Article

Quality Evaluation of Multi-Source Cropland Data in Alpine Agricultural Areas of the Qinghai-Tibet Plateau

Shenghui Lv [1,2,3,†], Xingsheng Xia [1,2,*,†], Qiong Chen [1,2] and Yaozhong Pan [1,3]

1   Academy of Plateau Science and Sustainability, Qinghai Normal University, Xining 810016, China
2   School of Geographical Sciences, Qinghai Normal University, Xining 810016, China
3   State Key Laboratory of Remote Sensing Science, Beijing Normal University, Beijing 100875, China
*   Correspondence: xxs@qhnu.edu.cn
†   These authors contributed equally to this work.

Abstract: Accurate cropland distribution data are essential for efficiently planning production layouts, optimizing farmland use, and improving crop planting efficiency and yield. Although reliable cropland data are crucial for supporting modern regional agricultural monitoring and management, cropland data extracted directly from existing global land use/cover products present uncertainties in local regions. This study evaluated the area consistency, spatial pattern overlap, and positional accuracy of cropland distribution data from six high-resolution land use/cover products from approximately 2020 in the alpine agricultural regions of the Hehuang Valley and middle basin of the Yarlung Zangbo River (YZR) and its tributaries (Lhasa and Nianchu Rivers) area on the Qinghai-Tibet Plateau. The results indicated that (1) in terms of area consistency analysis, European Space Agency (ESA) WorldCover cropland distribution data exhibited the best performance among the 10 m resolution products, while GlobeLand30 cropland distribution data performed the best among the 30 m resolution products, despite a significant overestimation of the cropland area. (2) In terms of spatial pattern overlap analysis, AI Earth 10-Meter Land Cover Classification Dataset (AIEC) cropland distribution data performed the best among the 10 m resolution products, followed closely by ESA WorldCover, while the China Land Cover Dataset (CLCD) performed the best for the Hehuang Valley and GlobeLand30 performed the best for the YZR area among the 30 m resolution products. (3) In terms of positional accuracy analysis, the ESA WorldCover cropland distribution data performed the best among the 10 m resolution products, while GlobeLand30 data performed the best among the 30 m resolution products. Considering the area consistency, spatial pattern overlap, and positional accuracy, GlobeLand30 and ESA WorldCover cropland distribution data performed best at 30 m and 10 m resolutions, respectively. These findings provide a valuable reference for selecting cropland products and can promote refined cropland mapping of the Hehuang Valley and YZR area.

Keywords: land use/cover products; cropland; consistency analysis; accuracy validation; influencing factors; Qinghai-Tibet Plateau

1. Introduction

Accurate cropland distribution data form the cornerstone of modern agricultural production management [1]. Currently, these data are extracted primarily from existing land use/cover products, such as WorldCover (WC) by the European Space Agency (ESA) [2], Sentinel-2 10-Meter Land Use/Land Cover (LC) by Esri [3], the AI Earth 10-Meter Land Cover Classification Dataset (AIEC) by the AI Earth team of DAMO Academy [4], GlobeLand30 (GL) by the National Geomatics Center of China [5], GLC_FCS30 (GLC) by the Chinese Academy of Sciences [6], and the China Land Cover Dataset (CLCD) by Wuhan University [7]. Although these products have become vital data sources for acquiring fundamental information on global or regional cropland distribution, they are generated using diverse classification standards, data processing methods, training samples, and

classification techniques. Moreover, inherent uncertainties in the application of Earth observation technology result in discrepancies in cropland distribution data extracted from various products, which introduce uncertainty in practical research and applications [8–12]. Therefore, whether existing cropland distribution data in specific areas can meet user research needs must be determined, and the criteria for selecting appropriate cropland distribution data must be established based on these needs.

Thus, a number of researchers have conducted quality evaluations of various cropland distribution datasets across different study areas to provide references for data application selection. For instance, Xue et al. [13] and Zhang et al. [14] performed comparative evaluations of cropland distribution datasets from multiple products covering mainland China for 2010 and 2015, respectively, and revealed varying degrees of discrepancy between the different products in terms of total area and spatial distribution consistency. Moreover, global accuracy evaluation results do not necessarily reflect local evaluation results, with plain areas exhibiting better area and spatial consistency between different products than mountainous regions, which present significant terrain variations. The size of the plots significantly influences the cropland data quality as well. Therefore, the quality of large-scale data products for local applications must be further evaluated to ensure the accuracy of basic data in local research or applications. Such evaluations will also aid in analyzing and discussing uncertainties in research or applications.

Alpine agricultural areas, which are the major grain supply areas in cold and arid regions, are responsible for ensuring regional grain self-sufficiency. However, owing to environmental constraints in cold and arid regions, most alpine agriculture areas are concentrated in valleys with steep slopes and deep gullies [15], leading to cropland resources characterized by small plot areas and fragmented distributions. Thus, accurate and reliable cropland distribution data are particularly crucial for supporting the modernization of alpine agricultural management and the continuous optimization of the human–land relationship. In practical research and applications, directly extracting cropland data from existing datasets is undoubtedly cost-effective. However, most existing datasets [2–7] were developed and produced on national or global scales. Although studies [9,10,13,14,16] have conducted evaluations at different scales and provided application recommendations, they have rarely focused on cold and arid regions. Moreover, single-element spatial distribution quality evaluations of croplands are rare. Consequently, clear guidance on data selection for practical applications is lacking.

This study aimed to further evaluate the consistency and accuracy of cropland distribution data within current medium- and high-resolution land use/cover datasets in two major alpine agricultural regions of the Qinghai-Tibet Plateau: Hehuang Valley in Qinghai, China, and the middle basin of the Yarlung Zangbo River and its tributaries (Lhasa and Nianchu Rivers; YZR area) in Tibet, China. The objective was to elucidate the quality of cropland distribution data in large-scale land use/cover products in the alpine agricultural regions of the Qinghai-Tibet Plateau, thereby providing a reference for data selection for research and application.

2. Materials and Methods

2.1. Study Area

The Hehuang Valley (Figure 1a,c) is located in the northeastern part of the Qinghai-Tibet Plateau and encompasses a total area of approximately 26,000 km^2. It consists of a section of the Yellow River Valley approximately 200 km long and several valleys of the Huangshui River system. The region is predominantly characterized by high or very high mountains, alluvial plains, hills, and terraces, and it features numerous wide valleys formed by river erosion.
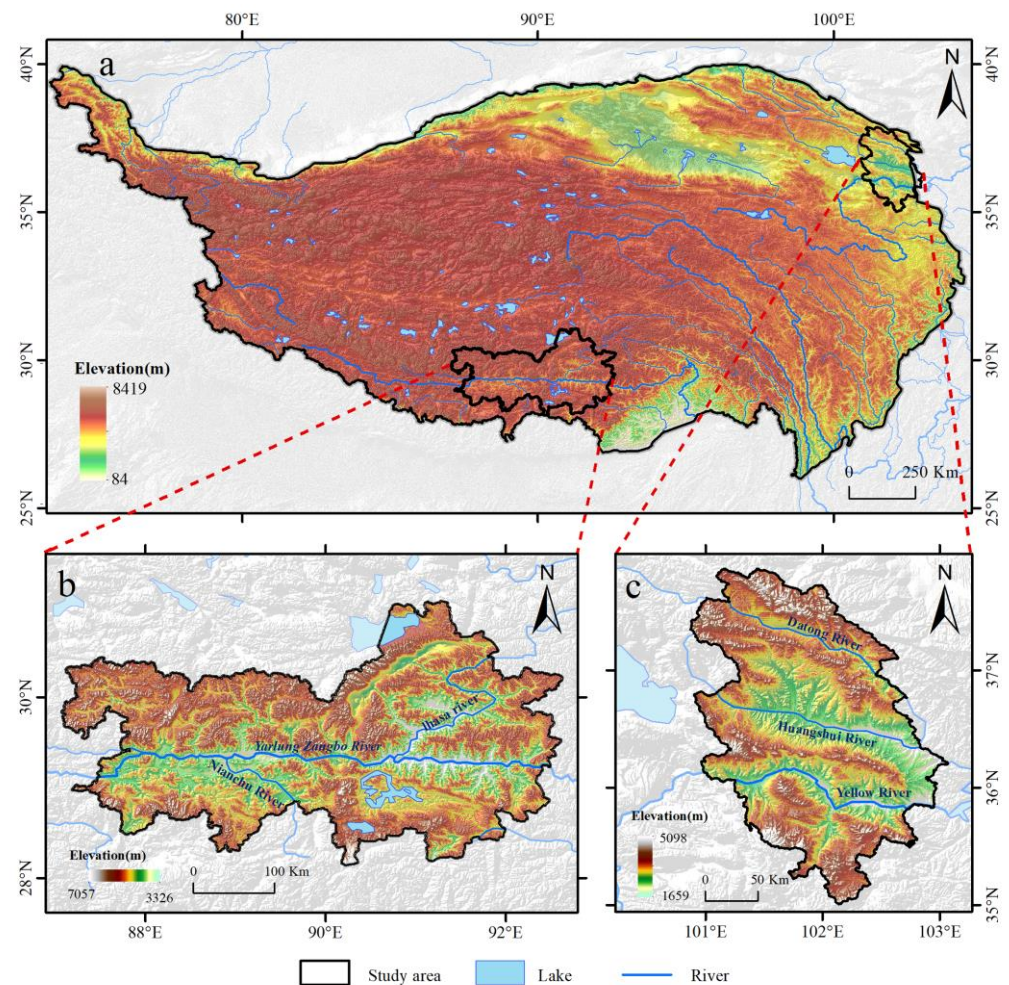
Figure 1. Overview of the study area: (**a**) study area location, (**b**) YZR area, and (**c**) Hehuang Valley. Note: elevation and slope may also vary on Earth due to its geological activity.

The YZR area (Figure 1a,b) is situated in the southern part of the Qinghai-Tibet Plateau. It primarily encompasses the middle reaches of the Yarlung Zangbo, Lhasa, and Nianchu Rivers, with a total area of approximately 66,500 km². This region belongs to the mountain plain broad-valley area, which is characterized by wide and gentle river floodplains, alluvial terraces, and alluvial fans along riverbanks, as well as a few gorges.

Owing to the valley terrain, both the Hehuang Valley and the YZR area have relatively low altitudes that meet the hydrothermal conditions required for alpine agriculture. These regions have become fertile lands with concentrated populations on the Qinghai-Tibet Plateau, with cropland in these two regions accounting for more than 60% of the total cropland area in the province/region. Agricultural production in these regions supports approximately two-thirds of the population of Qinghai and Tibet. However, constraints associated with the rugged valley terrain, which presents steep slopes and deep gullies [15], have led to a fragmented and scattered distribution of cropland resources in both regions. Therefore, accurate cropland distribution data are crucial for supporting the agricultural development and regulation in these two major alpine agricultural areas to ensure grain self-sufficiency and food security.

2.2. Data and Preprocessing

The data utilized in this study primarily consisted of land use/cover products, remote sensing imagery, basic geographical data, and cropland statistics.

The land use/cover products were produced by various institutions based on 2020 as the benchmark: WC [2], LC [3], AIEC [4], GL [5], GLC [6], and CLCD [7] (all released

by 2020). These products were primarily employed to extract cropland distribution data from the study area at varying spatial resolutions. These datasets were primarily derived from Landsat and Sentinel-2 series satellite data. Specifically, the GL, GLC, and CLCD datasets, which are based on Landsat series satellite data, have a spatial resolution of 30 m, whereas the WC, LC, and AIEC datasets, which are based on Sentinel-2 series satellite data, have a spatial resolution of 10 m. Detailed information regarding the releasing institutions, classification methods, and product accuracy is presented in Table 1. Data preprocessing primarily involved downloading and mosaicking data tiles based on the study area, extracting cropland distribution data according to land type codes, and unifying spatial references.

Table 1. Land use/cover products.

