component and the landscape metrics. Component 1 showed strong positive correlations with shape metrics (Table 3), specifically SHAPE\_AM and FRAC\_AM (loadings: 0.92 and 0.98, respectively), representing cropland shape complexity. Component 2 was closely associated with aggregation metrics, exhibiting a negative correlation with AREA\_MN (loading:  $-0.96$ ) and a positive correlation with PD (loading: 0.95), reflecting the degree of fragmented cropland patch aggregation. Component 3 correlated positively with diversity metrics, including SHDI and SHEI (loadings: 0.9 and 0.86), indicating crop diversity.

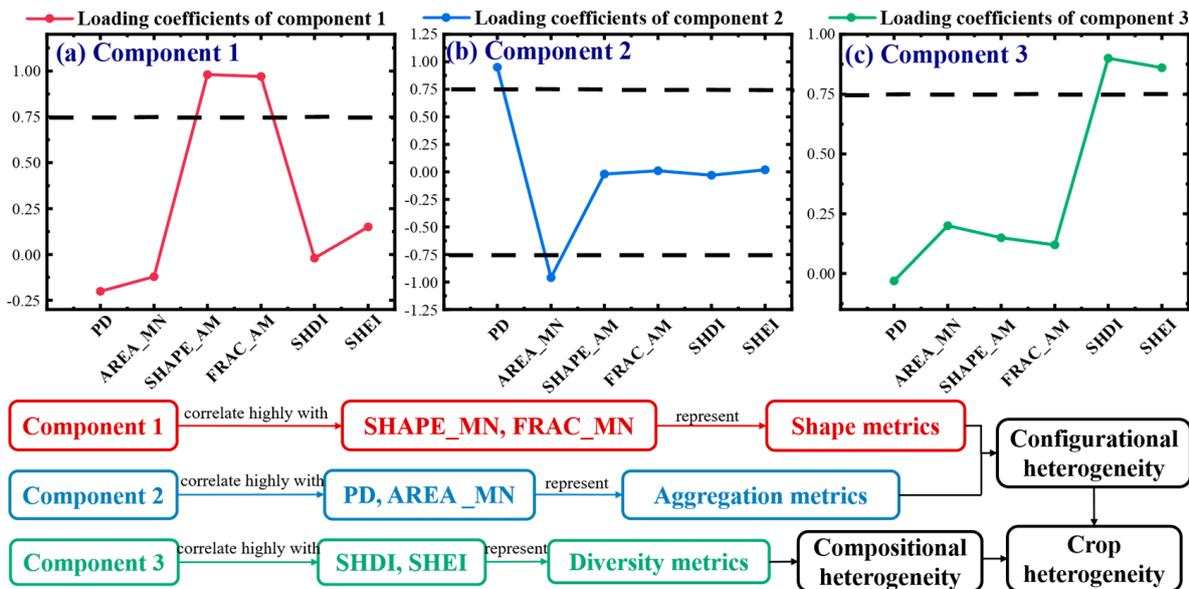
Components 1, 2, and 3 were selected as comprehensive landscape metrics and renamed as ‘Shape’, ‘Aggregation’, and ‘Diversity’, respectively, based on their dominant characteristics. These renamed components are consistently applied throughout the remainder of the paper for clarity of understanding.

### 4.2. CHZ Generation

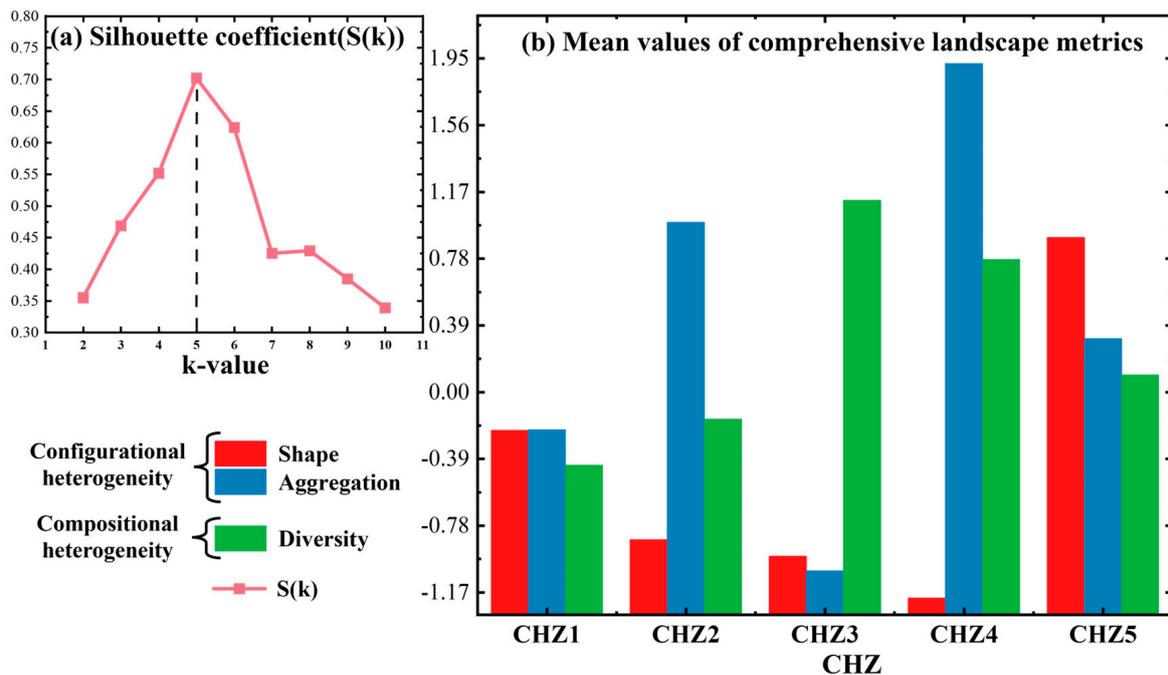
Five CHZs were identified through clustering based on the three comprehensive landscape metrics. The intermediate steps for determining the number of partitions are shown in Figure 9a. The optimal number of clusters was determined to be five, as indicated by the highest silhouette coefficient ( $S(5) = 0.702$ ).

Figure 9b presents the overall crop heterogeneity for each CHZ, while Figure 10 illustrates the spatial distribution of the CHZs, with Figure 10(a1–a5) providing details on cropland fragmentation. The degree of geomorphology and cropland fragmentation was analyzed for each CHZ to understand its unique characteristics. CHZs 1 and 3 exhibited low shape complexity and high aggregation (Figure 9b), appearing as large, regular cropland patches in Figure 10(a1,a3). These zones were primarily plains, where cropland patches were minimally influenced by topographic factors. CHZs 2 and 4 showed

