



## OPEN ACCESS

## EDITED BY

Sawaid Abbas,  
University of the Punjab, Pakistan

## REVIEWED BY

Chao Chen,  
Suzhou University of Science and  
Technology, China  
Josef Strobl,  
University of Salzburg, Austria

## \*CORRESPONDENCE

Yan Yan  
✉ [yanyan@lynu.edu.cn](mailto:yanyan@lynu.edu.cn)

RECEIVED 09 August 2023

ACCEPTED 30 October 2023

PUBLISHED 20 November 2023

## CITATION

Zhang W, Tian J, Zhang X, Cheng J and  
Yan Y (2023) Which land cover product  
provides the most accurate land use land  
cover map of the Yellow River Basin?  
*Front. Ecol. Evol.* 11:1275054.  
doi: 10.3389/fevo.2023.1275054

## COPYRIGHT

© 2023 Zhang, Tian, Zhang, Cheng and Yan.  
This is an open-access article distributed  
under the terms of the [Creative Commons  
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,  
distribution or reproduction in other  
forums is permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original publication in  
this journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted  
which does not comply with these terms.

# Which land cover product provides the most accurate land use land cover map of the Yellow River Basin?

Weige Zhang<sup>1,2</sup>, Junjie Tian<sup>2</sup>, Xiaohu Zhang<sup>2</sup>, Jinlong Cheng<sup>2</sup>  
and Yan Yan<sup>2\*</sup>

<sup>1</sup>Institute of Management and Business, Kyrgyz National University named after Jusup Balasagyn, Bishkek, Kyrgyzstan, <sup>2</sup>School of Land and Tourism, Luoyang Normal University, Luoyang, China

Precise land use land cover (LULC) data are essential for understanding the landscape structure and spatial pattern of land use/cover in the Yellow River Basin (YRB) to regulate scientific and rational territorial spatial planning and support sustainable development. However, differences in the multiple sets of LULC products in portraying the land composition of the YRB limit our understanding of the land cover composition in this region. To address this issue, this study chose five sets of open and high spatiotemporal LULC data in 2020, namely, CLCD, LSV10, ESRI10, CLC\_FCS30, and Globeland30, to evaluate the accuracy and consistency of classification in the YRB. Our results show that: (1) The LULC composition of the YRB in 2020 was mapped consistently by the five datasets. Grasslands, croplands, and woodlands constitute the major LULC types, accounting for 96% of the total area of the study area. (2) The correlation coefficients of the LULC types of any two of the five datasets ranged from 0.926 to 0.998, showing high land compositional consistency. However, among the five datasets, there were considerable differences in the areas of a single LULC type. (3) The classification consistencies of croplands, woodlands, grasslands, and water bodies were higher than 60% in any two datasets. The spatial consistencies of grasslands, croplands, and woodlands were higher than those of other LULC types. An area with better consistency can reach more than 50% of the average area of the corresponding land types, but grasslands were mixed with other LULC types in ESRI10 and GLC\_FCS30. (4) According to the accuracy assessments, LSV10 data have the highest overall classification accuracy, 79.32%, and the classification accuracy of major land types is also higher than 70%; GLC\_FCS30 data have the lowest overall accuracy, 70.14%. Based on these results, LSV10 can more accurately demonstrate LULC than the other four datasets. This study can be used as a reference for selecting land cover data, and it also highlights that the necessary assessments of consistency and accuracy are essential when conducting land use/cover change studies in a specific region.

## KEYWORDS

land use land cover datasets, accuracy assessment, classification consistency, spatial consistency, Yellow River Basin (YRB)

## 1 Introduction

The Yellow River Basin (YRB) is a vital ecological barrier and an important region for economic development in China. It is essential for national environmental security and economic construction (Xi, 2019). However, the fragile environment and the relatively crude mode of economic development, mainly in the form of energy-dependent industries, have put enormous pressure on the ecological environment in the basin. In the past few decades, the increasing population and intensity of human activities, combined with climate change, have resulted in degradation in the environmental functions of the YRB (Jin, 2019; Jin et al., 2020). In response to environmental degradation in the basin, Chinese governments at every level have taken a series of actions to improve the eco-environment, for example, implementing several major ecological restoration projects, including large-scale sand control and sand treatment, promoting comprehensive environmental improvement in beach areas and protecting and restoring the Yellow River Delta wetlands (China, 2021). These ecological restoration projects have had a profound and dramatic impact on the land cover of the YRB, driving complex spatiotemporal land cover changes (Zhao et al., 2018; Zhang et al., 2021) that have affected the quality of the ecological environment of the basin (Zhang et al., 2014). It has been stated that the land cover change in the YRB from 1986 to 2018 was dominated by an increase in orchards and terraces at the cost of decrease in woodlands, grasslands, and croplands (Ji et al., 2021). In addition, vegetation coverage significantly increased owing to the implementation of major ecological restoration projects such as the Yellow River Basin Ecological Protection Plan (Ji et al., 2021). Thus, accurate and detailed LULC data are helpful not only in understanding the LULC changes induced by human activities in the basin and their accompanying ecological and environmental effects but also in understanding the results of major ecological implementation. Additionally, such data are primary data for studies on biodiversity, sustainable resource use and soil carbon stocks (Watson et al., 2001; Yang et al., 2023).