Land Use/Cover Product	Major Research and Development Unit	Satellite Sensors	Extraction Method	Overall Accuracy	Cropland Code	Spatial Resolution
WC [2,17–19]	European Space Agency	Sentinel-2	Decision tree classification	74.40%	40	10 m
LC [3,20]	ESRI	Sentinel-2	Deep learning classification	75.00%	5	10 m
AIEC [4]	DAMO Academy	Sentinel-2	Deep learning classification	—	1	10 m
GL [5]	National Geomatics Center of China	Landsat, HJ-1	Maximum likelihood classification	85.72%	10	30 m
GLC [6]	Aerospace Information Research Institute, Chinese Academy of Sciences	Landsat	Random forest classification	—	10, 20	30 m
CLCD [7]	Wuhan University	Landsat	Random forest classification	79.31%	1	30 m

Note: "—" indicates that the publication accuracy was not clearly stated, or the data source was not found. WC, WorldCover; LC, Sentinel-2 10-Meter Land Use/Land Cover; AIEC, AI Earth 10-Meter Land Cover Classification Dataset; GL, GlobeLand30; GLC, GLC_FCS30; CLDC, China Land Cover Dataset.

Remote sensing imagery primarily consisted of Landsat 8, Sentinel-2A/B, and Google Earth Engine (GEE) imagery from approximately 2020, with a focus on the growing season (May–September). As this study evaluated the quality of historical data products, these imagery datasets were primarily utilized to visually interpret and obtain cropland and non-cropland samples to assess the accuracy of the cropland data products. Specifically, grid center points within the study area were employed as the sampling population, and a 10% random sample was drawn to obtain validation sample points, which were visually interpreted using multisource remote sensing imagery. The results are shown in Figure 2. In the Hehuang Valley, 1639 sample points were obtained, and they consisted of 349 cropland samples and 1290 non-cropland samples. In the YZR area, 1468 sample points were obtained, and they consisted of 151 cropland samples and 1317 non-cropland samples. Additionally, to compare the spatial performance of cropland distribution data during the same period, 0.75 m resolution Jilin-1 optical satellite imagery was utilized as a reference, specifically for the Hehuang Valley on 7 October 2020, and the YZR area on 17 October 2020.

Basic geographical data consisted of topographic and administrative information. Topographic data, primarily elevation and slope, were used to analyze the relationship between cropland distribution and topographic factors in the study area. Elevation data were directly obtained from the 30 m resolution NASADEM [21] digital elevation model (DEM), and slope data were derived from the DEM. Administrative data encompassing the county-level administrative boundaries of the study area were utilized to define the

study area boundaries, and they were sourced from the National Catalogue Service for Geographical Information of China [22].
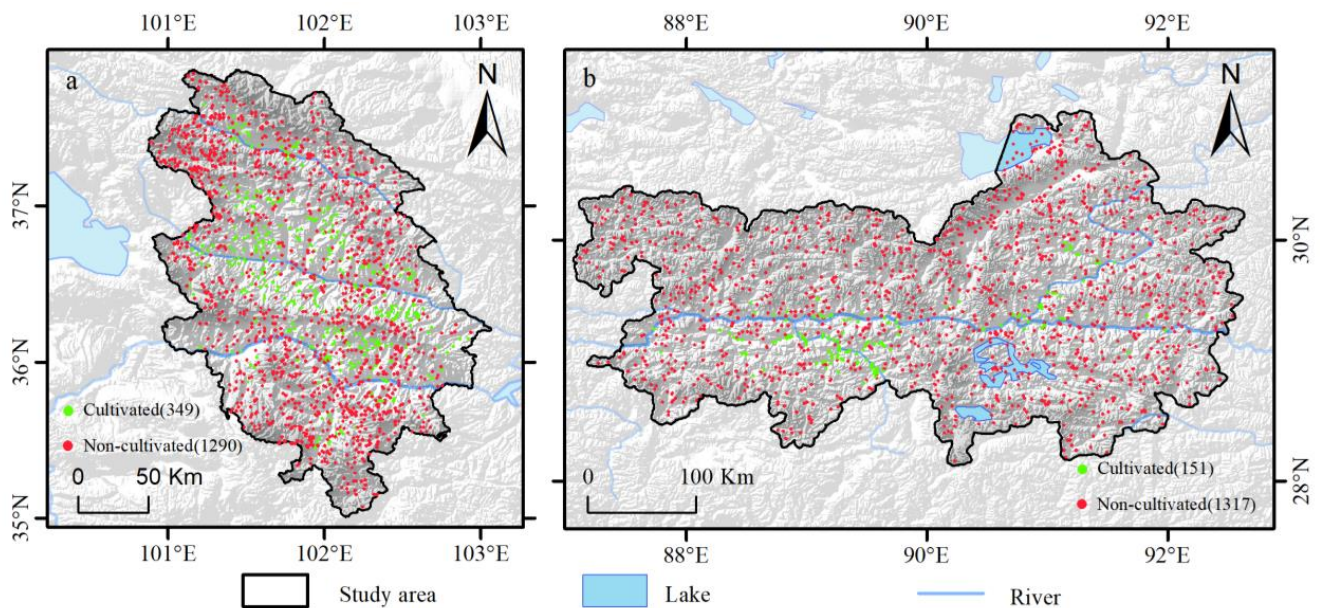

Figure 2. Verification sample points: (**a**) sample points in the Hehuang Valley and (**b**) sample points in the YZR area.

Cropland statistics specifically denote the cropland area or grain crop planting area of each county and were primarily employed to evaluate the area consistency of each cropland distribution dataset by comparing their degree of fitness and deviation. These data primarily originated from the agricultural or statistical departments of local governments.

2.3. Methods

2.3.1. Area Consistency

From a practical application standpoint, the agricultural sector is primarily concerned with the total amount of cropland resources, followed by the spatial distribution, quality, and production capacity of these resources. Consequently, the relative accuracy of an area is often the primary concern of users, and it is typically assessed by comparing different product data with reference data in terms of fitness and deviations.

Fitness evaluations are primarily performed via linear regression comparisons using the coefficient of determination (R^2; Equation (1)). R^2 values range from 0 to 1 and indicate the degree of fit between observed and predicted values. In this study, R^2 values closer to 1 indicated a greater fit between the cropland product data and reference data. The deviation of cropland product data from the reference data was primarily assessed using the root mean-square error (*RMSE*; Equation (2)), which measures the difference between the observed and predicted values, with a smaller *RMSE* indicating greater data accuracy. In this study, a smaller *RMSE* indicated a lower deviation between the cropland product data and reference data.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(x_i - y_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2} \tag{1}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(x_i - y_i)^2}{n}} \tag{2}$$

where x_i is the cropland area of the ith county in the various cropland data, y_i is the cropland area in the corresponding statistical data of that county, and $\overline{y}$ is the average cropland area in the statistical data.

2.3.2. Spatial Pattern Overlap Analysis

Area consistency evaluation can indicate the similarity of the total cropland area but cannot determine the consistency of different cropland distribution data in spatial terms. Spatial certainty is crucial for decision-making in research and applications. Therefore, this study employed a spatial overlay to obtain the spatial correspondence of different data on a per-pixel basis, subsequently assessing the per-pixel consistency. Thematic maps were generated to support the reliability assessment of the classification results. The spatial overlay results were categorized into high-consistency pixels, medium-consistency pixels, and low-consistency pixels. Specifically, high-consistency pixels indicated that all three products identified them as cropland, medium-consistency pixels indicated that two products classified them as cropland, and low-consistency pixels indicated that only one product identified them as cropland.

2.3.3. Positional Accuracy

The accuracy of classification categories is a crucial foundation for precision oversight in modern agriculture. In this context, this study primarily used samples to construct a confusion matrix for evaluation, which is also referred to as a sample accuracy evaluation. Constructing a confusion matrix from samples is a widely used method for accuracy assessment in remote sensing classification [23,24]. Specific indicators included accuracy (*ACC*), *precision*, Matthew's correlation coefficient (*MCC*), true positive rate (*TPR*), false positive rate (*FPR*), and comprehensive evaluation index (*CEI*).

ACC: *ACC* is a fundamental indicator that measures the performance of a classification model and represents the proportion of correctly classified samples compared to the total number of samples (Equation (3)). In this study, the higher *ACC* indicated that various data products successfully predicted most cropland classes during extraction:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

where *TP* represents true positive, which is the number of correctly classified positive samples, *FN* represents false negative, which is the number of incorrectly classified positive samples, *FP* represents false positive, which is the number of negative samples misclassified as positive, and *TN* represents true negative, which is the number of correctly classified negative samples. In this study, positive and negative samples refer to the cropland and non-cropland samples, respectively.

Precision: *Precision* measures the proportion of samples predicted as positive that are truly positive among all samples predicted as positive (Equation (4)). In this study, higher *precision* signified fewer instances of other land types being misclassified as cropland during extraction:

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

MCC: *MCC* is a comprehensive measure of a classification model's performance that considers the relationship between true positives, true negatives, false positives, and false negatives (Equation (5)). Its value ranges from −1 to 1, where 1 denotes perfect prediction, 0 represents random prediction, and −1 signifies completely inconsistent prediction:

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP) \times (TP + FN) \times (TN + TP) \times (TN + FN)}} \tag{5}$$

TPR: *TPR*, also known as recall, measures the proportion of actual positives correctly identified as positives among all actual positives (Equation (6)). In this study, a higher *TPR* reflects the ability of the various datasets to identify a greater number of cropland patches during extraction:

$$TPR = \frac{TP}{TP + FN} \tag{6}$$

FPR: *FPR* measures the proportion of actual negatives incorrectly classified as positive among all actual negatives (Equation (7)). A lower *FPR* signifies that various datasets are more effective at reducing the misclassification of other land types, such as cropland, during extraction:

$$FPR = \frac{FP}{TN + FP} \tag{7}$$

An excessive number of evaluation indicators can result in information overload, thereby increasing the difficulty of distilling key information and drawing conclusions from the data. Therefore, based on the aforementioned evaluation indicators, this study introduced a new *CEI* to assess the accuracy of the cropland extraction results from various distribution datasets. Specifically, the algorithm first organizes all evaluation indicators in descending order (with NaN values replaced by 0), and then assigns scores ranging from 6 to 1 to the evaluation results of various cropland datasets (with *FPR* scores ranging from −6 to −1). The sum of all assigned scores for each cropland dataset yields a *CEI*. A higher *CEI* indicates superior cropland distribution data extraction.

Additionally, according to the results of the spatial pattern analysis, data products with the same spatial resolution should ideally exhibit consistent spatial distribution results under the same natural environmental conditions. However, variations in the data, samples, and methodologies may lead to discrepancies in the spatial distribution of different products. High consistency suggests that data products from varying sources, samples, and methods maintain a uniform data quality within specific regions, whereas medium or low consistency indicates the products present variations in data quality under identical natural environmental conditions, potentially indicating the products that offer superior quality. Therefore, based on the sample accuracy evaluation, this study not only investigated the overall accuracy of cropland data categories at the same spatial resolution but concurrently assessed the category accuracy in pixel regions with medium and low consistency, as indicated by the pattern analysis results.

3. Results

3.1. Area Consistency Results

Figures 3 and 4 display the cropland area statistics derived from various cropland distribution datasets for Hehuang Valley and the YZR area. Theoretically, if all cropland distribution datasets are reliable, then the area statistics across these regions should yield consistent results. However, as illustrated in Figures 3 and 4, notable variations occurred in the cropland area statistics among the different datasets. Specifically, cropland area statistics from the 10 m resolution datasets for both the Hehuang Valley and the YZR area were relatively consistent, whereas the 30 m resolution datasets showed considerable discrepancies. Notably, cropland areas from the GL and GLC products showed substantially higher consistency than those from other datasets, while CLCD demonstrated relatively consistent results with the 10 m resolution data in the Hehuang Valley but showed markedly lower consistency in the YZR area. This indicates that the 10 m resolution datasets may offer more reliable results for both agricultural regions. Among the 30 m resolution datasets, only CLCD provided reliability comparable to the 10 m resolution datasets in the Hehuang Valley area, while its reliability in the YZR area may be poorer. Additionally, the GL and GLC datasets may demonstrate lower reliability across both agricultural regions.