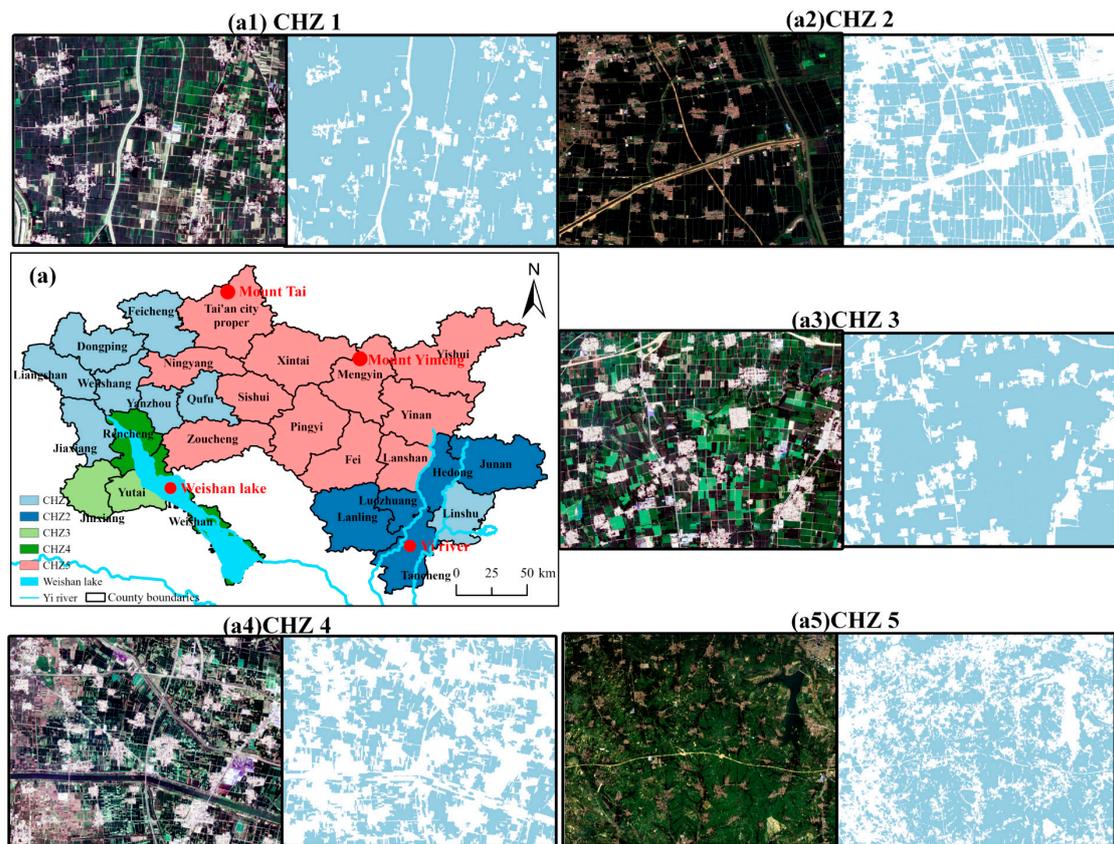
high aggregation of cropland patches characterized by regular shapes and numerous small patches (Figure 10(a2,a4)). While these zones were also dominated by plains, they were enriched with water resources, such as Weishan Lake and Yi River. The availability of water promotes irrigation-based cultivation along waterways, leading to the proliferation of small cropland patches in these areas. CHZ 5 was distinguished by fine cropland patches with intricate shapes due to the presence of rugged mountains, such as Mount Tai and Mount Yimeng, and the surrounding hills. These geographic features fragmented the cropland patches, as seen in Figure 10(a5).



**Figure 8.** Rotated principal component loadings extracted from landscape metrics. Landscape metrics with absolute loadings above 0.75, indicating significant contribution to their components, are marked with blue triangles.



**Figure 9.** Generation of CHZs and their crop heterogeneity. (a) shows the average silhouette coefficient based on Equation (4), assessing the clustering effectiveness. (b) shows the mean values of comprehensive landscape metrics for each CHZ.



**Figure 10.** The spatial distribution of CHZs. (a) shows the overall spatial distribution of CHZs. (a1–a5) are representative examples of each CHZ, showing the spatial details of arable patches. All images are derived from Sentinel 2 August mean composite images, and the cropland maps are from Section Cropland Extraction.

To analyze crop diversity in each CHZ, the crop statistics for each CHZ were surveyed (Figure 11). CHZ 1 and CHZ 2 were primarily dominated by maize, with maize accounting for 71% and 66% of the area, respectively. Other crops were cultivated to a lesser extent, contributing to the low crop diversity shown in Figure 9b. CHZ 3 had a more varied crop composition, with rice, cotton, maize, and soybean as the dominant crops, representing 31%, 36%, 16%, and 17% of the area, respectively. The even distribution of these crops within CHZ 3 resulted in a high crop diversity. CHZ 4 was predominantly planted with maize, with rice as a secondary crop, while CHZ 5 was also maize-dominated but had peanuts and potatoes as secondary crops. These two subzones shared similar crop proportions and crop diversity structures.

The above findings suggested that crop diversity and cropland fragmentation were significantly different between CHZs, highlighting the potential for CHZs in crop mapping applications.

#### 4.3. Optimal Classification Schemes Selection

The overall accuracies (OAs) of the seven classification schemes, which incorporated different feature combinations for each CHZ, are shown in Figure 12. In CHZ 1 and CHZ 2, S1 to S7 achieved OAs greater than 90%. The highest OA schemes were S3 and S6, with OAs of 96.23% and 94.66%, respectively. However, these schemes only demonstrated minimal improvements of 1.32% and 1.35% (average value of enhancements compared to other schemes), respectively. This limited enhancement may have been due to the relatively low crop heterogeneity in CHZ 1 and CHZ 2. By contrast, CHZ 3 (with the highest crop diversity in Figure 9b), CHZ 4 (with the highest fragmented cropland aggregation), and CHZ 5 (with

the highest shape complexity) exhibited higher crop heterogeneity. The schemes with the highest OAs in these zones were S7, S4, and S6, with OAs of 89.14%, 89.87%, and 93.24%, respectively. These schemes showed significant average improvements of 4.12%, 5.97%, and 4.20% over the other schemes, respectively. These results suggested that the selection of classification schemes could effectively compensate for the accuracy loss caused by high crop heterogeneity in these regions.

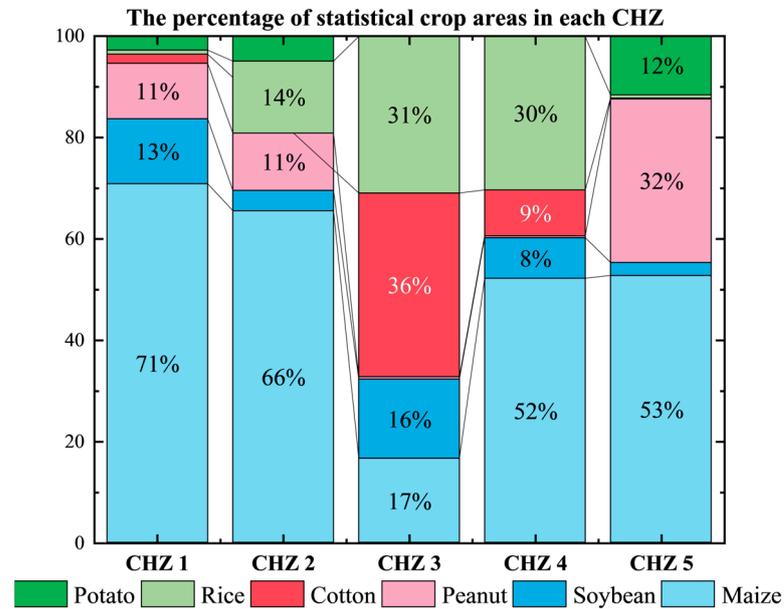


Figure 11. The percentage of statistical crop areas in each CHZ.

