land

MDPI

Article

Assessing the Accuracy and Consistency of Cropland Products in the Middle Yangtze Plain

Haixia Xu [1,2], Luguang Jiang [1,2,*] and Ye Liu [1]

[1] Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China
[2] University of Chinese Academy of Sciences, Beijing 100049, China
* Correspondence: jianglg@igsnrr.ac.cn; Tel.: +86-010-64889471

Abstract: With the evolution of remote sensing, more data products concerning cropland distribution are becoming available. However, the accuracy and consistency across all datasets in crucial regions are inherently uncertain. We delved into the Middle Yangtze Plain, a complex and vital agricultural area with relatively high cultivation intensities in China. We used confusion matrices and consistency analysis to compare the accuracy and consistency of four multi-year cropland distribution data products. These include Global Land Analysis & Discovery Cropland Data (GLAD), Annual Global Land Cover (AGLC), the China Land Cover Dataset (CLCD), and China's Annual Cropland Dataset (CACD). Key findings include the following: GLAD has the highest precision at 96.09%, the CLCD has the highest recall at 98.41%, and AGLC and CACD perform well in achieving a balance between precision and recall, with F1 scores of 90.30% and 90.74%, respectively. In terms of consistency, GLAD and the CLCD show inconsistency at 69.58%. When all four products unanimously classify a pixel as cropland, the identified cropland area closely corresponds to the statistical data reported in the yearbook. The Jianghan Plain holds the majority of cropland in the Middle Yangtze Plain, constituting 50.88%. From 2003 to 2019, the cropland area experienced fluctuating and ascending trends. Shangrao City witnessed the most notable rise in cropland area, with an increase of 323.0 km^2, whereas Wuhan City underwent the most substantial decline, amounting to 185.8 km^2. These findings contribute valuable insights into the precision and consistency of existing cropland distribution products, offering a foundation for further research.

Keywords: cropland distribution; Middle Yangtze Plain; GLAD (Global Land Analysis & Discovery); AGLC (Annual Global Land Cover); CLCD (China Land Cover Dataset); CACD (China's Annual Cropland Dataset)

Citation: Xu, H.; Jiang, L.; Liu, Y. Assessing the Accuracy and Consistency of Cropland Products in the Middle Yangtze Plain. *Land* **2024**, *13*, 301. https://doi.org/10.3390/land13030301

Academic Editor: Alexandru-Ionuţ Petrişor

Received: 9 January 2024
Revised: 15 February 2024
Accepted: 26 February 2024
Published: 28 February 2024

1. Introduction

Food security is a crucial cornerstone for maintaining social stability and economic development in China. Human activities exert a more significant influence on food security when contrasted with the impact of climate factors [1]. Cropland plays a crucial role in supporting food production and is also a vital land use category. The conversion between cropland and other land use types (such as forests and construction land) is closely related to the region's food security, ecological environment, and economic development. During development, human activities can lead cropland to expand into natural areas like forests or grasslands [2,3], while construction activities encroach upon cropland [4,5]. This urban development leads to the fragmentation of cropland in China [6]. To control the quantity and quality of cropland in our country, policies such as cropland occupation and compensation balance have been implemented [7], and cropland protection redlines and permanent basic farmland have been designated. In order to ensure ecological security, China has carried out a Grain to Green Project [8]. Long-term cropland distribution data form the basis for studying the response of cropland to relevant policies [9]. According to

several national land spatial surveys in China, there have been significant changes in the distribution of cropland in the past [10].

The Middle Yangtze Plain possesses high-quality cropland resources, serving as a crucial component of China's economic development and concurrently representing a significant region for ecological environment protection and restoration along the Yangtze River [11]. Since 1980, land use changes in cropland in China have primarily shifted toward construction land [12]. The Middle Yangtze Plain features a flat terrain, abundant sunlight, and ample water resources, making it suitable for crop cultivation. Encompassing the Jianghan Plain, Dongting Lake Plain, and Poyang Lake Plain, it is a vital grain production area among China's nine major grain bases [6,13]. Moreover, the urban cluster in the Middle Yangtze region, known as the Central Yangtze River City Cluster or the "Central Triangle," is an integral part of the Yangtze River Economic Belt. It undertakes the critical tasks of promoting the rise of the central region and facilitating new urbanization strategies. The Central Yangtze River City Cluster, including the Wuhan Metropolitan Area, the Changsha–Zhuzhou–Xiangtan Urban Agglomeration, and the Poyang Lake Urban Agglomeration, has emerged as a new growth pole for China's economic development [14]. Simultaneously, the region boasts two of China's largest freshwater lakes: Poyang Lake and Dongting Lake. Consequently, the Middle Yangtze Plain holds special significance in terms of ecological environment protection and restoration. Notably, from 1975 to 2015, the predominant shift in land use changes in lakes in the Middle Yangtze region was toward cropland [15]. In summary, the changes in cropland in this region have become a focal point of attention for multiple policies.

With the continuous advancement of remote sensing technology, the production of land cover and land use data products has witnessed substantial development. In terms of spatial resolution, existing data products have progressed from 1 km to 10 m (e.g., FROM-GLC10 [16], Esri Land Cover [17], WorldCover [18]). Moreover, in terms of temporal resolution, "near real-time" releases have been achieved (Dynamic World [19]). As a critical land use type, cropland has also seen the development of specialized datasets such as Global Land Analysis & Discovery Cropland Data (GLAD) [9] and China's Annual Cropland Dataset (CACD) [20]. However, despite the continuous improvement in the temporal and spatial resolution of cropland distribution data, the accuracy and consistency of various data products in key regions remain unclear. Additionally, remote sensing classification products often face challenges of accurately classifying similar objects in geographically heterogeneous areas [9]. In this study, we aim to analyze four 30 m datasets with high temporal resolution to assess their accuracy and consistency in the Middle Yangtze Plain. We also explore and discuss how to leverage existing datasets to enhance the accuracy of cropland distribution in specific region. The findings may offer valuable insights for the effective regional application of these large-region products.

2. Materials and Methods

2.1. Study Area

The Middle Yangtze Plain spans three provinces in central China: Hunan, Hubei, and Jiangxi, between 28° N and 31.5° N latitude and 111° E–117° E longitude (Figure 1). This area is characterized by a subtropical monsoon climate. The region receives average annual precipitation of 950–1800 mm, with an average annual temperature of 10–18 °C and average relative humidity of 75–82%. The annual sunshine hours range from 1400 to 2000. The study region is a typical alluvial plain, including the Dongting Lake Plain, Jianghan Plain, E'dong Plain Along the Yangtze River, and Poyang Lake Plain. The soil is fertile, with loose texture, primarily composed of human-modified soil, strongly leached soil, and alluvial soil. The terrain is predominantly low and gently sloping, with approximately 86% of the region below 100 m in elevation and 85% having a slope less than 4°. Moreover, this area boasts the largest number of lakes in China, spanning a total area of 2 million hectares, constituting 16.9% of the region.