Using data from agricultural management and statistical departments, this study assessed the relative discrepancies between cropland distribution data and county-level statistical data. The results revealed that in the Hehuang Valley, 10 m resolution cropland distribution data generally overestimated cropland areas, with notable overestimations in the Chengbei District, Chengxi District, and Datong Hui and Tu Autonomous County. Conversely, Huangyuan County and Ping'an District showed a tendency to underestimate cropland areas. In contrast, 30 m resolution cropland distribution data products generally overestimated cropland areas more severely than their 10 m resolution counterparts, with the Ping'an District showing a relatively lower degree of overestimation. In the YZR area,

although comprehensive statistical data references were lacking, analysis based on the available data indicated that all three 10 m resolution cropland distribution datasets tended to underestimate cropland areas. For the 30 m resolution data, while the GL, GLC, and CLCD products overestimated cropland areas in some counties, they severely underestimated cropland areas in others. These products extracted significantly less cropland area from the YZR area compared to the other cropland distribution datasets.

Additionally, the correlation between various cropland distribution data and statistical data was analyzed (Table 2). In the Hehuang Valley, for 10 m resolution cropland distribution data, the R^2 value for WC surpassed that of LC and AIEC. Similarly, the *RMSE* for WC was lower than that for AIEC and LC, although the differences were minimal. For the 30 m resolution cropland distribution data, GL exhibited the highest R^2 value, followed by GLC, while CLCD had the lowest R^2 value. Moreover, the *RMSE* of CLCD data was lower than that of the GLC and GL data.
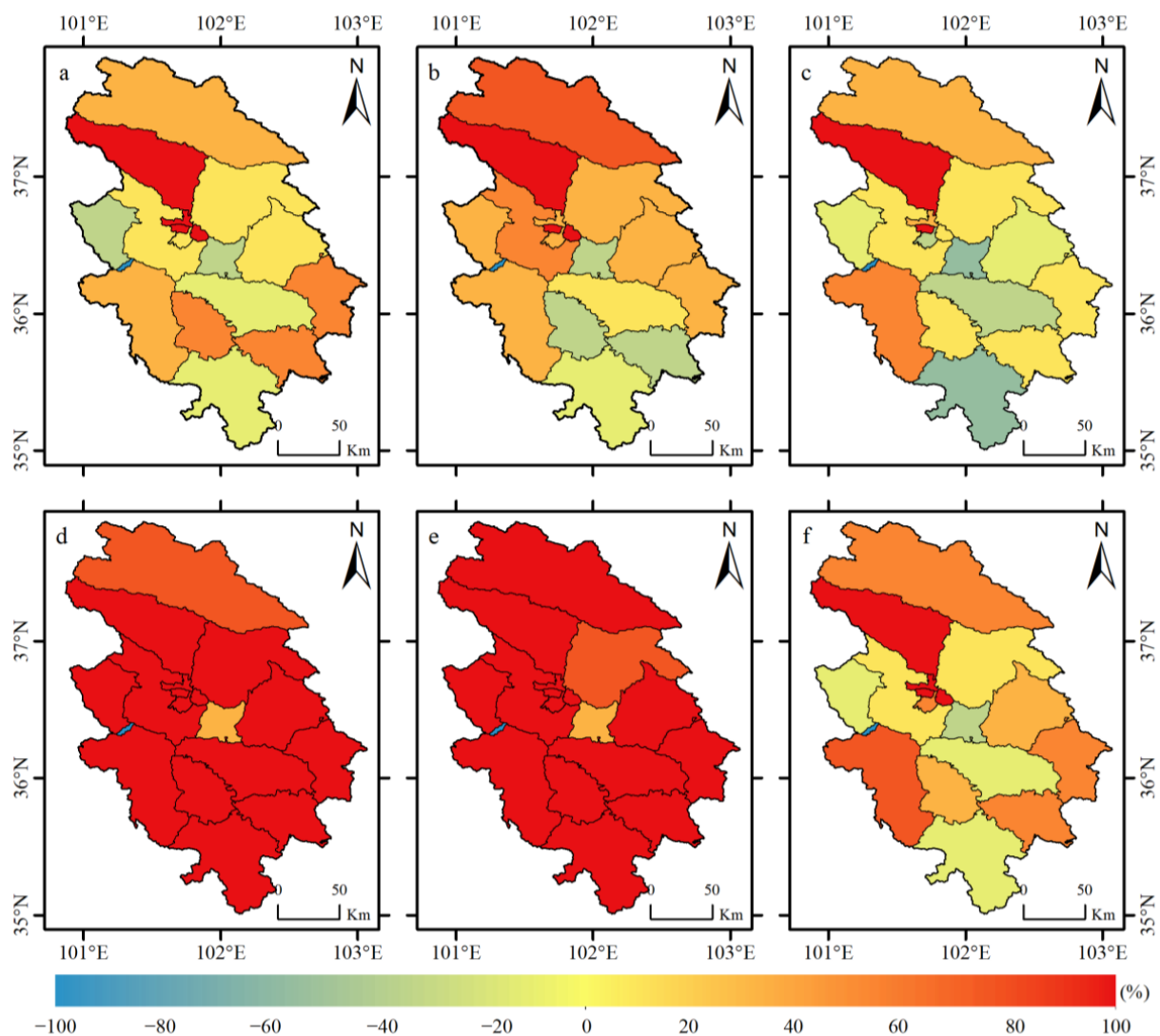

Figure 3. Illustration of the relative area difference (%) between the cropland distribution data products and statistical data in the Hehuang Valley. Note that the overestimation proportions exceeding 100% were truncated at 100% to maintain the balance of the color bar. The issue of severe overestimation was common in GL and GLC. (**a**) WC (10 m), (**b**) LC (10 m), (**c**) AIEC (10 m), (**d**) GL (30 m), (**e**) GLC (30 m), and (**f**) CLCD (30 m).
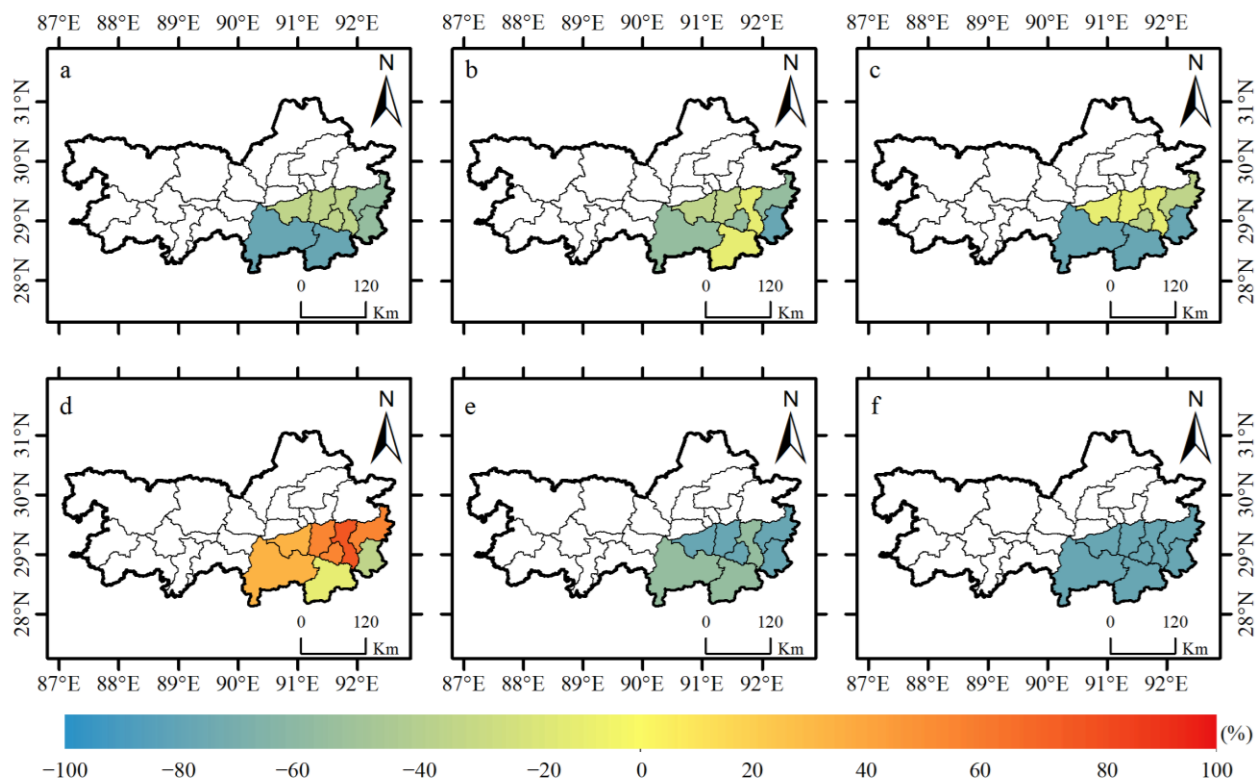
Figure 4. Illustration of the relative area difference (%) between the cropland distribution data products and statistical data in the YZR area. (**a**) WC (10 m), (**b**) LC (10 m), (**c**) AIEC (10 m), (**d**) GL (30 m), (**e**) GLC (30 m), and (**f**) CLCD (30 m).

Table 2. Fitness and deviation of data for the Hehuang Valley and the YZR area.

Data	WC (10 m)	LC (10 m)	AIEC (10 m)	GL (30 m)	GLC (30 m)	CLCD (30 m)
R^2	0.68/0.77	0.67/0.76	0.64/0.74	0.76/0.91	0.74/0.41	0.62/0.13
RMSE	1.25/0.38	1.65/0.34	1.33/0.31	3.46/0.19	3.39/0.51	1.37/0.63

In the YZR area, for 10 m resolution cropland distribution data, the R^2 value for WC was higher than that for LC and AIEC. Conversely, the *RMSE* of AIEC data was lower than that of LC and WC data, although the differences were relatively small. For the 30 m resolution cropland distribution data, GL achieved the highest R^2 value, followed by GLC, while CLCD had the lowest R^2 value. Additionally, the *RMSE* of GL data was lower than that of GLC and CLCD data.

Based on the comprehensive statistical results for area consistency, among the 10 m resolution cropland distribution data, WC demonstrated the best performance in terms of both similarity to and deviation from the statistical data. For the 30 m resolution cropland distribution data, GL showed the highest similarity with statistical data but also exhibited a greater degree of deviation, leading to a significant overestimation of cropland areas. Therefore, in terms of area consistency, WC was identified as the most suitable cropland distribution dataset for the study area.