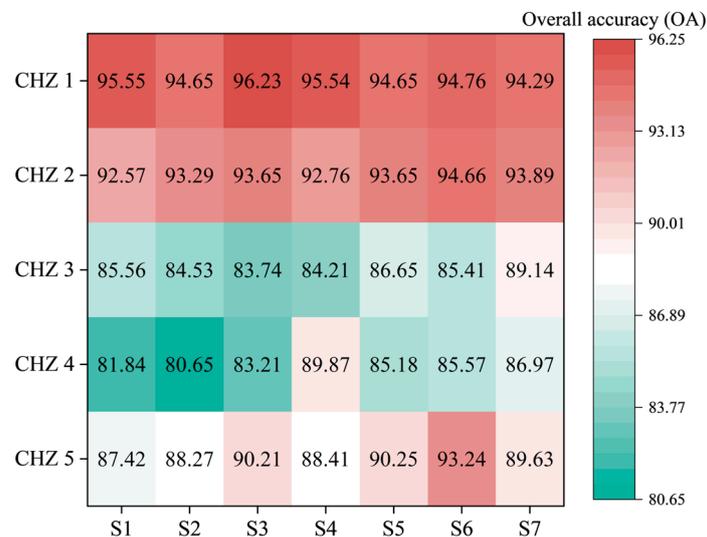
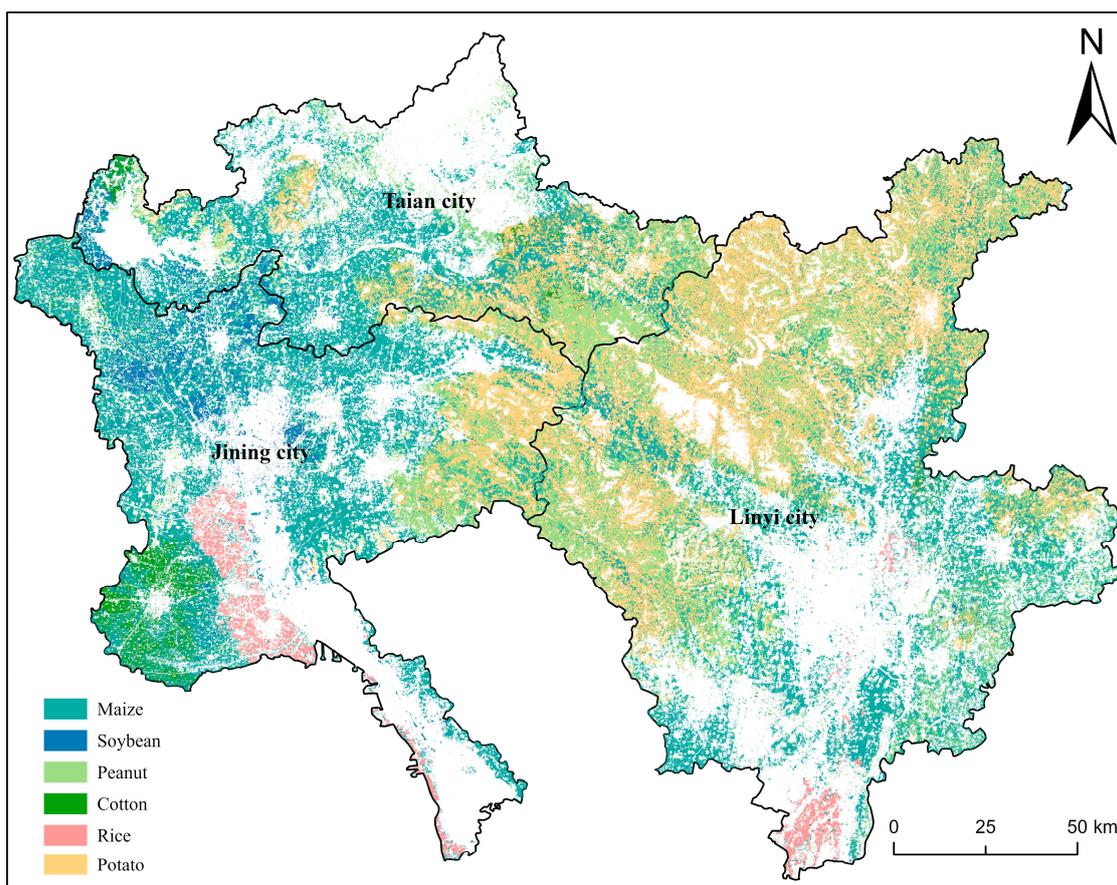


Figure 12. Optimal classification scheme selection for each CHZ. S1–S7 have the same meanings as in Table 4.

#### 4.4. Crop Mapping Results

Figure 13 shows the distribution of six crops in the cities of Tai’an, Jining, and Linyi. Maize and soybean are widely distributed across the plains and are the primary crops in all three cities. Cotton and rice are predominantly found in the southern part of the study area, where abundant water resources, such as Weishan Lake and the Yi River, and flat terrain provide favorable conditions for their cultivation. In the central part of the study area, peanuts and potatoes are the main crop types due to the hilly and mountainous terrain and dry climate.



**Figure 13.** Crop map of the study area obtained from landscape-clustering zoning strategy.

#### 4.5. Assessment of Crop Mapping

##### 4.5.1. Comparison with No-Zoning Methods

The comparison of confusion matrices for crop mapping between the landscape-clustering zoning method and the no-zoning method is shown in Table 5. The no-zoning method used S7, which integrated all available feature types in Table 4. The landscape-clustering zoning method achieved an OA of 93.52%, reflecting a 2.9% improvement over the no-zoning method. Similarly, the Kappa coefficient increased to 92.67%, representing a 3.82% improvement.

Regarding the accuracy of different crop types, the landscape-clustering zoning method achieved F1 scores exceeding 90% for all crops except potato, with an F1 of 88.74%. Compared to the no-zoning method, the landscape-clustering zoning method exhibited consistent improvements in F1 across all crops, especially for peanut and cotton, which increased by 4.42% and 5.56%, reaching F1 scores of 90.17% and 92.03%, respectively. The landscape-clustering zoning method also demonstrated balanced performance in UA and PA for all crop types. For maize, the most widely distributed crop in the study area, UA and PA exceeded 94%, while other crop types showed values around 90%. By contrast, in the no-zoning method, peanut exhibited significantly higher PA than UA and cotton exhibited higher UA than PA. This discrepancy could be attributed to spatial characteristics: peanuts are typically grown in hilly areas with fragmented arable land, while cotton is often cultivated in regions with diverse crop types. The landscape-clustering zoning method was divided into CHZ 3 and CHZ 5 based on these spatial characteristics, as detailed in Section 4.2. This zoning approach reduced the influence of crop heterogeneity on crop mapping and enhanced the accuracy of classification.