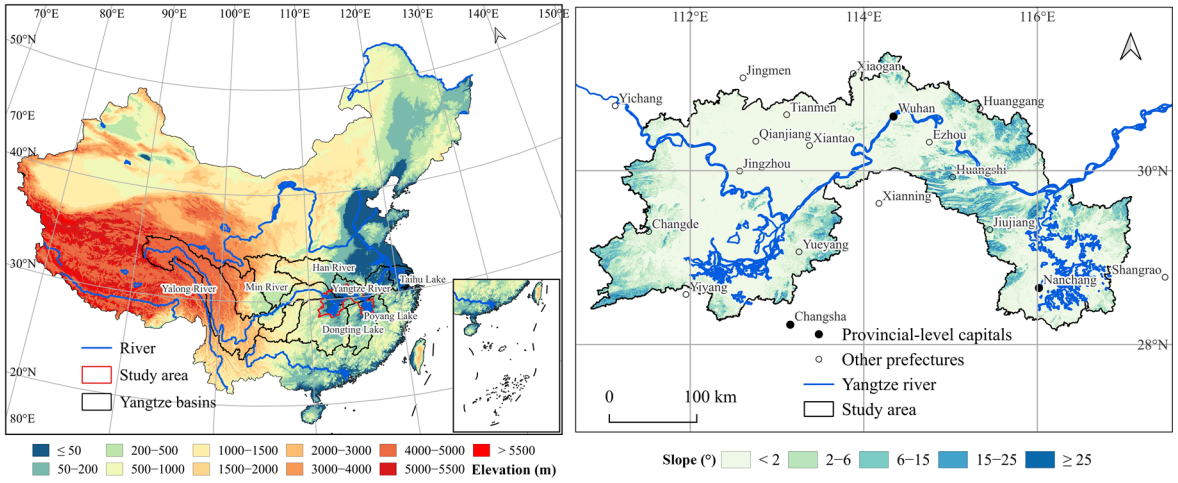

Figure 1. The geographical and topographic characteristics of the study area (Projection: Krasovsky 1940 Albers; Geodetic System: WGS 84).

2.2. Materials

Based on the study region and time interval, we selected four data products to compare (Table 1): Global Land Analysis & Discovery Cropland Data (GLAD) [9], Annual Global Land Cover (AGLC) [21], the China Land Cover Dataset (CLCD) [22], and China's Annual Cropland Dataset (CACD) [20]. We also conducted a comparison between the statistical data of administrative regions derived from the products and the corresponding level of China Yearbook statistics. The regions meeting these criteria are within Hubei Province. The cropland area data for these regions are sourced from the "Hubei Agricultural Statistical Yearbook," accessible from the China National Knowledge Infrastructure (CNKI) [23].

GLAD is a dataset that contains five periods (2003, 2007, 2011, 2015, 2019) of global 30 m cropland distribution data. The AGLC dataset covers global land cover and land use data from 2000 to 2015. To create this dataset, data from Globeland30, FROM-GLC, and GLC-FCS30 were combined to produce a single-year product called AGLC-2015. For the same year, a specific classification category was replaced with land use and land cover data from Global Annual Urban Dynamics (GAUD), Global Forest Change (GFC), and the Global Surface Water dataset (GSW). The CLCD contains annual land cover and land use data for China from 1990 to 2019. The dataset's samples are primarily based on the CLUDs (China's land use/cover datasets) and validated through samples from Geo-Wiki and GLCVSS. CACD contains annual cropland distribution data for China from 1986 to 2021. Potential sample points are generated based on CLCD and CLUD products. Then, Time-Weighted Dynamic Time Warping (TWDTW) is used to compare the spectral time series between potential sample points and high-quality sample points (obtained through visual interpretation), eliminating points with low similarity. To generate multi-year cropland distribution, the LandTrendr algorithm is employed, and the GSW and Global Impervious Surface dataset (GAIA) are used for non-cropland masking.

Each product uses different techniques and procedures to enhance accuracy and efficiency in the classification process. Regional differences are addressed by using regional models and diverse predictor variables. For example, GLAD selects 924 representative grid cells for various agricultural settings to consider the heterogeneity of geographical landscapes, while CACD introduces the concept of agricultural zones to better capture regional characteristics. GLAD incorporates phenological indicators based on NDVI in its classification features, while the CLCD integrates location features into its production process. CACD employs the TWDTW method for automated sample generation and precision control, aligning and generating samples based on spectral time series with high-quality samples to improve efficiency. AGLC uses the CCDC algorithm for land change detection, while CACD utilizes the LandTrendr algorithm for the generation of multi-year products.

Table 1. Overview of product datasets.

Dataset	Time Period	Reference Products	Classification Method	Features	Accuracy	Source
GLAD	2003, 2007, 2011, 2015, 2019	Global Food-and-Water Security-support Analysis Data (GFSAD)	Bagging Decision Trees (1° × 1°)	Bands and their linear combinations, phenological features, topographical features	98.3% ± 1.1% (OA)	https://glad.umd.edu/dataset/croplands (accessed on 25 February 2024)
AGLC	2000–2015	FROM-GLC, Globeland30, FROM-GLC, GLC-FCS30, GAUD, GFC, GSW, ESA CCI-LC	Random forest (4° × 4°)	Bands and their linear combinations	76.1% (AGLC-2015 OA); Cropland UA85.3%, PA74.23%	https://code.earthengine.google.com/?asset=users/xxc/GLC_2000_2015 (accessed on 25 February 2024)
CLCD	1990–2019	CLUDs	Random forest (0.5° hexagonal grid)	Bands and their linear combinations, time features, topographical features, location features	79.30% ± 1.99% (OA); Cropland UA77.73%, PA73.66%	https://zenodo.org/records/5816591 (accessed on 25 February 2024)
CACD	1986–2021	CLCD, CLUD, GSW, GAIA	Random forest (0.8° × 0.8°)	Bands and their linear combinations, topographical features	OA (93% ± 1%)	https://zenodo.org/records/7936885 (accessed on 25 February 2024)

Note: OA—Overall accuracy, UA—User's Accuracy, PA—Producer's Accuracy.

2.3. Methods

This research focused on the Middle Yangtze Plain to assess the accuracy and consistency of different cropland distribution products. Cropland sample point labels were derived through the visual interpretation of high-resolution imagery from Google Earth Pro. Considering the platform's image availability and the products' timeframe, the year of approximately 2011 was chosen for comparison. The comparison and analysis were performed using a GIS platform.

To guarantee sample point randomness and sufficiency, we partitioned the study region into hexagons and randomly selected five sample points within each hexagon (Figure 2). The use of hexagons could improve the sampling strategy by establishing a regular grid, ensuring uniform coverage, and reducing bias associated with irregular shapes. With fewer edges, hexagons decrease bias near partition edges and optimize nearest-neighbor distances, fostering a more representative distribution across the region for reliable and unbiased data. Points located outside the study region boundary were excluded, resulting in a total of 5986 sample points. Of these, 3452 were recognized as cropland through visual interpretation comparison. The assessment of various cropland distribution products utilized a confusion matrix with User's Accuracy (UA), Producer's Accuracy (PA), overall accuracy (OA), and F1 score as evaluation metrics.

Furthermore, we conducted a comparative analysis of the accuracy of multiple cropland distribution products under various identification scenarios. Leveraging multiple products for cropland identification overcomes limitations associated with relying on a single product. Depending solely on one product can introduce errors and incomplete data due to variations in accuracy, coverage, timing, resolution, and algorithmic biases. The considered scenarios were as follows: (1) All four products identifying the point as cropland. (2) Three or more products identifying the point as cropland. (3) Two or more products considering the point as cropland. (4) Products GLAD, AGLC, and CLCD collectively determining the point as cropland. (5) Products GLAD, AGLC, and CACD collectively determining the point as cropland. (6) Products GLAD, CLCD, and CACD collectively determining the point as cropland. (7) Products AGLC, CLCD, and CACD collectively determining the point as cropland.