3.2. Subsection Spatial Pattern Overlap Analysis Results

Figure 5 presents the spatial consistency analysis results of cropland distribution data at different resolutions in Hehuang Valley and the YZR area. Ideally, if the cropland distribution data from various products are accurate, then their spatial distributions should be consistent. However, Figure 5 indicates otherwise. For the 10 m resolution cropland distribution data, the proportion of high-consistency pixels was significantly lower than that of medium- and low-consistency pixels, with low-consistency pixels being the most

prevalent. For the 30 m resolution cropland distribution data, the proportion of high-consistency pixels was similarly lower than that of medium- and low-consistency pixels, and it was significantly lower than the results at the 10 m resolution.
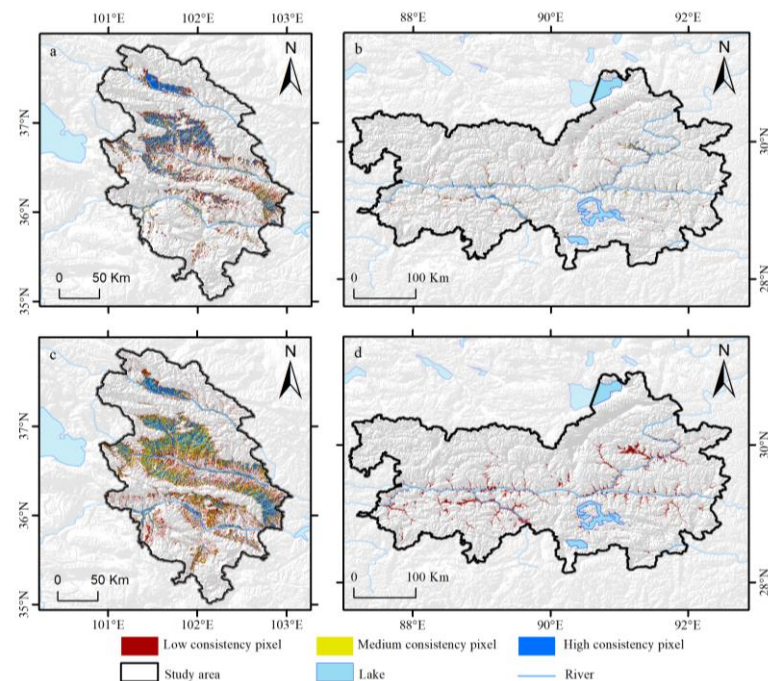

Figure 5. Spatial consistency among overlay results: (**a**) 10 m cropland distribution data in the Hehuang Valley area, (**b**) 10 m cropland distribution data in the YZR area, (**c**) 30 m cropland distribution data in the Hehuang Valley area, and (**d**) 30 m cropland distribution data in the YZR area.

In the Hehuang Valley, for the 10 m spatial resolution cropland distribution data overlay results (Figure 5a; Tables 3–5), the largest proportion of pixels marked as cropland was found for the low-consistency pixels (43.29%), followed by high-consistency pixels (33.73%) and medium-consistency pixels (22.98%). High-consistency pixels were primarily concentrated in the northern area along the northern bank of the Datong River and the central Huangshui River Basin, whereas medium- and low-consistency pixels were mainly found in the central Huangshui River Basin and the southern Yellow River Basin. Among all WC pixels, low-consistency pixels accounted for approximately 19.03%, medium-consistency pixels accounted for approximately 26.12%, and high-consistency pixels accounted for approximately 54.80%. For the LC pixels, low-consistency pixels accounted for approximately 35.87%, medium-consistency pixels accounted for approximately 17.89%, and high-consistency pixels accounted for approximately 46.21%. Among the AIEC pixels, low-consistency pixels represent approximately 9.65%, medium-consistency pixels represent approximately 30.06%, and high-consistency pixels represent approximately 60.29%. Additionally, when comparing pairs of cropland distribution data, WC and AIEC exhibited the highest consistency in this region at 59.13%, whereas WC and LC showed the lowest consistency at 42.17%.

Table 3. Overall consistency evaluation results for Hehuang Valley and the YZR area.

Pixels	10 m Resolution	30 m Resolution
Low-consistency pixels	43.29%/52.40%	39.67%/97.14%
Medium-consistency pixels	22.98%/23.68%	32.23%/2.83%
High-consistency pixels	33.73%/23.93%	28.10%/0.03%

Table 4. Proportion of consistent pixels for each 10 m cropland distribution data product for Hehuang Valley and the YZR area.

Pixels	WC	LC	AIEC
Low-consistency pixels	19.03%/32.43%	35.87%/39.73%	9.65%/17.87%
Medium-consistency pixels	26.12%/26.24%	17.89%/20.04%	30.06%/37.70%
High-consistency pixels	54.80%/41.34%	46.21%/40.23%	60.29%/44.43%

Table 5. Spatial consistency among cropland distribution data products for Hehuang Valley and the YZR area.

Index	Cropland Distribution Data A	Cropland Distribution Data B	Consistent Pixel Proportion
1	WC	LC	42.17%/30.21%
2	WC	AIEC	59.13%/47.03%
3	LC	AIEC	46.02%/39.92%
4	GL	GLC	53.13%/27.72%
5	GL	CLCD	40.54%/0.05%
6	GLC	CLCD	38.82%/1.44%

In the YZR area, for the 10 m spatial resolution cropland distribution data overlay results (Figure 5b; Tables 3–5), the largest proportion of pixels marked as cropland was found for low-consistency pixels (39.67%), followed by medium-consistency pixels (32.23%) and high-consistency pixels (28.10%). High-consistency pixels were mainly concentrated in the northeastern Lhasa River Basin and the eastern Yarlung Zangbo River and Nianchu River valleys, whereas medium- and low-consistency pixels were relatively widespread. Among all WC pixels, low-consistency pixels accounted for approximately 32.43%, medium-consistency pixels accounted for approximately 26.24%, and high-consistency pixels accounted for approximately 41.34%. For the LC pixels, low-consistency pixels accounted for approximately 39.73%, medium-consistency pixels accounted for approximately 20.04%, and high-consistency pixels accounted for approximately 40.23%. Among the AIEC pixels, low-consistency pixels represented approximately 17.87%, medium-consistency pixels represented approximately 37.70%, and high-consistency pixels represented approximately 44.43%. Additionally, an analysis of pairs of cropland distribution data showed that WC and AIEC exhibited the highest consistency in this region at 47.03%, while WC and LC showed the lowest consistency at 30.21%.

In the Hehuang Valley, for the 30 m spatial resolution cropland data overlay results (Figure 5c; Tables 3, 5 and 6), among all areas marked as cropland, the largest proportion was found for low-consistency pixels (52.40%), followed by high-consistency pixels (23.93%) and medium-consistency pixels (23.68%). High-consistency pixels were mainly concentrated in the northern area along the northern bank of the Datong River and the central and eastern Huangshui River Basin, although they presented broader coverage. Medium- and low-consistency pixels were mainly concentrated in the central Huangshui River Basin and southern Yellow River Basin, although they were more scattered and showed an increased area compared to the 10 m data. Among all GL pixels, low-consistency pixels accounted for approximately 23.98%, medium-consistency pixels accounted for approximately 38.24%, and high-consistency pixels accounted for approximately 37.78%. Among the GLC pixels, low-consistency pixels accounted for approximately 26.74%, medium-consistency pixels accounted for approximately 36.64%, and high-consistency pixels accounted for approximately 36.62%. For the CLCD pixels, low-consistency pixels accounted for approximately 3.51%, medium-consistency pixels accounted for approximately 21.18%, and high-consistency pixels accounted for approximately 75.27%. Among the pairwise overlay results of the three cropland distribution datasets, the highest consistency was observed between GL and GLC at 53.13%, while the lowest consistency was observed between GLC and CLCD at 38.82%.

Table 6. Proportion of consistent pixels for each 30 m cropland distribution data product for Hehuang Valley and the YZR area.

Pixels	GL	GLC	CLCD
Low-consistency pixels	23.98%/97.07%	26.74%/61.07%	3.51%/34.45%
Medium-consistency pixels	38.24%/2.89%	36.64%/38.51%	21.18%/49.43%
High-consistency pixels	37.78%/0.03%	36.62%/0.42%	75.27%/16.12%

In the YZR area, for the 30 m spatial resolution cropland data overlay results (Figure 5d; Tables 3, 5 and 6), among all areas marked as cropland, the vast majority were found among low-consistency pixels (97.14%), followed by medium-consistency pixels (2.83%), with very few found among high-consistency pixels (0.03%). The distribution of cropland pixels was generally consistent with the 10 m resolution data, although the overall area increased. Among all GL pixels, low-consistency pixels accounted for approximately 97.07%, medium-consistency pixels accounted for approximately 2.89%, and high-consistency pixels accounted for approximately 0.03%. Among all the GLC pixels, low-consistency pixels accounted for approximately 61.07%, medium-consistency pixels accounted for approximately 38.51%, and high-consistency pixels accounted for approximately 0.42%. Among all the CLCD pixels, low-consistency pixels accounted for approximately 34.45%, medium-consistency pixels accounted for approximately 49.43%, and high-consistency pixels accounted for approximately 16.12%. Among the pairwise overlay results of the three cropland distribution datasets, the highest consistency was observed between GL and GLC at 27.72%, while the lowest consistency was observed between GL and CLCD at 0.05%.

Regarding the spatial pattern analysis results mentioned above, among all the overlay results of cropland distribution data, the cropland pixels were primarily characterized by medium- to low-consistency pixels, with a smaller proportion characterized by high-consistency pixels. This indicates poor consistency among different products, which increases the difficulty of data application. Comparatively, the 10 m resolution cropland distribution data showed a more consistent performance across the two agricultural regions. AIEC performed the best, followed by WC, although the difference between the two was not significant, whereas LC performed the worst. For the 30 m resolution cropland distribution data, the performance varied significantly between the two agricultural regions. Performance in the Hehuang Valley area was better than that in the YZR area. CLCD performed the best in the Hehuang Valley area, whereas GL performed the best in the YZR area.

3.3. Positional Accuracy Results

Based on the validation sample points, this study constructed confusion matrices and obtained accuracy evaluation results for six cropland distribution datasets (*ACC*, *precision*, *MCC*, *TPR*, *FPR*, and *CEI*; Tables 7–10).

Table 7. Accuracy evaluation results of 10 m resolution cropland distribution data for Hehuang Valley and the YZR area.

	WC	LC	AIEC
ACC	0.87/0.94	0.86/0.93	0.86/0.94
Precision	0.76/0.94	0.78/0.84	0.81/0.96
MCC	0.58/0.64	0.55/0.56	0.54/0.63
TPR	0.57/0.48	0.51/0.41	0.47/0.45
FPR	0.05/0.0038	0.04/0.0091	0.03/0.0023
CEI	20/23	15/17	12/20

Table 8. Accuracy evaluation results of 30 m resolution cropland distribution data for Hehuang Valley and the YZR area.

	GL	GLC	CLCD
ACC	0.92/0.96	0.85/0.90	0.85/0.90
Precision	0.81/0.82	0.64/0.80	0.74/NaN
MCC	0.79/0.77	0.57/0.19	0.53/NaN
TPR	0.86/0.77	0.69/0.05	0.51/0.00
FPR	0.06/0.019	0.11/0.0015	0.05/0.00
CEI	28/26	17/8	11/11

Table 9. Accuracy evaluation results of 10 m resolution cropland distribution data in medium- and low-consistency areas for Hehuang Valley and the YZR area.

	WC	LC	AIEC
ACC	0.49/0.60	0.56/0.37	0.51/0.57
Precision	0.64/0.89	0.74/0.67	0.74/0.94
MCC	−0.06/0.24	0.14/−0.28	0.11/0.29
TPR	0.53/0.56	0.53/0.39	0.41/0.49
FPR	0.59/0.0027	0.38/0.0073	0.30/0.0013
CEI	6/16	13/4	11/17

Table 10. Accuracy evaluation results of 30 m resolution cropland distribution data in medium- and low-consistency areas for Hehuang Valley and the YZR area.

	GL	GLC	CLCD
ACC	0.79/0.81	0.44/0.21	0.47/0.18
Precision	0.78/0.82	0.49/0.80	0.51/NaN
MCC	0.59/0.10	−0.16/−0.03	−0.03/NaN
TPR	0.86/0.98	0.56/0.07	0.24/0.00
FPR	0.28/0.93	0.71/0.08	0.27/0.00
CEI	22/14	2/7	7/5

In the Hehuang Valley, for the evaluation results of 10 m resolution cropland distribution data (Table 7), the highest *ACC* was achieved by WC, and the lowest was found for LC and AIEC, with minimal differences among them. AIEC had the highest *precision*, whereas WC had the lowest. The highest *MCC* was observed for WC and the lowest was observed for AIEC, with only slight differences between the three products. WC exhibited the highest *TPR*, whereas LC exhibited the lowest. AIEC exhibited the lowest *FPR*, whereas WC exhibited the highest. WC exhibited the highest *CEI*, followed by LC, with AIEC exhibiting the lowest *CEI*. However, when considering only areas of medium and low consistency (Table 9), LC had the highest *ACC* and WC had the lowest. LC and AIEC had the highest *precision*, whereas WC had the lowest. LC showed the highest *MCC*, while WC showed the lowest. WC and LC had the highest *TPR*, while AIEC had the lowest. AIEC exhibited the lowest *FPR*, while WC exhibited the highest. LC had the highest *CEI*, while WC had the lowest. These results reveal that the performance of different 10 m resolution cropland distribution data varied across different indicators. Overall, WC was slightly superior to LC and AIEC. However, in areas with medium and low consistency, the *precision* of the cropland distribution data from WC was noticeably lower than that of LC and AIEC. In these regions, the cropland distribution data from LC performed relatively better.