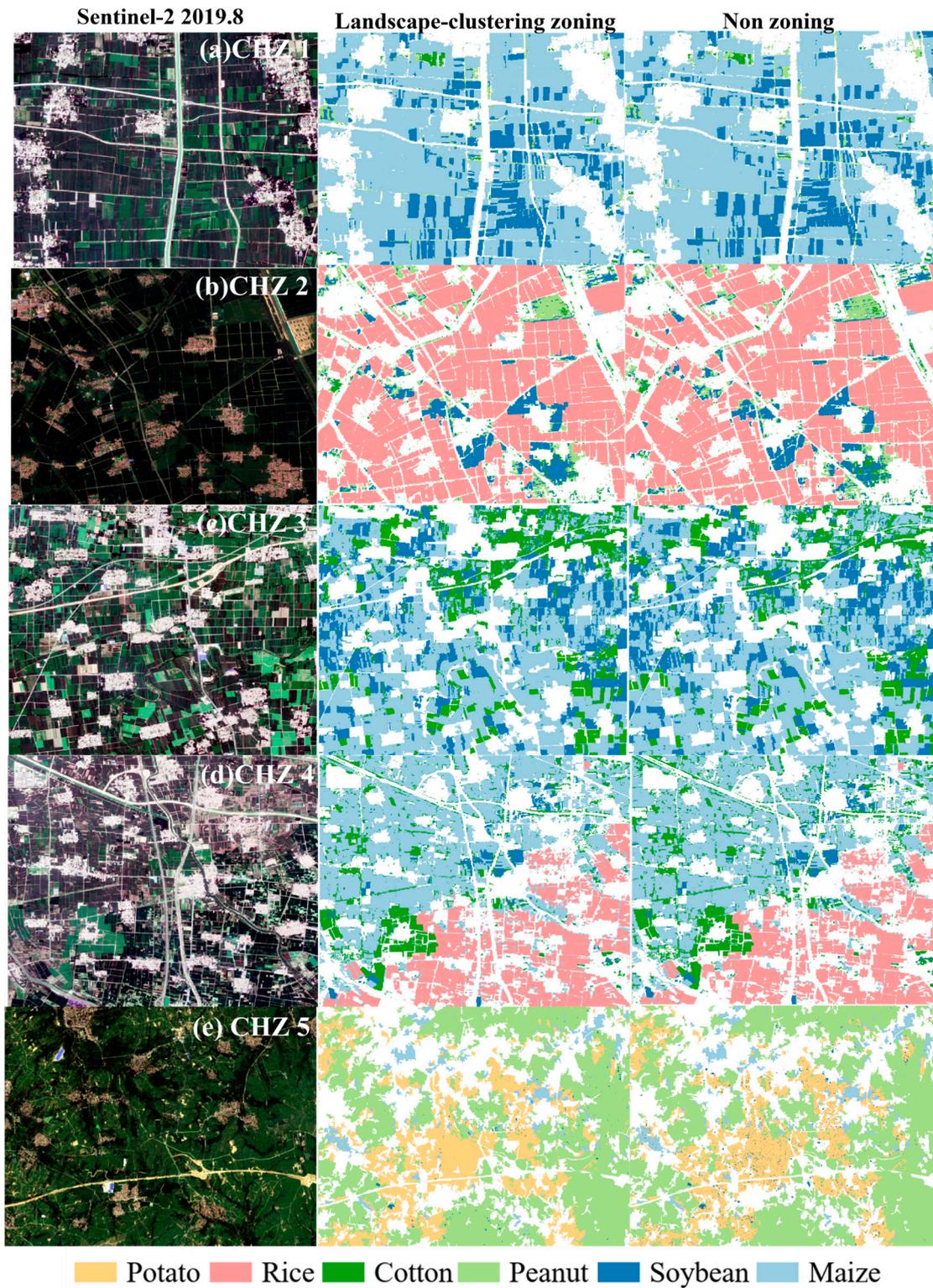
**Table 5.** Crop mapping accuracy for different methods.

|           |         | Landscape-Clustering Zoning | Non-Zoning with S7 | Improvement |
|-----------|---------|-----------------------------|--------------------|-------------|
| PA (%)    | Maize   | 95.80                       | 94.59              | 1.21        |
|           | Soybean | 93.43                       | 92.65              | 0.78        |
|           | Peanut  | 89.64                       | 88.46              | 1.18        |
|           | Cotton  | 91.60                       | 82.23              | 1.72        |
|           | Rice    | 92.72                       | 91.62              | 0.10        |
|           | Potato  | 89.23                       | 86.65              | 2.58        |
| UA (%)    | Maize   | 94.57                       | 94.41              | −0.02       |
|           | Soybean | 92.61                       | 91.35              | 0.66        |
|           | Peanut  | 90.70                       | 83.2               | 7.50        |
|           | Cotton  | 92.46                       | 89.88              | 9.15        |
|           | Rice    | 91.45                       | 89.77              | 3.68        |
|           | Potato  | 88.25                       | 87.64              | 0.61        |
| F1(%)     | Maize   | 94.16                       | 94.04              | 0.12        |
|           | Soybean | 92.71                       | 92.00              | 0.72        |
|           | Peanut  | 90.17                       | 85.75              | 4.42        |
|           | Cotton  | 92.03                       | 86.47              | 5.56        |
|           | Rice    | 95.08                       | 93.17              | 1.91        |
|           | Potato  | 88.74                       | 87.14              | 1.60        |
| OA (%)    |         | 93.52                       | 90.62              | 2.9         |
| Kappa (%) |         | 92.67                       | 88.85              | 3.82        |

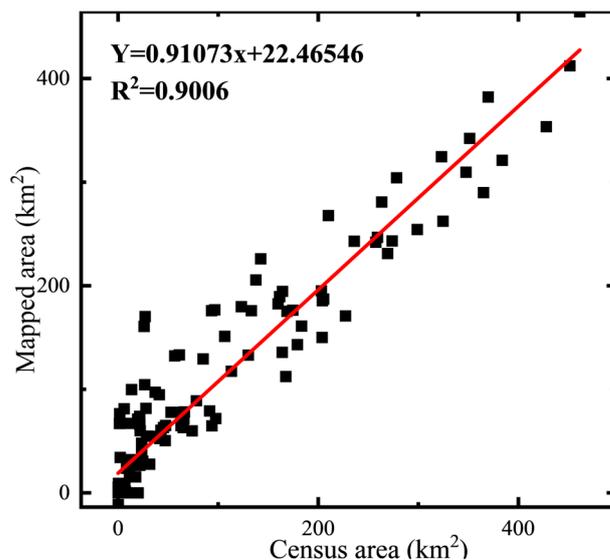
To further illustrate the advantages of the landscape-clustering zoning method for crop mapping in complex agricultural regions, we analyzed representative crop mapping results across five CHZs (Figure 14). In CHZ 1 and CHZ 2 (Figure 14a,b), characterized by regular cropland patches with low crop diversity, the landscape-clustering zoning method produced similar crop mapping results to the no-zoning approach. In CHZ 3 (Figure 14c), the landscape-clustering zoning method reduced the misclassification of cotton as maize compared to the no-zoning method. In CHZ 4 and CHZ 5 (Figure 14d,e), the landscape-clustering zoning method effectively resisted pepper noise and improved classification accuracy. These findings confirmed that the landscape-clustering zoning method yielded more accurate classification results.

#### 4.5.2. Comparison with Agricultural Statistical Data

The crop areas derived from the crop maps were compared with county-level statistical data. Because rice and cotton are distributed in only a small number of counties, the areas of all crop types were combined into a single equation to evaluate the overall fit. Figure 15 shows a strong consistency between the mapped crop area and the census data, with an  $R^2$  value of 0.9006. Most points are evenly distributed around the 1:1 line, with only a very small number of counties showing a higher mapped crop area than the agricultural statistics. These results underscored the reliability and accuracy of the area estimates of the landscape-clustering zoning method.



**Figure 14.** Spatial details of crop mapping results in five CHZs between the landscape-clustering zoning and non-zoning methods. All images were derived from Sentinel-2 satellites in August 2019.



**Figure 15.** Comparison of the mapped area of all crop types with census data at the county level. The red slash indicates that the ratio of the mapped area to census area is 1:1.