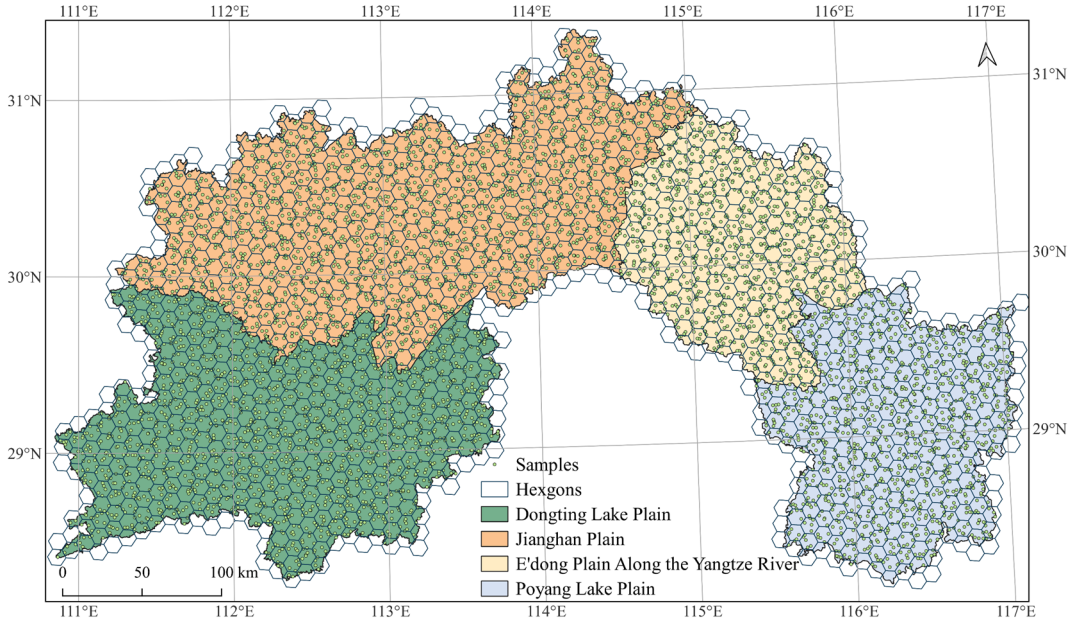

Figure 2. The distribution of sample points.

The accuracy was measured based on the sample and confusion matrix (Table 2) following previous methods [24]. The accuracy comparison metrics used included UA, PA, OA, and F1 score. UA, also known as precision, represents the ratio of correctly classified instances in the classification product and is calculated as 1 minus errors of commission. PA, also known as recall or sensitivity, is the ratio of correctly classified instances in the true class and is calculated as 1 minus errors of omission. OA is the ratio of correctly classified instances to all validation points. The F1 score is a composite metric of recall and precision. The specific formulas are as follows:

$$UA = \frac{TP}{TP + FP} \quad (1)$$

$$PA = \frac{TP}{TP + FN} \quad (2)$$

$$OA = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

$$F1 = \frac{2PA \times UA}{PA + UA} \quad (4)$$

where TP, TN, FP, and FN represent true positives, true negatives, false positives, and false negatives, respectively.

Table 2. Confusion matrix.

Predicted \ Labeled	True	False
True	TP	FP (error of commission)
False	FN (error of omission)	TN

The evaluation of consistency involves three main aspects: area comparison, per-pixel comparison, and comparison between products and statistical yearbooks [25,26]. Area comparison and per-pixel comparison primarily focus on comparative evaluations between the four products. Area comparison examines whether the proportions of cropland in

different products are consistent on the statistical grid, while per-pixel comparison assesses whether different products exhibit consistency at the pixel level. In area comparison, the products are typically gridded first, and then, the proportion of cropland is calculated at the grid scale. Per-pixel comparison involves direct spatial overlay analysis of the classified products to achieve every-pixel consistency. Additionally, we conducted a comparative evaluation between the products and statistical yearbook data.

3. Results

3.1. Accuracy Assessment

All four cropland distribution products in the study region had overall accuracies exceeding 85% (Table 3). GLAD, AGLC, the CLCD, and CACD achieved overall accuracies of 88.80%, 88.22%, 85.57%, and 88.60%, respectively. GLAD exhibited higher precision (i.e., lower errors of commission) at 96.09%, while the CLCD demonstrated higher recall (i.e., lower errors of omission) at 98.41%. This implies that GLAD has higher accuracy in determining cropland, while the CLCD provides a more comprehensive assessment of cropland. Although AGLC and CACD have some errors of commission and omission, they exhibit the highest F1 scores, indicating a balanced performance in precision and recall and higher comprehensive classification accuracy.

Table 3. Accuracy under different cropland identification scenarios.

Scenario	User's Accuracy (UA)	Producer's Accuracy (PA)	Overall Accuracy (OA)	F1 Score
GLAD	96.09%	83.95%	88.80%	89.61%
AGLC	85.81%	95.28%	88.22%	90.30%
CLCD	80.73%	98.41%	85.57%	88.69%
CACD	85.16%	97.10%	88.60%	90.74%
(1)	97.64%	81.63%	88.30%	88.92%
(2)	89.84%	95.57%	91.23%	92.62%
(3)	84.05%	98.29%	88.29%	90.61%
(4)	97.45%	81.92%	88.37%	89.01%
(5)	97.65%	81.72%	88.35%	88.98%
(6)	96.74%	83.49%	88.89%	89.63%
(7)	90.51%	93.34%	90.53%	91.90%

Combining multiple products for cropland identification enhances precision and overall accuracy. Depending on the specific requirements of precision prioritization, different cropland identification scenarios can be employed. When prioritizing precision, scenarios (1), (4), (5), and (6) achieved higher precision than GLAD. This indicates that cropland identification results in these scenarios are more accurate, with lower errors of commission. When prioritizing overall accuracy, scenarios (2), (6), and (7) achieved higher overall accuracy than GLAD. Prioritizing F1 score, scenarios (2) and (7) obtained higher F1 scores than CACD. This suggests that these scenarios exhibit better performance in balancing precision and recall, resulting in higher overall accuracy. Since AGLC lacks classification data for 2019, we utilized cropland identification scenarios (1) and (6) to obtain cropland distribution data with the highest precision for the years 2003, 2007, 2011, 2015, and 2019. In contrast, the CLCD provides cropland distribution data with the highest recall for these five years.

3.2. Consistency Evaluation

Consistency between products is closely related to the reference data used in the data generation process. The GLAD product only introduces GFSAD in its production process, resulting in lower consistency with the other three products (Figure 3). CACD introduces the CLCD in its production process, leading to the highest level of consistency between them. Both AGLC and CACD introduce GSW and GAIA in their production processes,

resulting in relatively high consistency between the two. Comparatively, GLAD exhibits higher consistency with CACD and AGLC, while its consistency with the CLCD is lower.