In the YZR area, the evaluation results for 10 m resolution cropland data (Table 7) showed that WC and AIEC both had the highest *ACC*, while LC had the lowest, with minimal differences among the three. AIEC and LC had the highest and lowest *precision*, respectively. The highest *MCC* was observed for WC, and the lowest was observed for LC. WC exhibited the highest *TPR*, whereas LC exhibited the lowest, with minimal differences among the three products. AIEC had the lowest *FPR*, whereas LC had the highest. WC

had the highest *CEI*, followed by AIEC, with LC having the lowest. When considering only areas of medium and low consistency (Table 9), WC had the highest *ACC*, LC had the lowest AIEC but highest *precision*, LC had the lowest *precision*, AIEC had the highest *MCC*, and LC had the lowest *MCC*. WC had the highest *TPR*, while LC had the lowest. AIEC exhibited the lowest *FPR*, while LC exhibited the highest. AIEC exhibited the highest *CEI*, followed by WC, and LC exhibited the lowest. These results indicate that in the YZR area, both the overall accuracy evaluation and the accuracy evaluation based on medium to low consistency showed relatively consistent performances across the different datasets. This indicates that the cropland distribution data from both WC and AIEC performed relatively well.

In the Hehuang Valley, the evaluation results for 30 m spatial resolution cropland data (Table 8) showed that GL had the highest *ACC* while GLC and CLCD both had the lowest. GL had the highest *precision*, whereas GLC had the lowest. GL had the highest *MCC*, whereas CLCD had the lowest. GL exhibited the highest *TPR*, while CLCD had the lowest. CLCD had the lowest *FPR*, while GLC has the highest. GL presented the highest GEI, followed by GLC, with CLCD exhibiting the lowest. When considering only areas of medium and low consistency (Table 10), GL had the highest *ACC*, while GLC had the lowest. GL had the highest *precision*, while GLC had the lowest. GL had the highest *MCC*, while GLC had the lowest. GL had the highest *TPR*, while CLCD had the lowest. CLCD exhibited the lowest *FPR*, while GLC exhibited the highest. GL presented the highest *CEI*, while GLC presented the lowest. Thus, in this region, GL performed the best, while GLC performed the worst. Moreover, the 30 m resolution cropland distribution data products in the Hehuang Valley showed consistent performance across various indicators, with GL having the best performance and GLC presenting the worst.

In the overlay analysis of the 30 m resolution cropland data in the YZR area, CLCD was excluded because of its poor cropland completeness and inability to meet the accuracy assessment criteria. For cropland distribution data obtained from GL and GLC, GL outperformed GLC in all accuracy evaluation metrics except for the *FPR* (Table 8). The same results were found in the evaluation based on medium and low consistency (Table 10). Therefore, the performance of GL was also superior in the 30 m resolution cropland data for the YZR area.

Regarding the accuracy evaluation results, for 10 m cropland distribution data, WC had the best performance, whereas for 30 m cropland distribution data, GL had the best performance.

4. Discussion

4.1. Differences in the Presentation of Details

To further verify the performance of various cropland distribution data in detail, this paper selected four representative areas in the Hehuang Valley region for comparison: slope cropland (Figure 6a), concentrated cropland distribution areas (Figure 6b), urban green spaces (Figure 6c), and cropland mixed with other land types (Figure 6d).

All cropland distribution data could be used to roughly extract the distribution of croplands; however, differences remained among the various cropland distribution data. Specifically, for the 10 m resolution cropland distribution data, WC incorrectly identified urban green spaces as croplands. AIEC and LC could not effectively distinguish between forest and grassland, incorrectly classified them as croplands, and misidentified some built-up areas as croplands. For the 30 m resolution cropland distribution data, GL also failed to distinguish among forest, grassland, and built-up areas and commonly misclassified the land types at the edges of cropland as cropland. This may explain why the extraction results were much higher than the statistical and other cropland distribution data. GLC incorrectly identified urban green spaces as croplands. CLCD exhibited the most severe errors in misclassification and omission, with many omissions in slope croplands, and it failed to distinguish forest and grassland well and incorrectly identified urban green spaces as croplands.
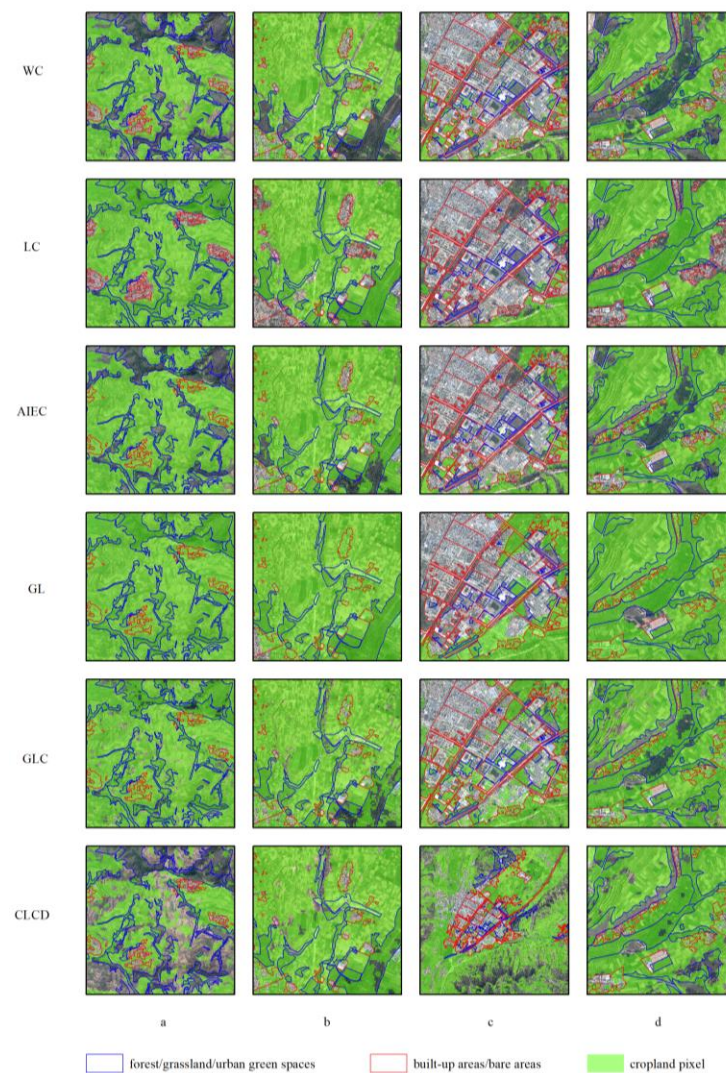
Figure 6. Details of cultivated land data in the Hehuang Valley. (**a**) Slope cropland, (**b**) areas with concentrated cropland distribution, (**c**) urban green spaces, and (**d**) areas with a mixture of cropland and other land types. WC (10 m), LC (10 m), AIEC (10 m), GL (30 m), GLC (30 m), and CLCD (30 m). Note: The process of manual visual interpretation primarily relies on sub-meter resolution remote sensing imagery, supplemented by auxiliary data, such as DEM data, ground survey samples, and other cropland distribution data. In this context, the primary focus is on the misclassification of cropland. Therefore, forest, grassland, and urban green spaces were categorized as a single class, while built-up areas and bare areas were grouped into another class.

Similarly, this study randomly selected two representative regions in the YZR area, as shown in Figure 7. For the 10 m resolution cropland distribution data, the overall results were consistent with those of the Hehuang Valley, which generally reflects the overall distribution characteristics of the cropland. However, compared with WC, both LC and AIEC showed more severe omission errors, even among adjacent plots. The 30 m resolution cropland distribution data differed significantly from the results in the Hehuang Valley. Notably, GLC and CLCD barely extracted any cropland, with only a few cropland pixels captured, whereas GL roughly reflected the distribution characteristics of the cropland but misclassified the land types at the cropland edges.
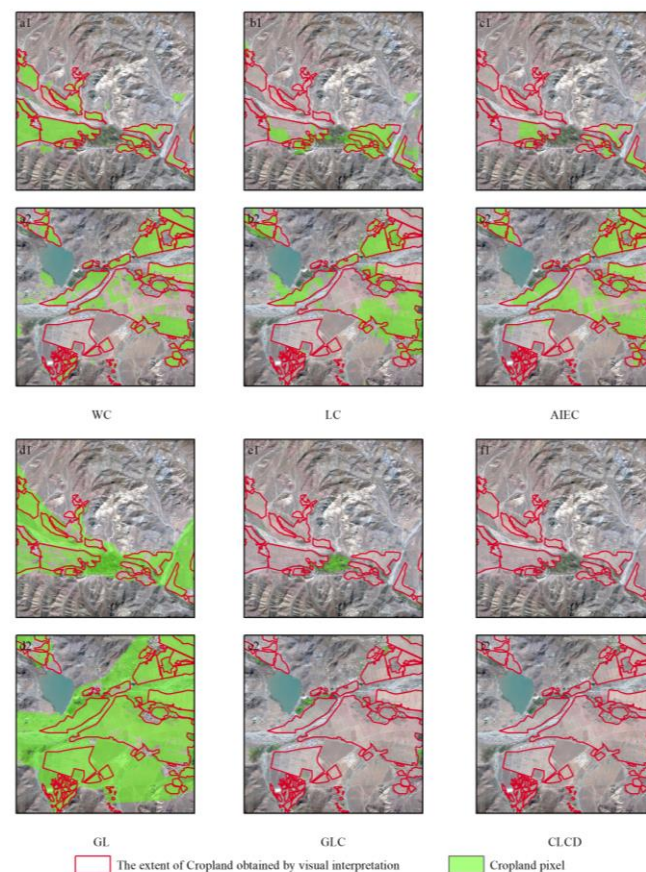
Figure 7. Details of cultivated land data in the YZR area. (**a1,a2**) WC (10 m), (**b1,b2**) LC (10 m), (**c1,c2**) AIEC (10 m), (**d1,d2**) GL (30 m), (**e1,e2**) GLC (30 m), and (**f1,f2**) CLCD (30 m). Note: The process of manual visual interpretation primarily relies on sub-meter resolution remote sensing imagery, supplemented by auxiliary data, such as DEM data, ground survey samples, and other cropland distribution data. In this case, the focus is on the omission of cultivated land. Consequently, only the cropland category was interpreted.

Detailed analysis of the map differences above indicated that the 10 m resolution cropland distribution data can generally depict the distribution of croplands in both sub-regions. In contrast, among the 30 m resolution cropland distribution data products, only GL provided an adequate representation. Furthermore, regardless of whether the spatial resolution was 10 m or 30 m, most cropland distribution data products misclassified grassland, woodland, and cropland. Therefore, to update and produce new cropland distribution data for the study area, a partitioned mapping strategy should be adopted. In addition, more attention should be paid to addressing the misclassification of croplands, grasslands, and woodlands.