## 5. Discussion

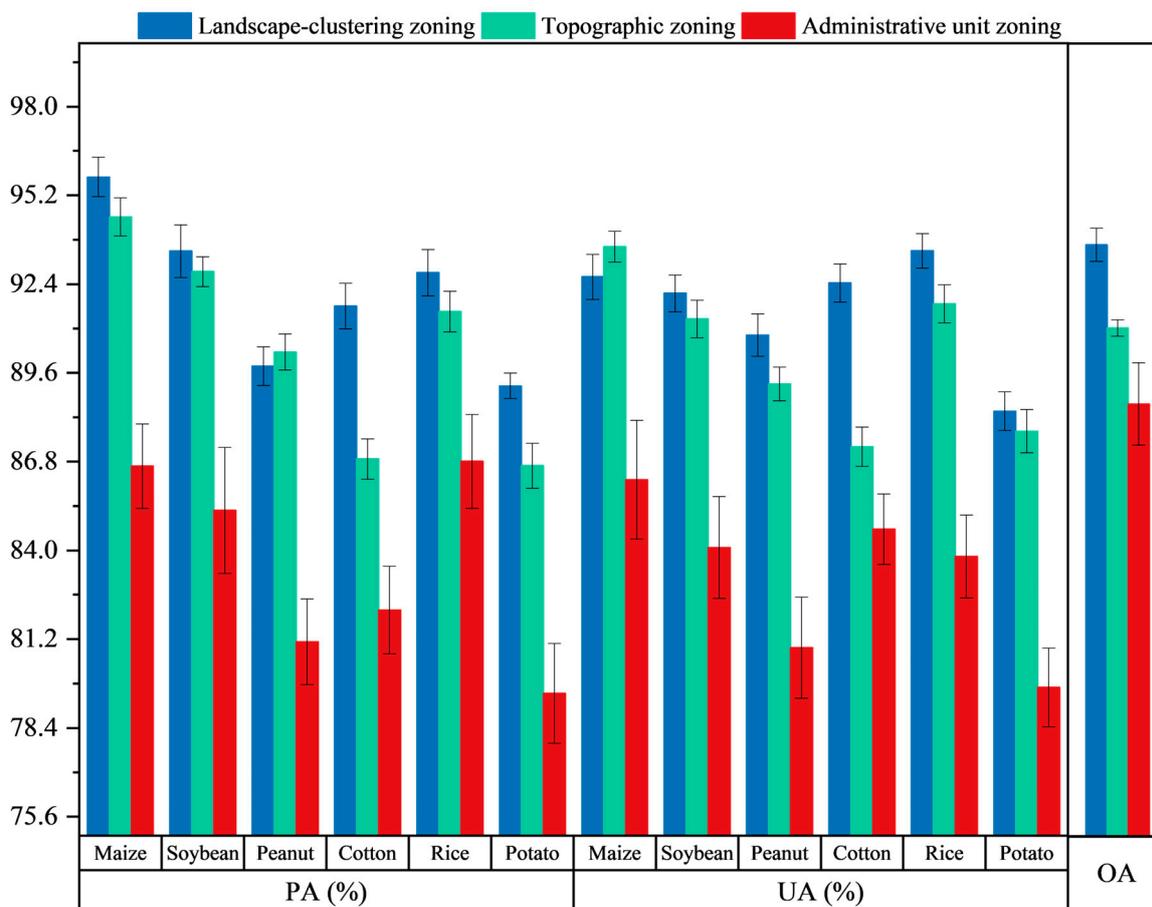
### 5.1. The Advantages of the Landscape-Clustering Zoning Strategy in Crop Mapping

Most previous studies on crop mapping have focused on large-scale regions with relatively simple crop cultivation structures [42,43]. However, the Huang-Huai-Hai region, with its complex agricultural landscape characterized by diverse crops and fragmented arable land, presents unique cropping conditions and a natural environment that are not representative of these earlier studies. Although zoning methods such as climate and topography have been proposed to address spatial heterogeneity in crop mapping [20,23], these zoning strategies are not applicable because complex agricultural landscapes are affected by multiple environmental factors.

In this study, we propose a landscape-clustering zoning strategy that divides the study area by quantifying crop diversity and cropland fragmentation through multidimensional landscape metrics. Based on this zoning, classification schemes tailored to different CHZs were selected to map crops in the Huang-Huai-Hai region's complex agricultural landscapes. To evaluate the effectiveness of the landscape-clustering zoning strategy, we compared it with two traditional zoning methods, including a topographic zoning strategy [44] based on elevation and a county-level administrative zoning strategy [17]. Each classification experiment was repeated 10 times and the mean and standard deviation of the classification accuracies are presented in Figure 16.

The landscape-clustering zoning strategy achieved the highest OA at  $93.25 \pm 0.52\%$ , followed by the topographic zoning strategy at  $91.03 \pm 0.26\%$ , and the administrative unit zoning strategy at  $88.63 \pm 1.3\%$ . The landscape-clustering zoning strategy was obviously superior to the other two strategies. In addition, the PA and UA of the landscape-clustering zoning strategy generally outperformed those of the other strategies. For peanut and potato, the PA and UA were similar to those of the topographic zoning strategy, as these crops are predominantly grown in hilly areas where topographic zoning helps to reduce spatial heterogeneity. However, for cotton, the UA and PA in the topographic zoning strategy were significantly lower than in the landscape-clustering zoning strategy, likely because the topographic zoning strategy failed to delineate the high crop diversity areas in the plains. The administrative unit zoning strategy had the lowest accuracy and the highest standard deviation for each crop type, which could be attributed to the limited number of training samples within each unit. This limitation hindered the model's ability to capture

crop-specific characteristics, resulting in underfitting, and yielded the lowest PA and UA among the three strategies, along with significantly higher errors.



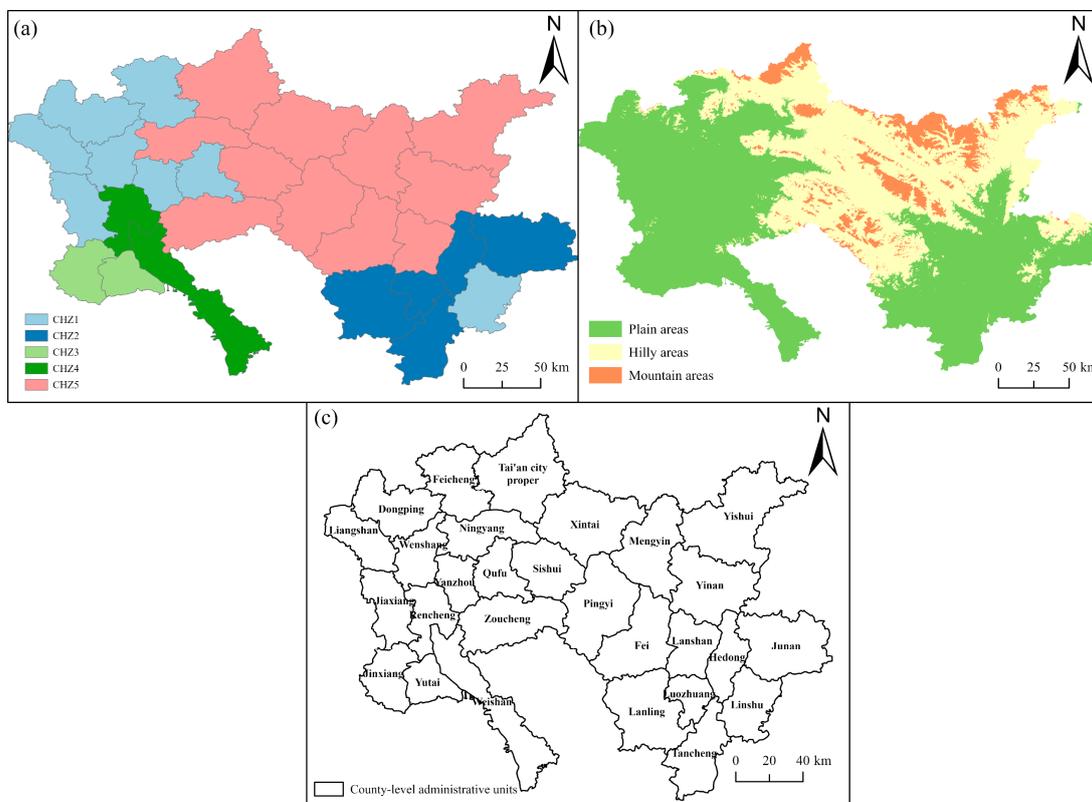
**Figure 16.** Accuracy evaluation accuracy based on different zoning methods.