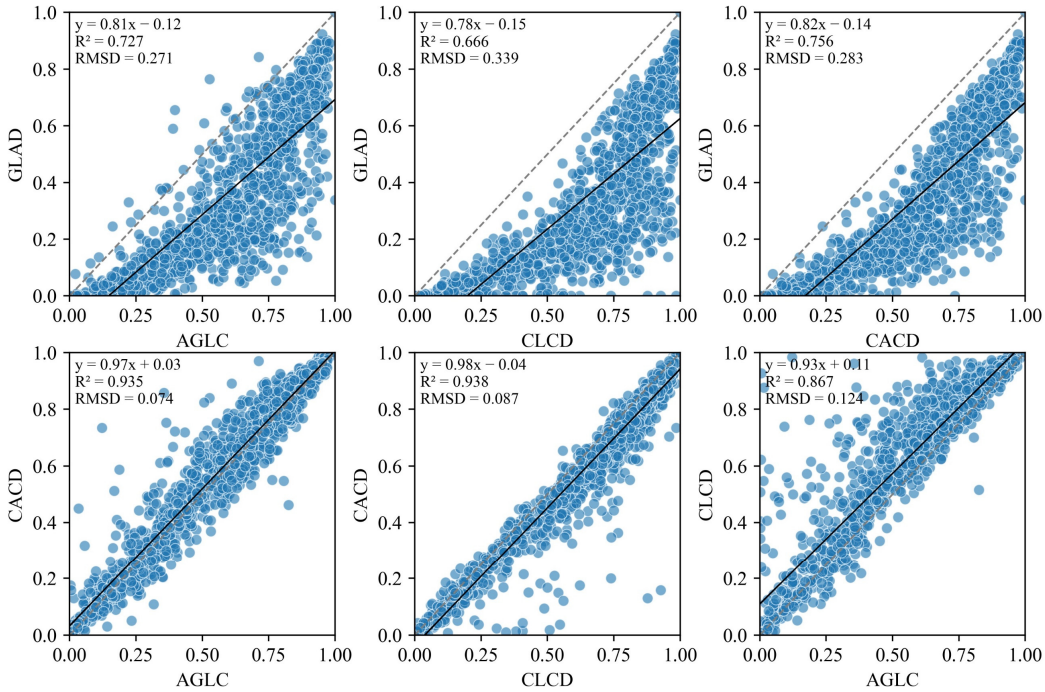

Figure 3. Consistency in cropland proportions across products (The blue points illustrate the percentage of cropland in small hexagons defined by the *x*-axis and *y*-axis. The black line represents the fitted regression line for these cropland percentage data points.).

Figure 3 illustrates the contrast in area consistency within the statistical grid among the four products. The *x* and *y* axes represent the proportion of cropland in hexagonal grids for individual products. AGLC, the CLCD, and CACD demonstrate high consistency, as they exhibit a higher degree of fit to the y = x line ($R^2 > 0.85$). The CLCD tends to estimate a larger cropland area compared to the other three products, as indicated by the higher number of data points between the CLCD *y*-axis and the y = x line. This aligns with CLCD's higher recall compared to the other three products. Similarly, GLAD estimates a smaller cropland area, consistent with its highest precision.

In terms of distribution consistency, the proportion of consistency among the four products in cropland classification reached 63.36%. The main inconsistency between the four products arose from GLADand the CLCD. Specifically, the consistency ratio between GLAD and the CLCD was 69.58% (Figure 4), while the consistency ratio among AGLC, the CLCD, and CACD was 81.92%. Among these three highly consistent products, AGLC and the CLCD exhibited relatively lower consistency at 85.68%. CACD showed higher consistency with AGLC and the CLCD, with agreement ratios of 88.26% and 89.90%, respectively.

Around 69.58% of GLAD and the CLCD were inconsistent. In regions with inconsistencies, the primary focus is on Wuhan City in Hubei, Yueyang City, and the northern part of Changde City in Hunan. Additionally, concentrated areas of inconsistency are found around Dongting Lake, Poyang Lake, Wuhan City, and the along the Jianghan Plain in eastern Hubei. Statistical data from county-level administrative regions reveal areas with notably high inconsistency between the two, including Yueyang County (1321.7 km^2), Taoyuan County (971.0 km^2), and Li County (825.5 km^2) in Hunan; Huangpi District (1260.4 km^2), Jiangxia District (905.0 km^2), and Xishui County (873.6 km^2) in Hubei; and Poyang County (1038.8 km^2) in Jiangxi.

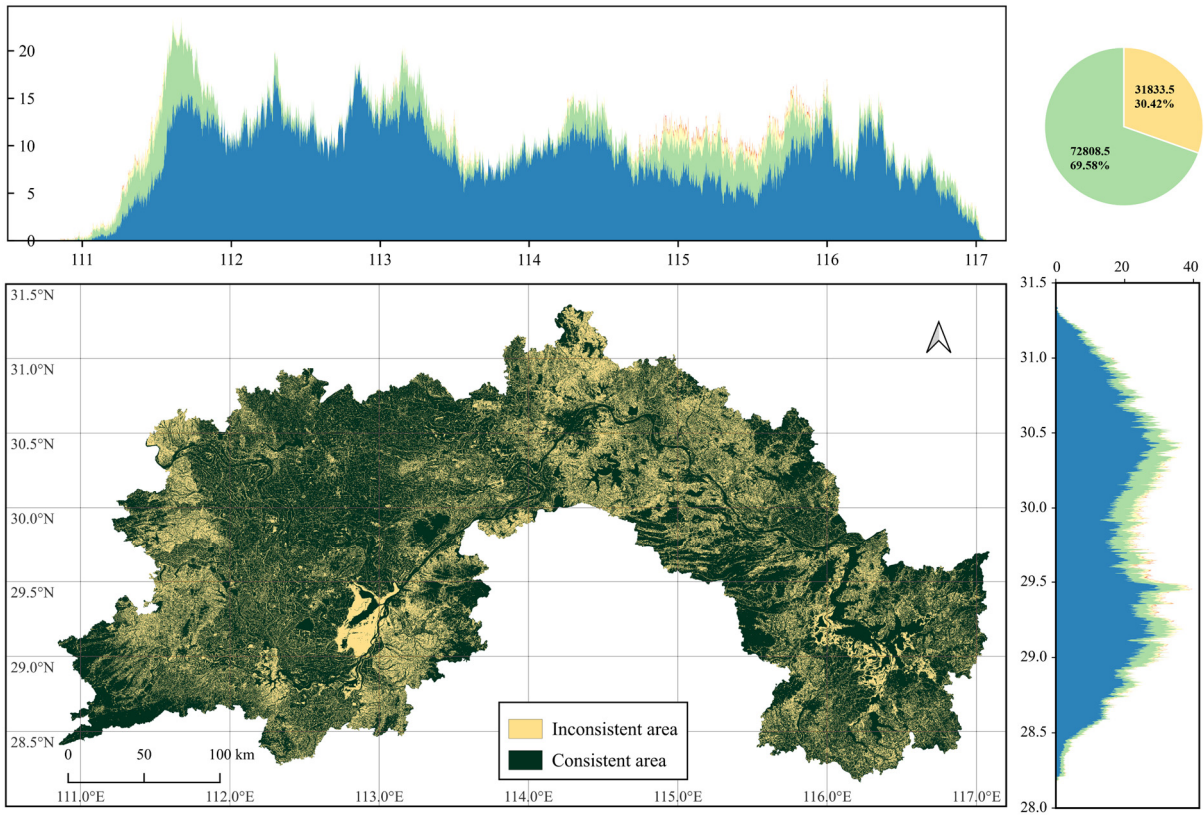

Figure 4. Consistency between GLAD and the CLCD (the distribution plot above and on the right illustrates the inconsistent area (in square kilometers) across longitude and latitude. Different colors represent varying slopes: blue corresponds to 0–2°, green to 2–6°, yellow to 6–15°, orange to 6–25°, and red to slopes greater than 25°).