4.2. Factors Influencing the Classification Results

Due to the complex spectral characteristics of croplands, their spectral characteristics early in the crop growth cycle or in poorly growing plots may resemble those of grasslands. In addition, the planting of shelterbelts or fruit trees along roadsides or between plots may cause the spectral characteristics of croplands to resemble those of forestland. In regions with complex land types, especially in transitional areas between cropland and grassland or forest land, significant "foreign matter with the same spectrum" phenomena may occur during the growing season, leading to misclassification [11,25,26]. Numerous studies [27–30] have shown that apart from the inherent spectral characteristics of land cover, the differences in classification results are related to the classification system used, spatial resolution of the original imagery, classification method, and samples selected [31,32]. This paper

explored the reasons for the significant differences in cropland data between the Hehuang Valley area and the YZR area.

First, from the perspective of classification systems, WC defines croplands as land cultivated with annual crops that are sown and harvested at least once within a 12-month period following the sowing or planting date. Annual cropland typically produces an herbaceous cover and may occasionally include trees or woody vegetation. Perennial woody crops are classified as either tree cover or shrub land, as appropriate. Greenhouses are categorized as built-up areas. LC defines cropland as land cultivated for cereals, grasses, and non-tree crops, such as corn, wheat, soybeans, and fallow land. GL defines cropland as land utilized for growing a variety of crops, including paddy fields, irrigated and rain-fed dry land, vegetable gardens, pastures, and greenhouses. It also encompasses areas primarily used for crops that are interspersed with fruit trees and other economic trees, such as tea gardens, coffee plantations, and shrub-based economic crops. GLC defines croplands as areas where natural vegetation has been removed or altered and replaced by anthropogenic vegetation cover that is maintained through human activities. CLCD defines croplands as paddy fields, greenhouse agriculture, and other types of farmlands (such as arable and cultivated land). Fruit trees are classified as forests, and pastures may transition from farmland to natural grassland. AIEC lacks a precise definition of cropland. The classification systems of the various datasets exhibited significant differences in defining croplands, particularly concerning forage land, fallow land, facility agriculture, and economic tree species. The study area, situated in the agro-pastoral transitional zone, features widespread forage land, facility agricultural land, and fallow land. These definitional differences resulted in considerable discrepancies in cropland area estimates. Furthermore, the categorization of facility agriculture and economic tree species influences the spatial patterns of cropland distribution. These discrepancies affect not only the consistency and comparability of the data but also the accuracy and reliability of the cropland extraction results in the study area. Because of the absence of statistical indicators at the county or district level, this study used the crop planting area as a reference for estimating the cropland area. Crop planting area refers to the actual land where crops are sown or transplanted, thereby encompassing any area with planted crops, regardless of whether it is classified as cropland or non-cropland. It also includes areas where crops are replanted or supplemented following disasters during the sowing season. Currently, crop planting areas encompass nine categories: cereals, cotton, oilseeds, sugar crops, fiber crops, tobacco, vegetables and melons, medicinal plants, and other crops. This reference value may have some uncertainties owing to fallow periods and crop rotation in facility agriculture. Consequently, the crop planting area is likely to be slightly larger than the actual cropland area; however, the difference should be minor and offset by fallow land. For regions dominated by single-season agriculture, the discrepancy between the crop planting area and actual cropland area should be minimal. Additionally, a comparison of multiyear statistical data revealed that the crop planting area used in this study showed stable trends over time, indicating that the statistical data employed in this research were reliable and had reference value.

A comparison of the statistical data with the cropland distribution data revealed that in the Hehuang Valley area, most cropland distribution datasets tended to overestimate the cropland area relative to the statistical figures, with the exception of AIEC. Conversely, in the YZR area, most cropland distribution datasets tended to underestimate the cropland area, except for GL. This discrepancy suggests that the accuracy of cropland distribution data is influenced not only by the definitions of cropland within the classification systems but also by other significant factors.

From the perspective of remote sensing image resolution, a comparison of the overlapping results between the 10 m and 30 m resolution cropland data showed that the proportion of high-consistency pixels was higher in the 10 m resolution data than in the 30 m resolution data (Table 3; Figure 5), with the distribution being more concentrated. This suggests that with increased spatial resolution, some pixels identified as cropland in

the 30 m data were reclassified as other land types. Therefore, spatial resolution is likely a key factor influencing the spatial consistency of cropland data. Additionally, discrepancies in the overlapping results may have been caused by different processing methods applied to remote sensing images by the various products during the registration process. However, during the registration process, discrepancies in the methods used by different products to process remote sensing imagery may introduce uncertainty in the overlapping results.

From the mechanism of classification methods, this study categorized the classification methods used by the six cropland data products into two types: traditional machine learning and deep learning. Traditional machine learning methods simply extract features based on the samples and require fewer samples, whereas deep learning methods can perform deeper feature extraction but need a large number of samples for support. Among the 10 m resolution cropland distribution data, only WC employed traditional machine learning methods, whereas LC and AIEC used deep learning classification schemes. Using the same series of satellite remote sensing images, the performance of these three cropland distribution data products in extracting cropland did not show significant differences between the two sub-study areas. This indicates that cropland extraction did not show notable differences between the traditional machine learning and deep learning methods. However, for the 30 m resolution cropland distribution data, GL, GLC, and CLCD all used traditional machine learning classification methods and nearly identical series of satellite remote sensing images. While GL's classification performance did not show significant differences between the two sub-study areas, GLC and CLCD extracted a relatively complete range of cropland distribution in the Hehuang Valley (Figure 8) but performed poorly in the YZR area (Figure 9). Although GL used a different classifier than GLC and CLCD, numerous studies indicate that the differences in classification performance between various traditional machine learning classifiers are not very pronounced [33–35]. This suggests that in addition to differences in the classification method, differences in training samples and environmental factors in the sub-study areas may influence the classification results.
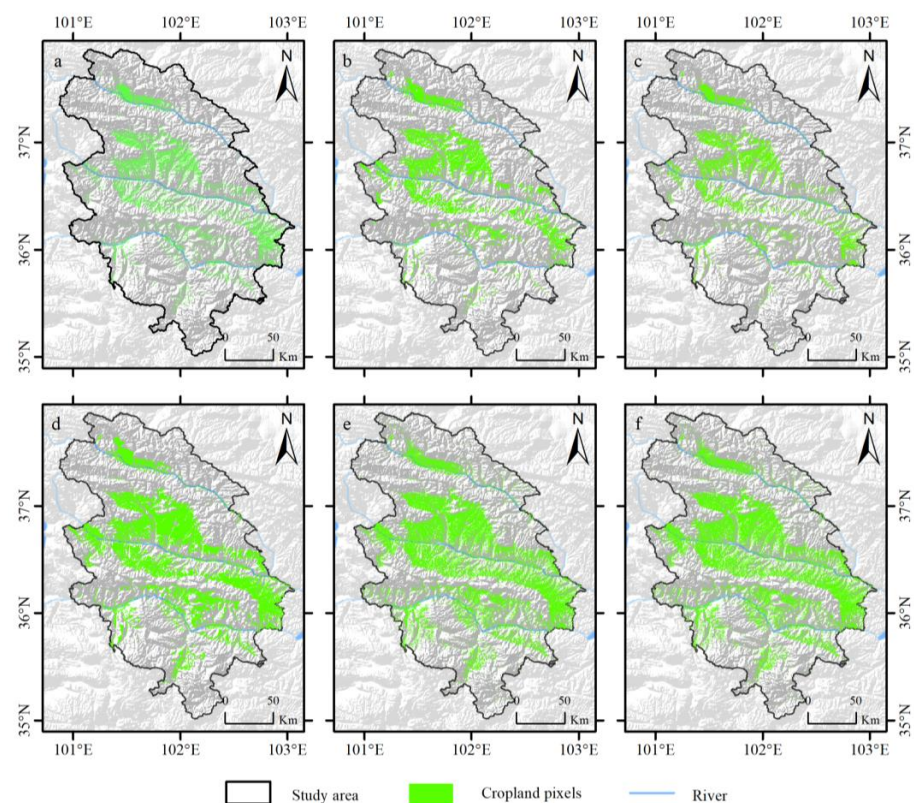

Figure 8. Pixel distribution of cropland distribution data in the Hehuang Valley. (**a**) WC (10 m), (**b**) LC (10 m), (**c**) AIEC (10 m), (**d**) GL (30 m), (**e**) GLC (30 m), and (**f**) CLCD (30 m).
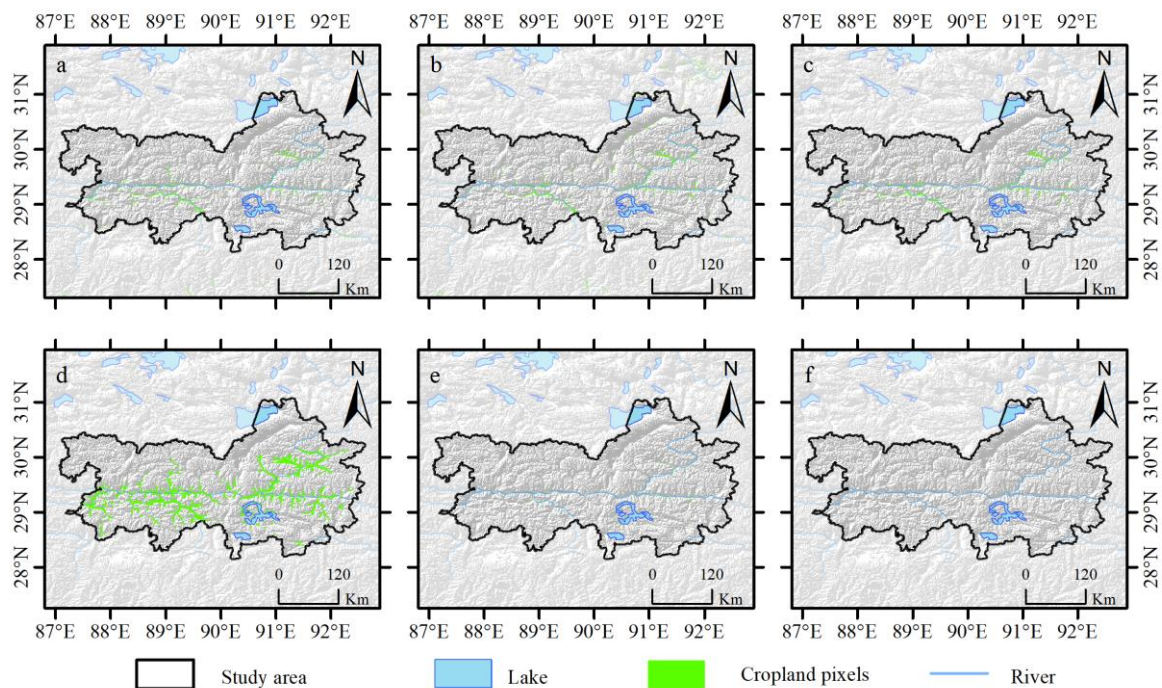
Figure 9. Pixel distribution of cropland distribution data in the YZR area. (**a**) WC (10 m), (**b**) LC (10 m), (**c**) AIEC (10 m), (**d**) GL (30 m), (**e**) GLC (30 m), and (**f**) CLCD (30 m).

From the perspective of training samples, among the 10 m resolution cropland distribution data products, LC gathered tens of billions of training samples during production. Although AIEC has not explicitly disclosed information about its training samples, its performance in both study areas was generally consistent with that of LC (Figures 8 and 9). Given that LC and AIEC use deep learning, which still requires a massive amount of training samples, the training samples likely covered the two study areas well, leading to generally consistent evaluation results (Figures 8 and 9). In the 30 m resolution cropland distribution data, GL primarily relied on expert visual interpretation. GLC has built a spatiotemporal spectral library with training samples mainly derived from the GlobCover2009 and CCI_LC datasets, whereas CLCD training samples are largely randomly drawn from the invariant areas of China's Land Use/Cover Datasets (CLUDs). Since the samples were obtained from large-scale datasets, this study examined these source datasets and the training sample sets of GLC and CLCD. Because of the lack of effective classification of cropland in the YZR area in these source datasets, cropland sample points were not included in the training sample sets of GLC and CLCD. This could be attributed to issues with the sampling strategy because the relatively small cropland area in the study area may have resulted in few or no sample points being allocated. Therefore, by comparing the extraction of cropland by LC, AIEC, GL, GLC, and CLCD in the two sub-study areas, this study concluded that when using the same resolution remote sensing imagery and classification methods, although the impact of the classification method cannot be completely ruled out, the training samples will likely represent the determining factor for the differences in the classification results within the study area.