We also analyzed the spatial distribution similarities and differences of subzones between the landscape-clustering and topographic zoning strategies (Figure 17). Both methods effectively delineated the mixed hilly and mountainous areas in the central part of the study region. In the landscape-clustering zoning strategy, this central region was classified as CHZ5, characterized by complex and finely fragmented cropland patches divided by hills and mountains (Figure 10(a5)). The topographic zoning strategy further subdivided this central region into smaller hilly and mountainous subzones. However, due to harsh environmental conditions, such as steep slopes and poor soil quality in the mountainous areas, crop cultivation in these regions is minimal. Additionally, the cropland mask applied in this study excluded the influence of mountainous land features, thereby providing limited benefits for crop classification. In the plains, the landscape-clustering zoning strategy demonstrated greater capability by dividing the area into four subzones (CHZ1–CHZ4) based on the degree of crop diversity and cropland fragmentation. By contrast, the topographic zoning strategy was limited to delineating a single plain subzone.

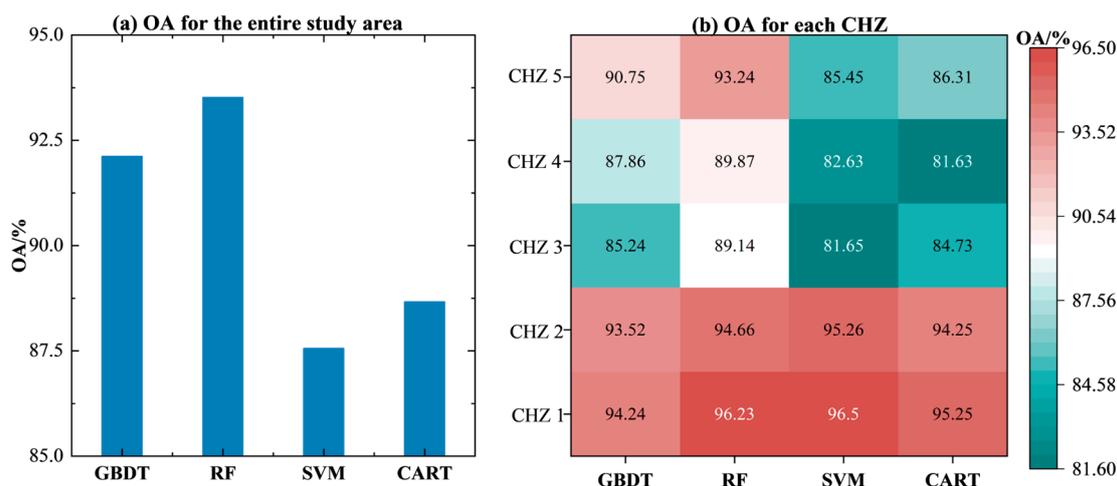
### 5.2. The Effect of Classifiers on Crop Mapping in the Landscape-Clustering Zoning Strategy

Classifiers used in crop mapping can be significantly influenced by spatial heterogeneity [45,46]. In this study, we compared the adaptability of Random Forest (RF) and other classifiers, including Gradient Boosting Decision Tree (GBDT), Support Vector Machine (SVM), and Classification and Regression Tree (CART), within the landscape-clustering zoning strategy, based on the optimal classification scheme for CHZs determined in Section 4.3.

The parameters for the classifiers were optimized using the grid search method (using a brute-force method to filter the optimal parameters of a classifier). The OA for the entire study area is presented in Figure 18a, where RF achieved the highest OA (93.52%), followed by GBDT (92.12%), while both SVM and CART had OA values below 90%.



**Figure 17.** The spatial distribution of zoning results of different zoning strategies. (a) Landscape-clustering zoning strategy (b), topographic zoning strategy, and (c) county-level administrative zoning strategy.



**Figure 18.** The OA of the entire study area and each CHZ.

The accuracy of the classifiers for each CHZ was also analyzed (Figure 18b). In CHZ 1 and CHZ 2, which had low crop heterogeneity, all four classifiers achieved OAs greater than 90% and displayed similar accuracy levels, indicating that all classifiers adapted well in areas with low crop heterogeneity. However, in CHZ 3, CHZ 4, and CHZ 5,

which exhibited higher crop heterogeneity, RF demonstrated slightly higher OAs than GBDT and significantly outperformed SVM and CART. These results highlighted RF's superior robustness in regions with high crop heterogeneity, followed by GBDT, whereas SVM and CART were less effective in addressing the challenges posed by heterogeneous crop landscapes.

### 5.3. Limitations

The landscape-clustering zoning strategy proposed in this study has demonstrated its usefulness, but several challenges remain. First, because the smallest statistical scale for agricultural statistics is at the county level, landscape metrics could only be calculated at this scale. However, we observed differences in the spatial distribution of crop diversity and cropland fragmentation within counties. Although county-level zoning yielded satisfactory results in this study, there remains room for improvement. Future studies could integrate multi-source remote sensing and statistical data to downscale agricultural statistics to finer grids (e.g., [15]). Second, this study used 10 m spatial resolution images but did not account for the mixed pixel effect. The mixed pixel effect not only influences classification accuracy but also impacts crop area statistics, especially in fragmented agricultural landscapes. The comparison between mapped areas and statistical data in this study can only be interpreted as a reference rather than a strict validation [18]. Hybrid pixel decomposition methods, such as linear hybrid models, may offer a potential solution to address the mixed pixel effect [47]. However, challenges in selecting end-members and the high computational requirements limited their application in this study. Third, the study considered only crop diversity and cropland fragmentation as factors influencing crop mapping but did not include crop phenology. When scaling up to larger study areas, the inconsistency of crop phenology across regions should be further analyzed. Finally, investigating the application of deep learning models within the landscape-clustering zoning strategy could offer better adaptability and reduce the impact of crop heterogeneity on crop mapping in future studies.

## 6. Conclusions

This study proposes a landscape-clustering zoning strategy to map crops in fragmented cropland regions using feature combinations. The comprehensive landscape metrics obtained could effectively quantify the localized complex agricultural areas in the Huang-Huai-Hai region from multiple dimensions (including cropland patch shape, aggregation, and crop diversity) and reflected the association of complex natural environments with crop diversity and cropland fragmentation. The generated CHZs divided the whole study area into different homogeneous units and selecting feature combinations tailored to each CHZ effectively compensated for accuracy losses caused by crop heterogeneity, particularly in high crop heterogeneity regions. Compared to no-zoning and other zoning strategies, the landscape-clustering zoning strategy demonstrated superior accuracy and great consistency with agricultural statistics. Unlike previous methods that rely solely on pre-acquired mapping results for calculating landscape metrics, this strategy directly computes metrics using auxiliary data, providing a methodological reference for the a priori quantification of complex agricultural landscapes. Additionally, the zoning-based feature combination approach offers a novel perspective for enhancing crop mapping accuracy in these regions. With promising results, the landscape-clustering zoning strategy shows potential for application in large-scale regions and diverse cropping environments. Future studies will evaluate its integration with sample generation strategies to further refine its utility and scalability.