To evaluate the cropland area, we compared two scenarios: scenario 1, which has the highest precision, and the CLCD, which has the highest recall. We compared these products separately against the cropland area from the statistical yearbook of 2011. The comparison was made in four prefectures (Wuhan, Huangshi, Ezhou, and Jingzhou) and 22 counties, which have the same statistical scale and available data for comparison.

In general, scenario 1 aligned more closely with the 2011 yearbook statistics, whereas the CLCD tended to provide higher estimations (Figure 5). The cropland areas for the four prefectures in 2011 were 2065.20 km^2, 1204.40 km^2, 413.90 km^2, and 6586.50 km^2, respectively. Scenario 1's estimated cropland areas closely matched these figures, measuring 2089.82 km^2, 592.49 km^2, 442.95 km^2, and 7734.35 km^2, with minimal differences except for Huangshi. However, the CLCD's estimates significantly surpassed the statistical data, measuring 5881.52 km^2, 2154.59 km^2, 1080.52 km^2, and 11,250.21 km^2, respectively. At the county level, scenario 1 occasionally overestimated cropland areas in certain regions, but overall, most areas aligned well with the yearbook statistics (scatter points close to the y = x line). Conversely, the CLCD consistently overestimated cropland areas at the county level, with scatter points consistently below the y = x line.

3.3. Cropland Distribution and Its Changes

In Scenario 1, we conducted a statistical analysis of cropland area changes at different boundary levels from 2003 to 2019. In 2019, the Jianghan Plain had the largest cropland area, covering 18,436.1 km^2, which accounted for 50.88% of the total cropland area. The Dongting Lake Plain, Poyang Lake Plain, and E'Dong Plain Along the Yangtze River accounted for 24.59%, 14.70%, and 9.82%, respectively. When examining cropland area statistics at the provincial and municipal levels for 2019, we found that Hubei, Hunan, and Jiangxi had cropland areas of 21,830.8 km^2, 8910.9 km^2, and 5492.2 km^2, respectively. Notably, Hubei

contributed more than half, reaching 60.25% of the total. Among prefectures, Jingzhou City had the largest cropland area, making up 21.29% of the total cropland.

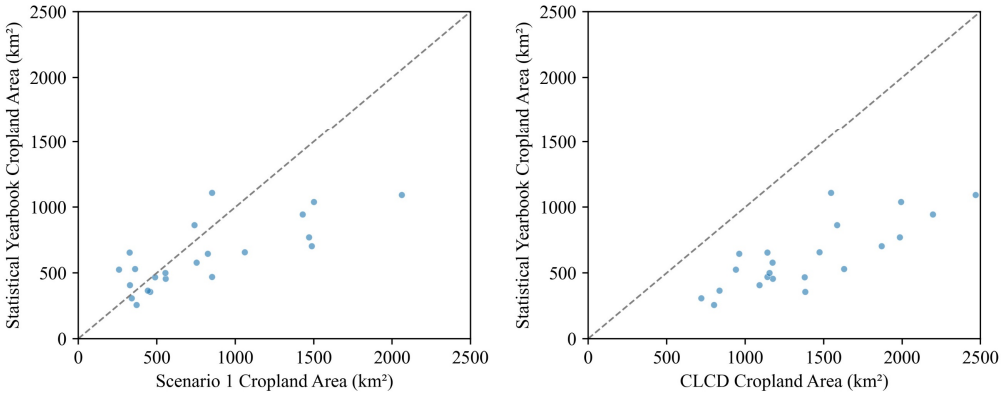

Figure 5. Comparison between different cropland scenarios and yearbook statistical data (The blue points represent the cropland area obtained by the *x*-axis and *y*-axis. The dotted line corresponds to the 1:1 relationship.).

From 2003 to 2019, the cropland area in the study region showed a fluctuating upward trend. Specifically, there was a noticeable increase in cropland area (548.6 km^2, 1.55%) from 2003 to 2007, followed by a slight decrease (−57.2 km^2, −0.16%) from 2007 to 2011. Subsequently, there was a modest increase in cropland area (3.6 km^2, 0.01%) from 2011 to 2015. Finally, a significant increase in cropland area (440.6 km^2, 1.12%) occurred from 2015 to 2019. Significant regional differences were observed in cropland changes. The Dongting Lake Plain and Poyang Lake Plain showed fluctuating increases in cropland area, with a rise of 616.21 km^2 (7.42%) and 374.34 km^2 (7.56%), respectively. However, the cropland area in the E'dong Plain Along the Yangtze River decreased by 46.00 km^2 (1.28%), and in the Jianghan Plain, it decreased by 8.96 km^2 (0.05%).

In 2019, the cropland area in the study region increased by 935.6 km^2 compared to 2003. Notable changes in cropland area were observed at the provincial level. Hunan Province and Jiangxi Province saw an increase in cropland area by 616.2 km^2 and 372.7 km^2, respectively, while Hubei Province experienced a decrease of 53.4 km^2. Within Hunan and Jiangxi, various cities witnessed an upward trend in cropland areas from 2003 to 2019 (Figure 6). Among them, Shangrao City in Jiangxi experienced the largest increase, adding 323.0 km^2, representing a 19.62% increase. In Hunan, Yueyang City (290.6 km^2, 11.29%), Yiyang City (159.9 km^2, 7.32%), and Changde City (165.7 km^2, 4.68%) also saw significant increases in cropland area. Certain areas in Hubei showed distinct patterns, with Jingzhou City having the most significant increase in cropland area, adding 238.7 km^2. Cities experiencing a decrease in cropland area included Wuhan City, Ezhou City, Xiaogan City, Huanggang City, Qianjiang City, and Xiaogan City. Wuhan City had the largest decrease in cropland area, reaching 185.8 km^2, followed by Qianjiang City (−94.3 km^2), Xiaogan City (−81.3 km^2), and Huanggang City (−70.2 km^2). At the county level, Poyang County and Yugan County in Shangrao City, Jiangxi, saw the most significant increases in cropland area, reaching 166.8 km^2 and 156.2 km^2, respectively. Additionally, Xiangyin County in Yueyang City, Yuanjiang City in Yiyang City, and Hanshou County in Changde City in Hunan experienced substantial increases in cropland area. In Hubei, Songzi City in Jingzhou City had the largest increase in cropland area at the county level, adding 68.8 km^2. Significant decreases in cropland area occurred in Qianjiang City, Jiangxia District in Wuhan City, and Hanchuan City in Xiaogan City.