4.3. Factors Driving Spatial Consistency Results

Previous studies have demonstrated a strong correlation between cropland distribution and terrain factors, such as slope, elevation, and surface complexity [16]. In this context, this study examined the relationship between the spatial consistency of cropland data and terrain factors.

An analysis of the proportion of overlapping cropland data at different slopes for the two spatial resolutions (Figures 10a,b and 11a,b) revealed that as the slope increased,

the proportion of high-consistency pixels gradually decreased, while the proportion of low-consistency pixels gradually increased. This finding is consistent with the conclusions of other studies [14].
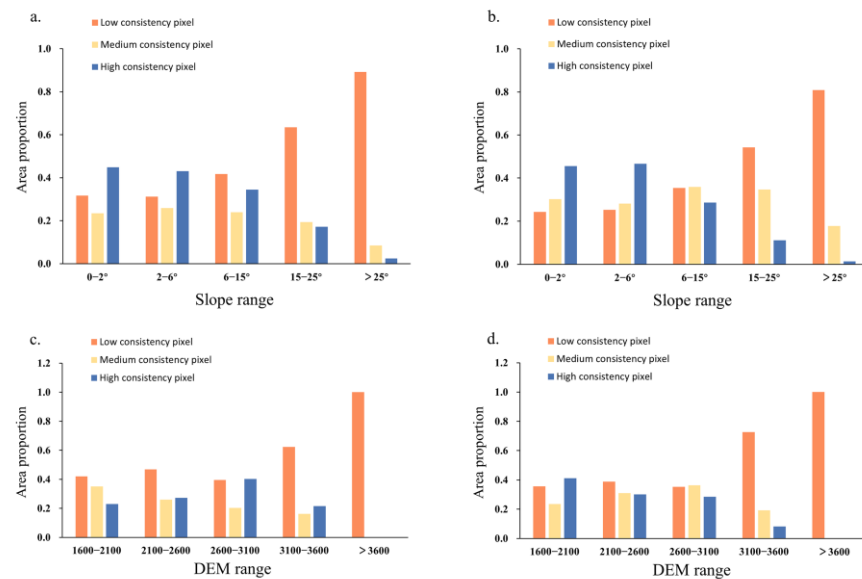

Figure 10. Area proportion of pixels with different consistencies in different terrain factor ranges in the Hehuang Valley region. (**a**) Proportion of consistent pixels among the 10 m cropland distribution data products at different slope ranges. (**b**) Proportion of consistent pixels among the 30 m cropland distribution data products at different slope ranges. (**c**) Proportion of consistent pixels among the 10 m cropland distribution data products at different elevation ranges. (**d**) Proportion of consistent pixels among the 30 m cropland distribution data products at different elevation ranges.
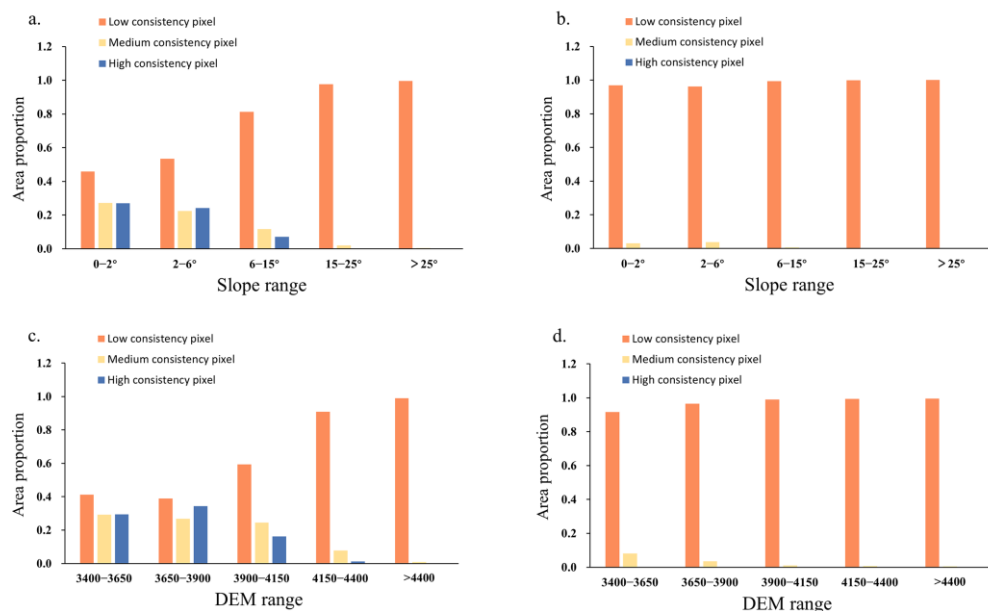

Figure 11. Area proportion of pixels with different consistencies among the different terrain factor ranges in the YZR area. (**a**) Proportion of consistent pixels among the 10 m cropland distribution data products at different slope ranges. (**b**) Proportion of consistent pixels among the 30 m cropland distribution data products at different slope ranges. (**c**) Proportion of consistent pixels among the 10 m cropland distribution data products at different elevation ranges. (**d**) Proportion of consistent pixels among the 30 m cropland distribution data products at different elevation ranges.

However, an analysis of the proportion of overlapping cropland data at different elevations (Figures 10c,d and 11c,d) revealed that in the Hehuang Valley area, for the 10 m resolution overlapping results, the proportion of high-consistency pixels initially decreased, and the proportion of low-consistency pixels increased with the increasing elevation. However, within the elevation range of 2600–3100 m, the proportion of high-consistency pixels increased, and the proportion of low-consistency pixels decreased. For the 30 m resolution overlapping results, although less pronounced, the proportion of low-consistency pixels also decreased within the 2600–3100 m elevation range, whereas the proportion of medium-consistency pixels increased, and high-consistency pixels only slightly decreased. A similar phenomenon was observed in the YZR area. For the 10 m resolution overlapping results, within the elevation range of 2650–3900 m, the proportion of high-consistency pixels increased, and low-consistency pixels decreased. This finding is inconsistent with previous studies [16]. This study suggests that in addition to topographic factors, other driving factors also influence the extraction results of croplands.

A plausible explanation for this is that although low-altitude areas have better natural conditions that are more suitable for crop cultivation, human activities are more frequent in these areas. Increased urbanization and non-agricultural activities have led to more common land use changes from cropland to other types, resulting in a lower proportion of high-consistency pixels between different datasets owing to the complexity of surface types. In contrast, in mid-altitude areas, natural conditions worsen with increasing elevation, although human activities are relatively less frequent, which reduces the occurrence of land use changes from croplands to other types. Therefore, the stability of the surface cover led to a higher proportion of high-consistency pixels. As the altitude increases, the natural conditions for crop growth deteriorate, and farmers are more likely to plant forage to meet their livestock feed needs (Figure 12). This results in more regular textures in these areas. Additionally, differences in cropland definitions and spectral confusion may have led to the misidentification of these areas as cropland, causing a decline in the spatial consistency of cropland classes. In summary, the extraction results of cropland distribution data were significantly influenced by topographic factors as well as the frequency of human activities and variations in cropland definitions. The interactions of these elements across various altitude ranges created intricate patterns of spatial consistency in cropland classification. Future research should focus on further exploring and quantifying these driving factors to enhance the accuracy and reliability of cropland distribution data extraction.
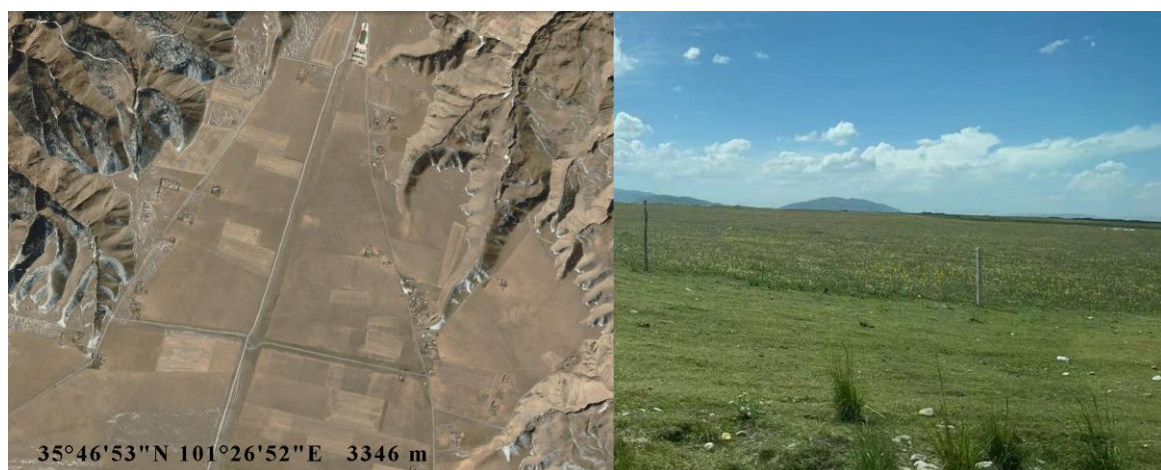

Figure 12. Forage fields with cultivated grain at high elevations.

4.4. Applicability of Existing Cropland Distribution Data

This study evaluated the quality of six cropland distribution data products in three dimensions: area consistency, spatial pattern analysis, and sample accuracy. For the 10 m resolution cropland distribution data, all three data products performed well in

both study areas, although LC and AIEC did not perform as well as WC. For the 30 m resolution cropland distribution data, GLC and CLCD performed worse than GL and showed significant discrepancies between the two study areas, making them unsuitable for application in alpine agricultural regions. Overall, cultivated land distribution data at a 10 m resolution generally provide more detailed information compared to 30 m resolution data and are preferable for single-period analysis. However, since these products are typically derived from Sentinel-2 imagery, their temporal coverage is limited, making them less suitable for long-term time-series analysis, particularly as WC only covers 2020 and 2021. In contrast, 30 m resolution data are based on Landsat imagery, offering sufficient temporal coverage for time-series analysis. Nevertheless, its accuracy may be inferior to that of the 10 m resolution data.

While WC and GL showed good performance, they also presented certain issues. For example, WC performed poorly in low-consistency areas in Hehuang Valley, and GL significantly overestimated cropland areas. In general, the 10 m resolution WC and 30 m resolution GL can be used under the current conditions. In practice, however, a single set of cropland distribution data may not be highly suitable and reliable for both agricultural regions to meet the needs of precision agricultural management. Therefore, to support the increasingly sophisticated alpine agricultural production management needs, high-quality localized cropland distribution data must be developed for detailed applications.

4.5. Innovations and Limitations

The novelty of this study lies in several key aspects. First, previous studies [36] often resampled high spatial resolution products to match lower spatial resolution products before conducting comparisons. They demonstrated that, across different scales, even when the 10 m product was resampled to match the 30 m product for comparison, the superior accuracy and stability of the 10 m product, due to the higher spatial resolution of its corresponding remote sensing data, remained evident. Moreover, the 10 m product consistently exhibited finer detail. Resampling did not fully negate the advantages conferred by the higher resolution of the 10 m product. Our comparison revealed that the 10 m product demonstrated significantly greater spatial stability compared to the 30 m product. For example, some 30 m products showed substantial differences between the two study areas. Moreover, our discussion on the applicability of existing cultivated land distribution data suggested that when opting for single-period products, 10 m resolution products should be prioritized, whereas for long-term series, 30 m products are more appropriate. Thus, conducting separate comparisons for the two resolutions is undeniably a more effective strategy, as it allows for a more comprehensive demonstration of differences between products at the same spatial resolution. Additionally, the cultivated land distribution products selected for this study are of relatively recent production years, rendering the evaluation results more aligned with current data selection requirements.