**Author Contributions:** G.F.: Project administration, Supervision, Writing—original draft. C.W.: Resources, investigation. T.D.: Supervision, Writing—review & editing. Z.W., M.L., J.C. and C.C.: Resources, investigation. H.Z.: Conceptualization, Supervision, Writing review & editing, Funding acquisition. All authors have read and agreed to the published version of the manuscript.

**Funding:** The authors acknowledge financial support from the National Natural Science Foundation of China (Grant No. 42471364), the Key Laboratory of Spatial Data Mining & Information Sharing of the Ministry of Education, Fuzhou University (Grant No. 2023LSDMIS05), the Natural Science Foundation of Shandong Province (Grant No. ZR2024MD004), and the Jinan City Municipal-School Integration Development Strategic Project (Grant No. JNSX2023036).

**Institutional Review Board Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available upon request from the corresponding author. The data are not publicly available due to privacy concerns. Requests for data should include a brief description of the intended use or research purpose for evaluation of sharing eligibility.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Zhang, X.; Wu, B.; Ponce-Campos, G.E.; Zhang, M.; Chang, S.; Tian, F. Mapping Up-to-Date Paddy Rice Extent at 10 M Resolution in China through the Integration of Optical and Synthetic Aperture Radar Images. *Remote Sens.* **2018**, *10*, 1200. [[CrossRef](#)]
- Kuenzer, C.; Knauer, K. Remote Sensing of Rice Crop Areas. *Int. J. Remote Sens.* **2013**, *34*, 2101–2139. [[CrossRef](#)]
- Bauer, M.E. The Role of Remote Sensing in Determining the Distribution and Yield of Crops. *Adv. Agron.* **1975**, *27*, 271–304.
- Taylor, J.C.; Wood, G.A.; Thomas, G. Mapping Yield Potential with Remote Sensing. In Proceedings of the Papers presented at the First European Conference on Precision Agriculture, Coventry, UK, 7–10 September 1997.
- Ma, Z.; Li, W.; Warner, T.A.; He, C.; Wang, X.; Zhang, Y.; Guo, C.; Cheng, T.; Zhu, Y.; Cao, W.; et al. A Framework Combined Stacking Ensemble Algorithm to Classify Crop in Complex Agricultural Landscape of High Altitude Regions with Gaofen-6 Imagery and Elevation Data. *Int. J. Appl. Earth Obs. Geoinf.* **2023**, *122*, 103386. [[CrossRef](#)]
- Song, X.-P.; Potapov, P.V.; Krylov, A.; King, L.; Di Bella, C.M.; Hudson, A.; Khan, A.; Adusei, B.; Stehman, S.V.; Hansen, M.C. National-Scale Soybean Mapping and Area Estimation in the United States Using Medium Resolution Satellite Imagery and Field Survey. *Remote Sens. Environ.* **2017**, *190*, 383–395. [[CrossRef](#)]
- Boryan, C.; Yang, Z.; Mueller, R.; Craig, M. Monitoring US Agriculture: The US Department of Agriculture, National Agricultural Statistics Service, Cropland Data Layer Program. *Geocarto Int.* **2011**, *26*, 341–358. [[CrossRef](#)]
- Kang, X.; Huang, C.; Chen, J.M.; Lv, X.; Wang, J.; Zhong, T.; Wang, H.; Fan, X.; Ma, Y.; Yi, X.; et al. The 10-m Cotton Maps in Xinjiang, China during 2018–2021. *Sci. Data* **2023**, *10*, 688. [[CrossRef](#)] [[PubMed](#)]
- Mondal, S.; Jeganathan, C. Effect of Scale, Landscape Heterogeneity and Terrain Complexity on Agriculture Mapping Accuracy from Time-Series NDVI in the Western-Himalaya Region. *Landsc. Ecol.* **2022**, *37*, 2757–2781. [[CrossRef](#)]
- Mondal, S.; Jeganathan, C. Evaluating the Performance of Multi-Class and Single-Class Classification Approaches for Mountain Agriculture Extraction Using Time-Series NDVI. *J. Indian Soc. Remote Sens.* **2018**, *46*, 2045–2055. [[CrossRef](#)]
- Zhang, H.; Wang, Y.; Shang, J.; Liu, M.; Li, Q. Investigating the Impact of Classification Features and Classifiers on Crop Mapping Performance in Heterogeneous Agricultural Landscapes. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *102*, 102388. [[CrossRef](#)]
- Chen, Y.; Song, X.; Wang, S.; Huang, J.; Mansaray, L.R. Impacts of Spatial Heterogeneity on Crop Area Mapping in Canada Using MODIS Data. *ISPRS J. Photogramm. Remote Sens.* **2016**, *119*, 451–461. [[CrossRef](#)]
- Han, J.; Zhang, Z.; Cao, J.; Luo, Y. Mapping Rapeseed Planting Areas Using an Automatic Phenology- and Pixel-Based Algorithm (APPA) in Google Earth Engine. *Crop J.* **2022**, *10*, 1483–1495. [[CrossRef](#)]
- Cano, E.; Denux, J.-P.; Bisquert, M.; Hubert-Moy, L.; Chéret, V. Improved Forest-Cover Mapping Based on MODIS Time Series and Landscape Stratification. *Int. J. Remote Sens.* **2017**, *38*, 1865–1888. [[CrossRef](#)]
- Mohammed, I. A Blended Census and Multiscale Remote Sensing Approach to Probabilistic Cropland Mapping in Complex Landscapes. *ISPRS J. Photogramm. Remote Sens.* **2020**, *161*, 233–245. [[CrossRef](#)]
- Vintrou, E.; Desbrosse, A.; Bégué, A.; Traoré, S.; Baron, C.; Lo Seen, D. Crop Area Mapping in West Africa Using Landscape Stratification of MODIS Time Series and Comparison with Existing Global Land Products. *Int. J. Appl. Earth Obs. Geoinf.* **2012**, *14*, 83–93. [[CrossRef](#)]