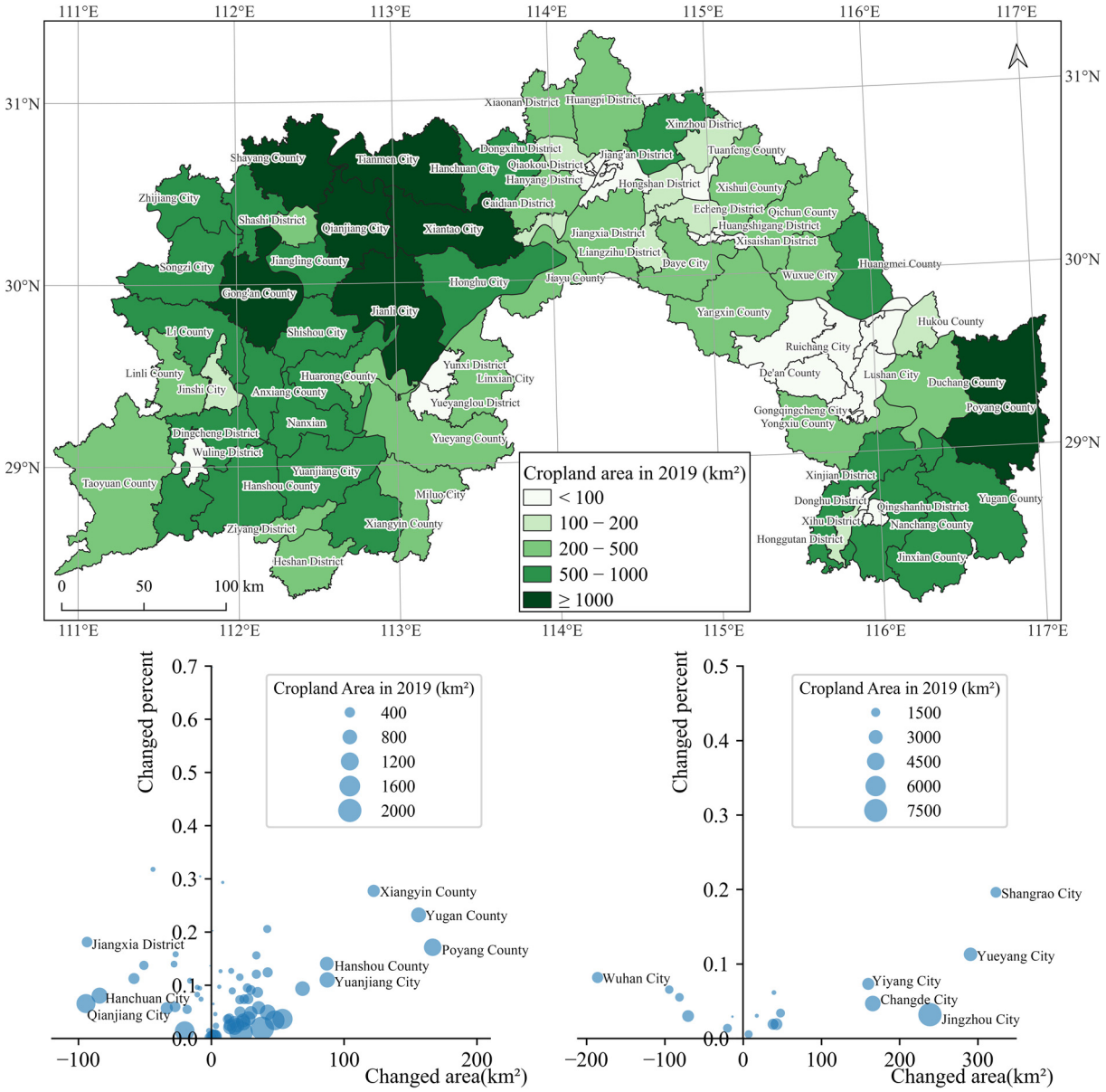

Figure 6. The changed area and proportion in counties and prefectures.

4. Discussion

The potential disparity in accuracy comparison results between GLAD and the CLCD may be attributed to differences in the definitions of cropland. GLAD categorizes cropland as areas dedicated to the cultivation of both annual and perennial herbaceous crops intended for human consumption, fodder (including hay), and biofuel production. This definition excludes lands utilized for cultivating perennial woody crops, permanent pastures, and areas left fallow for more than four years due to crop rotation. On the other hand, the CLCD classifies cropland as encompassing paddy rice fields, greenhouses, and other cultivated areas, along with orchards, managed grasslands, and temporarily bare farmland. Besides this, the complexity of the observed surface may affect the result as well. When the observed object is complicated, nonlinear interactions of light would happen, which would result in the spectral-mixing problem [27]. Different strategies used in these products could help mitigate the issue. GLAD chooses 924 representative grid cells for diverse agricultural settings, and CACD introduces the concept of agricultural zones.

Other factors, including the sample data source [28], predictor variables, and model algorithm, may influence the accuracy of the results. GLAD employs visual interpretation

to generate samples, demanding substantial human and time resources. In contrast, AGLC and CACD utilize algorithms for land use and land cover change detection in both sample and result migration, potentially affecting the ultimate accuracy. GLAD employs phenological indicators as predictors, while the CLCD integrates location features, both of which could impact the final outcome. Three model algorithms utilize random forest, with GLAD utilizing bagging decision trees.

Consistency was proven to be strongly correlated with factors such as topography and farming practices in previous studies [29,30]. In this study, there is a clear inconsistency as the CLCD designates certain beach areas around Dongting Lake as cropland, despite the predominant presence of poplar trees, reeds, cattails, and lake grass in these areas [31]. We also observed the phenomenon that the consistency of the products can be influenced by the reference data. GLAD exclusively utilized GFSAD as the reference data, presenting a relatively independent source compared to the other three datasets. As a result, the consistency between GLAD and the other three is comparatively lower. In contrast, AGLC and CACD both referenced GSW, and CACD referenced the CLCD and CLUDs. This direct alignment in data sources contributes to higher consistency among these three products, with the utmost consistency observed between CACD and the CLCD. The mixed pixels issue in fragmented landscapes [32] may result in the difference between the area derived from distribution products and yearbook statistics. For a more accurate cropland distribution, it is advised to utilize the intersection of all four products, emphasizing the highest precision. When dealing with marginal cropland regions, it is preferable to employ the scenario that yields the highest recall.

The dynamics of cropland are influenced by various factors aligned with stakeholder decisions. In response to the 1998 flood, the Grain for Green Project was initiated to address soil erosion and land degradation in China. The project aims to convert farmland into forests or grasslands to improve the ecological environment and promote sustainable development. As part of the project, approximately 447.48 km^2 of farmland in the Dongting Lake area was transformed into lakes. Additionally, in the northern and western parts of Dongting Lake, some cropland was repurposed for planting poplar trees [33]. The transformation of cropland into non-agricultural uses is significantly influenced by the agricultural labor force [34]. The swift progress in economic development and continuous urbanization has resulted in a substantial workforce migrating from rural regions to urban areas and industrial sectors [35]. In areas where croplands experience poor endowment conditions, leading to limited output and profits, abandonment is observed [36,37].

There are several limitations to this study, which can be explored in future research. The paper could include additional products for a more comprehensive assessment of accuracy and consistency, a task that could be undertaken in future studies. Google Earth provides free and easily accessible satellite imagery with global coverage, widely applied in land use and land cover classification [38]. However, its horizontal positional accuracy varies across regions [39]. Evidence suggests that in specific areas, it is sufficient for obtaining very accurate ground truth samples and measurements, suitable for producing large-scale planimetric maps [40]. Moreover, the time resolution of images on Google Earth is lower than that of satellite images, resulting in limited information about the temporal dynamics of cropland. Nonetheless, it can efficiently obtain relatively accurate samples overall.