Nevertheless, this study also presents several limitations. First, while the six products selected each covered multiple periods, only 2020 served as the common overlap. Thus, 2020 was chosen as the baseline year. However, relying on the evaluation results from a single year to represent each product may result in occasionality. Second, validation sample points were acquired through visual interpretation. Although this process incorporated a comprehensive analysis of remote sensing imagery from the 2020 growing season (Landsat 8, Sentinel-2, and Google Earth), DEM data, ground survey samples, and other cultivated land distribution data to ensure interpretative accuracy, the presence of rotational farming, fallowing, and phenomena, such as 'same object, different spectra' or 'different objects, same spectrum', may still result in misclassification. This could potentially influence the accuracy assessments of certain products. Moreover, the cultivated land reference data for some counties within the YZR area were unavailable, as only partial data could be obtained. This limitation could result in potential deviations between the current R^2 and *RMSE* evaluation results and the actual values. Lastly, this study hypothesized that training samples may serve as the key factor influencing classification results. However, we only compared

classification systems, remote sensing image resolutions, and classification methods independently, without investigating the interactive effects of these factors. Additionally, as none of the products explicitly disclosed the training sample data they employed, we could only infer sample quantity based on their classification methods and estimate sample quality from their production processes, making it impossible to definitively determine whether sample quantity or quality is the crucial factor.

5. Conclusions

Based on the support of the GEE, this study evaluated the accuracy, spatial consistency, and area consistency of six high-resolution cropland distribution datasets from both domestic and international sources generated in 2020 for the Hehuang Valley of Qinghai Province and the YZR area of the Tibet Autonomous Region. Additionally, the factors that influenced the classification results were analyzed. The results indicated that:

(1) In terms of area consistency, WC performed best among the 10 m resolution cropland distribution data in terms of both similarity and deviation from the statistics, while GL performed the best among the 30 m resolution data in terms of similarity, although it also showed a higher degree of deviation and led to the most serious overestimation of cultivated land area. Therefore, from the perspective of area consistency, the WC product may be the most suitable cultivated land distribution data product for this region.

(2) In terms of spatial pattern overlap, the proportion of low-consistency pixels was the highest among all overlaid cropland distribution data. In the Hehuang Valley area, high-consistency pixels were primarily concentrated in the northern area along the northern bank of the Datong River and central Huangshui River Basin, whereas medium- and low-consistency pixels were mainly found in the central Huangshui River Basin and southern Yellow River Basin. In the YZR area, high-consistency pixels were mainly concentrated in the northeastern Lhasa River basin and eastern Yarlung Zangbo River and Nianchu River valleys, whereas medium- and low-consistency pixels were relatively widespread. For the 10 m resolution cropland distribution data, AIEC performed the best, followed by WC, although the difference between the two was not significant. For the 30 m resolution cropland distribution data, CLCD performed the best in the Hehuang Valley area, while GL performed the best in the YZR area.

(3) In terms of positional accuracy, WC showed the best overall performance among the 10 m resolution cropland distribution data across various accuracy indicators, while GL showed the best overall performance among the 30 m resolution cropland distribution data.

In summary, GL and WC may be the best-performing cropland distribution data products in the alpine agricultural areas of the Qinghai-Tibet Plateau; however, a single cropland distribution dataset was not able to provide the best performance across all evaluation metrics for both agricultural regions. Therefore, when conditions permit, high-quality cropland distribution data should be further developed to support the evolving needs of high-altitude agricultural production management.

Author Contributions: Conceptualization, X.X.; methodology, S.L. and X.X.; software, S.L.; validation, S.L.; formal analysis, S.L. and X.X.; investigation, S.L. and X.X.; resources, Q.C. and Y.P.; data curation, S.L.; writing—original draft preparation, S.L. and X.X.; writing—review and editing, S.L., X.X. and Y.P.; visualization, S.L.; supervision, X.X. and Y.P. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: WC: https://esa-worldcover.org/en (accessed on 30 July 2024) (in Chinese); LC: https://livingatlas.arcgis.com/landcover/ (accessed on 30 July 2024) (in Chinese); AIEC: https://engine-aiearth.aliyun.com/#/dataset/DAMO_AIE_CHINA_LC (accessed on 30 July

Conflicts of Interest: The authors declare that they have no competing financial interests or personal relationships that may have influenced the work reported in this study.

References

1. Wu, W.; Shibasaki, R. Remotely sensed estimation of cropland in China: A comparison of the maps derived from four global land cover datasets. *Can. J. Remote Sens.* **2008**, *34*, 467–479. [CrossRef]
2. 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. *Zenodo* **2021**. [CrossRef]
3. 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), Online, 12–16 July 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 4704–4707.
4. Liu, S.; Wang, H.; Hu, Y. Land Use and Land Cover Mapping in China Using Multi-modal Fine-grained Dual Network. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 1–19.
5. Jun, C.; Ban, Y.; Li, S. Open access to Earth land-cover map. *Nature* **2014**, *514*, 434. [CrossRef]
6. Zhang, X.; Liu, L.; Chen, X.; Gao, Y.; Xie, S.; Mi, J. GLC_FCS30: Global land-cover product with fine classification system at 30 m using time-series Landsat imagery. *Earth Syst. Sci. Data* **2021**, *13*, 2753–2776. [CrossRef]
7. 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]
8. Bai, Y.; Feng, M. Data fusion and accuracy evaluation of multi-source global land cover datasets. *Acta Geogr. Sin.* **2018**, *73*, 2223–2235. (In Chinese)
9. Wu, Z.; Cai, Z.; Guo, Y. Accuracy evaluation and consistency analysis of multi-source remote sensing land cover data in the Yellow River Basin. *Chin. J. Eco Agric.* **2023**, *31*, 917–927. (In Chinese)
10. Hu, Y.; Zhang, Q.; Dai, Z.; Huang, M.; Yan, H. Agreement analysis of multi-sensor satellite remote sensing derived land cover products in the Europe Continent. *Geogr. Res.* **2015**, *34*, 1839–1852. (In Chinese)
11. Huang, Y.; Liao, S. Regional accuracy assessments of the first global land cover dataset at 30 m resolution: A case study of Henan province. *Geogr. Res.* **2016**, *35*, 1433–1446. (In Chinese)
12. Venter, Z.; Barton, D.; Chakraborty, T.; Simensen, T.; Singh, G. Global 10 m Land Use Land Cover Datasets: A Comparison of Dynamic World, World Cover and Esri Land Cover. *Remote Sens.* **2022**, *14*, 4101. [CrossRef]
13. 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.* **2024**, *23*, 283–297. [CrossRef]
14. 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]
15. Cheng, Y.; Zhao, W.; Jiao, J.; Zhang, L.; Cao, X.; Chen, T.; Li, J.; Zhang, Z. Soil conservation effect of cropland use change in the Yellow River-Huangshui River Valley over the past 20 years. *Sci. Soil Water Conserv.* **2023**, *21*, 55–63. (In Chinese)
16. Chen, Y.; Hua, S.; Li, Y. Consistency analysis and accuracy assessment of multi-source land cover products in the Yangtze River Delta. *Trans. Chin. Soc. Agric. Eng.* **2021**, *37*, 142–150. (In Chinese)
17. Arino, O.; Bicheron, P.; Achard, F.; Latham, J.; Witt, R.; Weber, J.L. The most detailed portrait of Earth. *Eur. Space Agency* **2008**, *136*, 25–31.
18. ESA. Land Cover CCI Product User Guide Version 2. Tech. Rep. 2017. Available online: https://maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-Ph2-PUGv2_2.0.pdf (accessed on 30 July 2024).
19. Buchhorn, M.; Lesiv, M.; Tsendbazar, N.; Bertels, L.; Smets, B. Copernicus global land cover layers—Collection 2. *Remote Sens.* **2020**, *12*, 1044. [CrossRef]
20. Sentinel-2 10 m Land Use/Land Cover Time Series. Available online: https://www.arcgis.com/home/item.html?id=cfcb7609de5f478eb7666240902d4d3d (accessed on 30 July 2024).
21. Description of the NASADEM_HGT v001. Available online: https://lpdaac.usgs.gov/products/nasadem_hgtv001/ (accessed on 30 July 2024).
22. National Catalogue Service For Geographic Information. Available online: https://www.webmap.cn/main.do?method=index (accessed on 30 July 2024). (In Chinese).
23. Yang, Y.; Xiao, P.; Feng, X.; Li, H. Accuracy assessment of seven global land cover datasets over China. *ISPRS J. Photogramm. Remote Sens.* **2017**, *125*, 156–173. [CrossRef]
24. Stehman, S. Estimating area from an accuracy assessment error matrix. *Remote Sens. Environ.* **2013**, *132*, 202–211. [CrossRef]
25. Zhu, J.; Su, J.; Qin, F.; Wang, H. Accuracy Assessment of the 1:100 000 Land Cover Data of Henan Province in 2015. *China Land Sci.* **2019**, *33*, 59–67. (In Chinese)

26. Tong, R.; Yang, Y.; Chen, X. Consistent Analysis and Accuracy Evaluation of Multisource Land Cover Datasets in 30 m Spatial Resolution over the Mongolian Plateau. *J. Geo-Inf. Sci.* **2023**, *24*, 2420–2434. (In Chinese)
27. Zhang, X.; Shi, W.; Lv, Z. Uncertainty assessment in multitemporal land use/cover mapping with classification system semantic heterogeneity. *Remote Sens.* **2019**, *11*, 2509. [CrossRef]
28. Gao, Y.; Liu, L.; Zhang, X.; Chen, X.; Mi, J.; Xie, S. Consistency analysis and accuracy assessment of three global 30-m land-cover products over the European Union using the LUCAS dataset. *Remote Sens.* **2020**, *12*, 3479. [CrossRef]
29. Verburg, P.; Neumann, K.; Nol, L. Challenges in using land use and land cover data for global change studies. *Glob. Change Biol.* **2011**, *17*, 974–989. [CrossRef]
30. Mas, J.; Kolb, M.; Paegelow, M.; Olmedo, M.T.C.; Houet, T. Inductive pattern-based land use/cover change models: A comparison of four software packages. *Environ. Modell. Softw.* **2014**, *51*, 94–111. [CrossRef]
31. Stehman, S.; Foody, G. Key issues in rigorous accuracy assessment of land cover products. *Remote Sens. Environ.* **2019**, *231*, 111199. [CrossRef]
32. Pengra, B.W.; Stehman, S.V.; Horton, J.A.; Dockter, D.J.; Schroeder, T.A.; Yang, Z.; Cohen, W.B.; Healey, S.P.; Loveland, T.R. Quality control and assessment of interpreter consistency of annual land cover reference data in an operational national monitoring program. *Remote Sens. Environ.* **2020**, *238*, 111261. [CrossRef]
33. Talukdar, S.; Singha, P.; Mahato, S.; Pal, S.; Liou, Y.A.; Rahman, A. Land-use land-cover classification by machine learning classifiers for satellite observations—A review. *Remote Sens.* **2020**, *12*, 1135. [CrossRef]
34. Abdi, A. Land cover and land use classification performance of machine learning algorithms in a boreal landscape using Sentinel-2 data. *GISci. Remote Sens.* **2020**, *57*, 1–20. [CrossRef]
35. Jamali, A. Evaluation and comparison of eight machine learning models in land use/land cover mapping using Landsat 8 OLI: A case study of the northern region of Iran. *SN Appl. Sci.* **2019**, *1*, 1448. [CrossRef]
36. Wang, Z.; Mountrakis, G. Accuracy Assessment of Eleven Medium Resolution Global and Regional Land Cover Land Use Products: A Case Study over the Conterminous United States. *Remote Sens.* **2023**, *15*, 3186. [CrossRef]