17. Bamud, S. A Remotely Sensed Based Comparison of Three Area Stratification Methods to Improve Estimation of Crop Area Statistics, A Case Study in Fragmented Landscapes of Ethiopia. Master's Thesis, University of Twente, Enschede, The Netherlands, 2022.
18. Dong, J.; Fu, Y.; Wang, J.; Tian, H.; Fu, S.; Niu, Z.; Han, W.; Zheng, Y.; Huang, J.; Yuan, W. Early-Season Mapping of Winter Wheat in China Based on Landsat and Sentinel Images. *Earth Syst. Sci. Data* **2020**, *12*, 3081–3095. [[CrossRef](#)]
19. Phalke, A.R.; Özdoğan, M.; Thenkabail, P.S.; Erickson, T.; Gorelick, N.; Yadav, K.; Congalton, R.G. Mapping Croplands of Europe, Middle East, Russia, and Central Asia Using Landsat, Random Forest, and Google Earth Engine. *ISPRS J. Photogramm. Remote Sens.* **2020**, *167*, 104–122. [[CrossRef](#)]
20. You, N.; Dong, J.; Huang, J.; Du, G.; Zhang, G.; He, Y.; Yang, T.; Di, Y.; Xiao, X. The 10-m Crop Type Maps in Northeast China during 2017–2019. *Sci. Data* **2021**, *8*, 41. [[CrossRef](#)]
21. Cordero-Sancho, S.; Sader, S.A. Spectral Analysis and Classification Accuracy of Coffee Crops Using Landsat and a Topographic-environmental Model. *Int. J. Remote Sens.* **2007**, *28*, 1577–1593. [[CrossRef](#)]
22. Liu, Y.; Yu, Q.; Zhou, Q.; Wang, C.; Bellingrath-Kimura, S.D.; Wu, W. Mapping the Complex Crop Rotation Systems in Southern China Considering Cropping Intensity, Crop Diversity, and Their Seasonal Dynamics. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2022**, *15*, 9584–9598. [[CrossRef](#)]
23. Ren, T.; Xu, H.; Cai, X.; Yu, S.; Qi, J. Smallholder Crop Type Mapping and Rotation Monitoring in Mountainous Areas with Sentinel-1/2 Imagery. *Remote Sens.* **2022**, *14*, 566. [[CrossRef](#)]
24. Donmez, E.; Yilmaz, M.T.; Yucel, I. Added Utility of Temperature Zone Information in Remote Sensing-Based Large Scale Crop Mapping. *Remote Sens. Appl. Soc. Environ.* **2024**, *35*, 101264. [[CrossRef](#)]
25. Sirami, C.; Gross, N.; Baillod, A.B.; Bertrand, C.; Carrié, R.; Hass, A.; Henckel, L.; Miguet, P.; Vuillot, C.; Alignier, A.; et al. Increasing Crop Heterogeneity Enhances Multitrophic Diversity across Agricultural Regions. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 16442–16447. [[CrossRef](#)] [[PubMed](#)]
26. Alignier, A.; Solé-Senan, X.O.; Robleño, I.; Baraibar, B.; Fahrig, L.; Giral, D.; Gross, N.; Martin, J.-L.; Recasens, J.; Sirami, C.; et al. Configurational Crop Heterogeneity Increases Within-Field Plant Diversity. *J. Appl. Ecol.* **2020**, *57*, 654–663. [[CrossRef](#)]
27. Priyadarshana, T.S.; Lee, M.; Ascher, J.S.; Qiu, L.; Goodale, E. Crop Heterogeneity Is Positively Associated with Beneficial Insect Diversity in Subtropical Farmlands. *J. Appl. Ecol.* **2021**, *58*, 2747–2759. [[CrossRef](#)]
28. Oliphant, A.J.; Thenkabail, P.S.; Teluguntla, P.; Xiong, J.; Gumma, M.K.; Congalton, R.G.; Yadav, K. Mapping Cropland Extent of Southeast and Northeast Asia Using Multi-Year Time-Series Landsat 30-m Data Using a Random Forest Classifier on the Google Earth Engine Cloud. *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *81*, 110–124. [[CrossRef](#)]
29. Segarra, J.; Buchailot, M.L.; Araus, J.L.; Kefauver, S.C. Remote Sensing for Precision Agriculture: Sentinel-2 Improved Features and Applications. *Agronomy* **2020**, *10*, 641. [[CrossRef](#)]
30. Yang, J.; Huang, X. The 30 m Annual Land Cover Dataset and Its Dynamics in China from 1990 to 2019. *Earth Syst. Sci. Data* **2021**, *13*, 3907–3925. [[CrossRef](#)]
31. Nagraj, G.M.; Karegowda, A.G. Crop Mapping Using SAR Imagery: An Review. *Int. J. Adv. Res. Comput. Sci.* **2016**, *7*, 47–52.
32. Haralick, R.M.; Shanmugam, K.; Dinstein, I. Textural Features for Image Classification. *IEEE Trans. Syst. Man Cybern.* **1973**, *SMC-3*, 610–621. [[CrossRef](#)]
33. Hao, P.; Zhan, Y.; Wang, L.; Niu, Z.; Shakir, M. Feature Selection of Time Series MODIS Data for Early Crop Classification Using Random Forest: A Case Study in Kansas, USA. *Remote Sens.* **2015**, *7*, 5347–5369. [[CrossRef](#)]
34. Khosravi, I.; Safari, A.; Homayouni, S. MSMD: Maximum Separability and Minimum Dependency Feature Selection for Cropland Classification from Optical and Radar Data. *Int. J. Remote Sens.* **2018**, *39*, 2159–2176. [[CrossRef](#)]
35. Zanaga, D.; Van De Kerchove, R.; De Keersmaecker, W.; Souverijns, N.; Brockmann, C.; Quast, R.; Wevers, J.; Grosu, A.; Paccini, A.; Vergnaud, S.; et al. ESA WorldCover 10 m 2020 V100 2021. Available online: <https://esa-worldcover.org/en/data-access> (accessed on 13 January 2025).
36. Cumming, S.; Vernier, P. Statistical Models of Landscape Pattern Metrics, with Applications to Regional Scale Dynamic Forest Simulations. *Landsc. Ecol.* **2002**, *17*, 433–444. [[CrossRef](#)]
37. MacQueen, J. Some Methods for Classification and Analysis of Multivariate Observations. *Berkeley Symp. Math. Statist. Prob.* **1967**, *1*, 281–297.
38. Chen, C.; Dong, T.; Wang, Z.; Wang, C.; Song, W.; Zhang, H. Exploring Optimal Features and Image Analysis Methods for Crop Type Classification from the Perspective of Crop Landscape Heterogeneity. *Remote Sens. Appl. Soc. Environ.* **2024**, *36*, 101308. [[CrossRef](#)]
39. Wang, X.; Liu, J.; Peng, P.; Chen, Y.; He, S.; Yang, K. Automatic Crop Classification Based on Optimized Spectral and Textural Indexes Considering Spatial Heterogeneity. *Remote Sens.* **2023**, *15*, 5550. [[CrossRef](#)]
40. Ok, A.O.; Akar, O.; Gungor, O. Evaluation of Random Forest Method for Agricultural Crop Classification. *Eur. J. Remote Sens.* **2012**, *45*, 421–432. [[CrossRef](#)]

41. Hripcsak, G. Agreement, the F-Measure, and Reliability in Information Retrieval. *J. Am. Med. Inform. Assoc.* **2005**, *12*, 296–298. [[CrossRef](#)]
42. Ashourloo, D.; Shahrabi, H.S.; Azadbakht, M.; Rad, A.M.; Aghighi, H.; Radiom, S. A Novel Method for Automatic Potato Mapping Using Time Series of Sentinel-2 Images. *Comput. Electron. Agric.* **2020**, *175*, 105583. [[CrossRef](#)]
43. Xun, L.; Zhang, J.; Cao, D.; Yang, S.; Yao, F. A Novel Cotton Mapping Index Combining Sentinel-1 SAR and Sentinel-2 Multispectral Imagery. *ISPRS J. Photogramm. Remote Sens.* **2021**, *181*, 148–166. [[CrossRef](#)]
44. Peng, X.; Wang, J.; Raed, M.; Gari, J. Land Cover Mapping from RADARSAT Stereo Images in a Mountainous Area of Southern Argentina. *Can. J. Remote Sens.* **2003**, *29*, 75–87. [[CrossRef](#)]
45. You, J.; Pei, Z.; Wang, D. Crop Mapping of Complex Agricultural Landscapes Based on Discriminant Space. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 4356–4367. [[CrossRef](#)]
46. Saini, R.; Ghosh, S.K. Crop Classification in a Heterogeneous Agricultural Environment Using Ensemble Classifiers and Single-Date Sentinel-2A Imagery. *Geocarto Int.* **2021**, *36*, 2141–2159. [[CrossRef](#)]
47. Ozdogan, M.; Woodcock, C.E. Resolution Dependent Errors in Remote Sensing of Cultivated Areas. *Remote Sens. Environ.* **2006**, *103*, 203–217. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.