5. Conclusions

All cropland distribution products have an overall accuracy of above 85% in the study region. GLAD has a high precision rate of 96.09%, indicating its accuracy in identifying croplands. The CLCD achieves a recall rate of 98.41%, indicating its thorough assessment of croplands. AGLC and CACD perform well in balancing precision and recall. To enhance precision and overall accuracy, the utilization of multiple products is recommended. Specifically, when all four products classify a point as cropland, higher precision is achieved. To enhance overall accuracy, it is recommended to select three or more products to classify a point as cropland or to directly use the joint assessment of AGLC, the CLCD, and CACD.

The four products have a consistent proportion of 63.36% in cropland classification in the study region. The primary inconsistency is observed between GLADand the CLCD. Specifically, the consistency between GLADand the CLCD is 69.58%, while the consistency between the CLCD and CACD is notably high at 89.90%. Since CACD incorporates the CLCD in its production process, the consistency between CACD and the CLCD is the highest. This suggests that the level of consistency between products is closely tied to the data used in the generation process. Areas of inconsistency are predominantly concentrated in Wuhan City, Hubei, Yueyang City in Hunan, and the northern part of Changde City. Additionally, inconsistent distribution is also concentrated around Dongting Lake, Poyang Lake, Wuhan City, and the E'dong Plain Along the Yangtze River.

In the study region, the largest cropland area is in Jianghan Plain, which is about half of the total cropland. Hubei has the most cropland, contributing 60.25% to the total. Among cities, Jingzhou City has the biggest cropland area, making up 21.29%. From 2003 to 2019, the overall cropland in the study region increased by 935.6 km^2. However, different regions experienced varied changes. The Dongting Lake Plain and Poyang Lake Plain saw increases, while the Jianghan Plain and E'dong Plain Along the Yangtze River had decreases. Notably, Poyang County and Yugan County in Jiangxi Province saw significant cropland increases, reaching 166.8 km^2 and 156.2 km^2. On the other hand, some cities, including Wuhan City, Qianjiang City, Xiaogan City, and Huanggang City, witnessed considerable reductions, with Wuhan City having the most significant decrease of 185.8 km^2.

Author Contributions: Conceptualization, H.X. and L.J.; methodology, H.X. and L.J.; formal analysis, H.X.; resources, H.X., L.J. and Y.L.; data curation, H.X., L.J. and Y.L.; writing of original draft, H.X., L.J. and Y.L.; writing, review, and editing, L.J.; supervision, H.X. and L.J.; project administration, L.J.; funding acquisition, L.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by a grant from National Natural Science Foundation of China (Grant Number: 42071253).

Data Availability Statement: Publicly available datasets were analyzed in this study. The cropland distribution data are available as follows: GLAD (https://glad.umd.edu/dataset/croplands (accessed on 25 February 2024)), AGLC (https://code.earthengine.google.com/?asset=users/xxc/GLC_2000_2015 (accessed on 25 February 2024)), the CLCD (https://zenodo.org/records/5816591 (accessed on 25 February 2024)), and CACD (https://zenodo.org/records/7936885 (accessed on 25 February 2024)). For regional crop statistics in Hubei, refer to the "Hubei Agricultural Statistical Yearbook" on the China National Knowledge Infrastructure (CNKI): http://data.cnki.net/ (accessed on 25 February 2024).

Acknowledgments: The authors are grateful to the reviewers for their comments and suggestions.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Lv, F.; Deng, L.; Zhang, Z.; Wang, Z.; Wu, Q.; Qiao, J. Multiscale Analysis of Factors Affecting Food Security in China, 1980–2017. *Environ. Sci. Pollut. Res.* **2022**, *29*, 6511–6525. [CrossRef] [PubMed]
2. Eigenbrod, F.; Beckmann, M.; Dunnett, S.; Graham, L.; Holland, R.A.; Meyfroidt, P.; Seppelt, R.; Song, X.-P.; Spake, R.; Václavík, T.; et al. Identifying Agricultural Frontiers for Modeling Global Cropland Expansion. *One Earth* **2020**, *3*, 504–514. [CrossRef] [PubMed]
3. Yu, P.; Liu, S.; Han, K.; Guan, S.; Zhou, D. Conversion of Cropland to Forage Land and Grassland Increases Soil Labile Carbon and Enzyme Activities in Northeastern China. *Agric. Ecosyst. Environ.* **2017**, *245*, 83–91. [CrossRef]
4. Cai, H.; Yang, X.; Xu, X. Spatiotemporal Patterns of Urban Encroachment on Cropland and Its Impacts on Potential Agricultural Productivity in China. *Remote Sens.* **2013**, *5*, 6443–6460. [CrossRef]
5. Ju, H.; Zhang, Z.; Zhao, X.; Wang, X.; Wu, W.; Yi, L.; Wen, Q.; Liu, F.; Xu, J.; Hu, S.; et al. The Changing Patterns of Cropland Conversion to Built-up Land in China from 1987 to 2010. *J. Geogr. Sci.* **2018**, *28*, 1595–1610. [CrossRef]
6. Yuan, Z.; Zhou, L.; Sun, D.; Hu, F. Impacts of Urban Expansion on the Loss and Fragmentation of Cropland in the Major Grain Production Areas of China. *Land* **2022**, *11*, 130. [CrossRef]
7. Song, W.; Pijanowski, B.C. The Effects of China's Cultivated Land Balance Program on Potential Land Productivity at a National Scale. *Appl. Geogr.* **2014**, *46*, 158–170. [CrossRef]
8. Xu, J.-Y.; Chen, L.-D.; Lu, Y.-H.; Fu, B.-J. Sustainability Evaluation of the Grain for Green Project: From Local People's Responses to Ecological Effectiveness in Wolong Nature Reserve. *Environ. Manag.* **2007**, *40*, 113–122. [CrossRef]

9. Potapov, P.; Turubanova, S.; Hansen, M.C.; Tyukavina, A.; Zalles, V.; Khan, A.; Song, X.-P.; Pickens, A.; Shen, Q.; Cortez, J. Global Maps of Cropland Extent and Change Show Accelerated Cropland Expansion in the Twenty-First Century. *Nat. Food* **2022**, *3*, 19–28. [CrossRef]
10. Jin, Y.; Liu, X.; Yao, J.; Zhang, X.; Zhang, H. Mapping the Annual Dynamics of Cultivated Land in Typical Area of the Middle-Lower Yangtze Plain Using Long Time-Series of Landsat Images Based on Google Earth Engine. *Int. J. Remote Sens.* **2020**, *41*, 1625–1644. [CrossRef]
11. Yao, X.; Wu, D. Spatiotemporal Changes and Influencing Factors of Rural Settlements in the Middle Reaches of the Yangtze River Region, 1990–2020. *Land* **2023**, *12*, 1741. [CrossRef]
12. Liu, J.; Kuang, W.; Zhang, Z.; Xu, X.; Qin, Y.; Ning, J.; Zhou, W.; Zhang, S.; Li, R.; Yan, C.; et al. Spatiotemporal Characteristics, Patterns, and Causes of Land-Use Changes in China since the Late 1980s. *J. Geogr. Sci.* **2014**, *24*, 195–210. [CrossRef]
13. Feng, Z. Agricultural Development Potential in the Lake Region of the Middle Yangtze Plain. *Resour. Environ. Yangtze Basin* **1994**, *2*, 114–120.
14. Wang, L.; Zhang, S.; Liu, Y.; Liu, Y. Interaction between Construction Land Expansion and Cropland Expansion and Its Socioeconomic Determinants: Evidence from Urban Agglomeration in the Middle Reaches of the Yangtze River, China. *Front. Environ. Sci.* **2022**, *10*, 882582. [CrossRef]
15. Xie, C.; Huang, X.; Mu, H.; Yin, W. Impacts of Land-Use Changes on the Lakes across the Yangtze Floodplain in China. *Environ. Sci. Technol.* **2017**, *51*, 3669–3677. [CrossRef] [PubMed]
16. Gong, P.; Liu, H.; Zhang, M.; Li, C.; Wang, J.; Huang, H.; Clinton, N.; Ji, L.; Li, W.; Bai, Y.; et al. Stable Classification with Limited Sample: Transferring a 30-m Resolution Sample Set Collected in 2015 to Mapping 10-m Resolution Global Land Cover in 2017. *Sci. Bull.* **2019**, *64*, 370–373. [CrossRef] [PubMed]
17. Karra, K.; Kontgis, C.; Statman-Weil, Z.; Mazzariello, J.C.; Mathis, M.; Brumby, S.P. Global Land Use/Land Cover with Sentinel 2 and Deep Learning. In Proceedings of the 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, Brussels, Belgium, 11–16 July 2021; pp. 4704–4707.
18. Zanaga, D.; Van De Kerchove, R.; Daems, D.; De Keersmaecker, W.; Brockmann, C.; Kirches, G.; Wevers, J.; Cartus, O.; Santoro, M.; Fritz, S. ESA WorldCover 10 m 2021 V200. 2022. Available online: https://zenodo.org/records/7254221 (accessed on 27 February 2024). [CrossRef]
19. Brown, C.F.; Brumby, S.P.; Guzder-Williams, B.; Birch, T.; Hyde, S.B.; Mazzariello, J.; Czerwinski, W.; Pasquarella, V.J.; Haertel, R.; Ilyushchenko, S.; et al. Dynamic World, Near Real-Time Global 10 m Land Use Land Cover Mapping. *Sci. Data* **2022**, *9*, 251. [CrossRef]
20. Tu, Y.; Wu, S.; Chen, B.; Weng, Q.; Gong, P.; Bai, Y.; Yang, J.; Yu, L.; Xu, B. A 30 m Annual Cropland Dataset of China from 1986 to 2021. *Earth Syst. Sci. Data Discuss.* **2023**, 1–34, *[preprint], in review*. [CrossRef]
21. Xu, X.; Li, B.; Liu, X.; Li, X.; Shi, Q. Mapping Annual Global Land Cover Changes at a 30 m Resolution from 2000 to 2015. *Yaogan Xuebao/J. Remote Sens.* **2021**, *25*, 1896–1916. [CrossRef]
22. 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]
23. China National Knowledge Infrastructure (CNKI). Available online: https://www.cnki.net/index/ (accessed on 14 February 2024).
24. Stehman, S.V. Estimating Area and Map Accuracy for Stratified Random Sampling When the Strata Are Different from the Map Classes. *Int. J. Remote Sens.* **2014**, *35*, 4923–4939. [CrossRef]
25. Chen, Y.; Shao, H.; Li, Y. Consistency Analysis and Accuracy Asssessment of Multi-Source Land Cover Products in the Yangyze River Delta. *Trans. Chin. Soc. Agric. Eng.* **2021**, *37*, 142–150.
26. Liu, L.; Zhang, X.; Gao, Y.; Chen, X.; Shuai, X.; Mi, J. Finer-Resolution Mapping of Global Land Cover: Recent Developments, Consistency Analysis, and Prospects. *J. Remote Sens.* **2021**, *2021*, 5289697. [CrossRef]
27. Yu, L.; Wang, J.; Gong, P. Improving 30 m Global Land-Cover Map FROM-GLC with Time Series MODIS and Auxiliary Data Sets: A Segmentation-Based Approach. *Int. J. Remote Sens.* **2013**, *34*, 5851–5867. [CrossRef]
28. Foody, G.M.; Arora, M.K. An Evaluation of Some Factors Affecting the Accuracy of Classification by an Artificial Neural Network. *Int. J. Remote Sens.* **1997**, *18*, 799–810. [CrossRef]
29. Xue, J.; Zhang, X.; Chen, S.; Hu, B.; Wang, N.; Shi, Z. Quantifying the Agreement and Accuracy Characteristics of Four Satellite-Based LULC Products for Cropland Classification in China. *J. Integr. Agric.* **2023**, *23*, 283–297. [CrossRef]
30. Zhang, C.; Dong, J.; Ge, Q. Quantifying the Accuracies of Six 30-m Cropland Datasets over China: A Comparison and Evaluation Analysis. *Comput. Electron. Agric.* **2022**, *197*, 106946. [CrossRef]
31. Yu, S.; Li, C.; Yu, D.; He, Q.; Luo, W.; Xiang, F. Land Cover Change on Beach of Dongting Lake's Beach. *Earth Sci.* **2020**, *45*, 1918–1927. [CrossRef]
32. Zhao, X.; Wu, T.; Wang, S.; Liu, K.; Yang, J. Cropland Abandonment Mapping at Sub-Pixel Scales Using Crop Phenological Information and MODIS Time-Series Images. *Comput. Electron. Agric.* **2023**, *208*, 107763. [CrossRef]
33. Zhu, L.; Liu, X.; Wu, L.; Tang, Y.; Meng, Y. Long-Term Monitoring of Cropland Change near Dongting Lake, China, Using the LandTrendr Algorithm with Landsat Imagery. *Remote Sens.* **2019**, *11*, 1234. [CrossRef]
34. Zhang, J.; Liu, Y.; Zhang, E.; Chen, J.; Tan, Q. Dynamics and Driving Mechanisms of Cultivated Land at County Level in China. *Acta Geogr. Sin.* **2023**, *78*, 2105–2127.

35. Li, D.; Duo, L.; Bao, C.; Zhang, X.; Zou, Z. Spatiotemporal Distribution and Fragmentation Driving Mechanism in Paddy Fields and Dryland of Urban Agglomeration in the Middle Reaches of the Yangtze River. *Land* **2024**, *13*, 58. [CrossRef]
36. Li, S.; Li, X. Economic Characteristics and the Mechanism of Farmland Marginalization in Mountainous Areas of China. *Acta Geogr. Sin.* **2018**, *73*, 803–817.
37. Zhong, H.; Liu, Z.; Wang, J. Understanding Impacts of Cropland Pattern Dynamics on Grain Production in China: An Integrated Analysis by Fusing Statistical Data and Satellite-Observed Data. *J. Environ. Manag.* **2022**, *313*, 114988. [CrossRef] [PubMed]
38. Yu, L.; Gong, P. Google Earth as a Virtual Globe Tool for Earth Science Applications at the Global Scale: Progress and Perspectives. *Int. J. Remote Sens.* **2012**, *33*, 3966–3986. [CrossRef]
39. Potere, D. Horizontal Positional Accuracy of Google Earth's High-Resolution Imagery Archive. *Sensors* **2008**, *8*, 7973–7981. [CrossRef] [PubMed]
40. Pulighe, G.; Baiocchi, V.; Lupia, F. Horizontal Accuracy Assessment of Very High Resolution Google Earth Images in the City of Rome, Italy. *Int. J. Digit. Earth* **2016**, *9*, 342–362. [CrossRef]