Over recent decades, many institutions and scholars have developed multiple sets of LULC products with multi-spatiotemporal resolutions that scanning of the YRB. However, there are some disadvantages in the LULC datasets generated in the early ages. First, the spatial resolution is 300–1000 m, which is too coarse to map LULC types with smaller patches. Examples include the UMD Land Cover Classification 1998 produced by the University of Maryland based on AVHRR images (Hansen et al., 2000) and the Land Cover Type/Dynamics data developed by Boston University based on MODIS data (Friedl et al., 2002). Second, the temporal resolution is 5 to 10 years, which is an interval that has failed to draw the gradual changes in LULC over time. The development of remote sensing and information technology has promoted LULC products toward a high time-frequency and high spatial resolution and released several sets of spatiotemporal data, such as the global land cover data GLC\_FCS30 (with a 30 m resolution, from 1985 to 2020, 5-year interval) (Zhang et al., 2020), Globeland30 (with a 30 m spatial resolution, from 2000–2020, 10-year interval) (Chen et al., 2014b), and the annual Chinese land use data CLCD released by Wuhan University (1990–2021) (Yang and Huang, 2021). Furthermore, the

spatial resolution of the specified LULC products can up to 10 m or even 1 m, such as the 10 m global land use data LSV10 (2020) released by the European Space Agency (ESA) (Zanaga et al., 2021), the 10 m global land use data released by ESRI (2020) (Karra et al., 2021), and the 1 m resolution China land use data SinoLC-1 (Supplementary Table S1) released by Wuhan University (Li et al., 2023). These data provide a solid data basis for understanding spatiotemporal LULC changes at multiple scales, including in the YRB, and associated ecological environmental effects.

However, the remote sensing data sources of the current LULC datasets differ in terms of the sensor type, the time the data were acquired, and classification schemes and systems. These differences have resulted in regional differences in classification accuracy among the different LULC datasets (Hansen et al., 2000; Friedl et al., 2002; Chen et al., 2014a; Zhang et al., 2020; Karra et al., 2021; Yang and Huang, 2021; Zanaga et al., 2021; Li et al., 2023). Additionally, they have led to considerable discrepancies in drawing the land composition and temporal changes in LULC in the YRB. For instance, the Chinese Land Use/Cover (CNLUCC) data developed by the Institute of Geographical Sciences and Resources of the Chinese Academy of Sciences show that the LULC of the YRB was dominated by grasslands and croplands during 1980 and 2020. Croplands, woodlands, and water bodies declined until 2000 but increased after 2000. In contrast, the continuous decreasing tendency of grasslands, the primary LULC type in the basin, was halted until 2015 (Ji et al., 2021). Based on GLC\_FCS30, the LULC composition of the YRB and the tendency of major LULC types, such as croplands, woodlands, and water bodies, were similar to those in CNLUCC. However, the temporal pattern of grasslands from 1985 to 2020 differed from that in CNLUCC. That is, the decline in grasslands reached a turning point in 2000, after which an increasing tendency was found (Xu et al., 2018). When choosing the CLCD dataset, the area of different LULC types in the YRB, as well as the temporal pattern of LULC changes, showed enormous differences from CNLUCC and GLC\_FCS30 (Supplementary Figure 1).

In summary, uncertainties exist in the quantitative study of the LULC composition of the YRB, making a comprehensive analysis of the similarities and differences between different data products necessary. Further, there is an urgent need to assess the classification accuracy of the LULC data of the YRB to select the optimal data. This work will be significant for improving the reliability of the results of eco-environmental assessment based on LULC data but also could support the formulation of scientific and reasonable territorial spatial planning based on reliable LULC data to support the sustainable development of the basin. In this study, we evaluated five open access LULC datasets in 2020 with a higher spatial resolution, namely, (1) ESA WorldCover 10 m – v100 (LSV10), with a 10 m spatial resolution and provided by the European Space Agency; (2) ESRI10, with a 10 m resolution and produced by the ESRI company; (3) the Global Geo-information Public Product (Globeland30), produced by the National Geomatics Center of China; (4) Global 30 m land-cover dynamic monitoring products, with a fine classification system (GLC\_FCS30) and produced by the Aerospace Information Research Institute, Chinese Academy of Sciences; and (5) China's Land-use/cover

Datasets (CLCD), annual data with a 30 m spatial resolution released by Wuhan University. Three indicators, i.e., LULC composition similarity, classification consistency, and accuracy assessment, were selected for the comparative analysis of multisource LULC data in the YRB. Our objectives were to (1) present a reliable map of the LULC composition and spatial patterns of the YRB, (2) verify the classification and spatial consistency of five prevailing high-resolution LULC datasets in the YRB, and (3) assess the classification accuracy of the five LULC datasets in the YRB to provide a basis for data selection for further spatiotemporal changes in LULC research on the YRB and for the simulation of different scenarios in the future.

## 2 Data and methodology

### 2.1 Study area

The Yellow River is the second largest river in China. It originates at the northern foot of the Bayan Har Mountains on the Qinghai-Tibet Plateau and flows into the Bohai Sea in Kenli County, Shandong Province. It flows from west to east through 9 provinces: Qinghai, Sichuan, Gansu, Ningxia, Inner Mongolia, Shaanxi, Shanxi, Henan and Shandong. Additionally, it has a total length of 5464 km and forms a total basin area of 79,500 km<sup>2</sup>, which is the significant ecological barrier of China, spanning from 32°–42°N to 96°–119°E (Figure 1). The terrain of the basin is high in the west and low in the east, spanning three major terrain steps in China. From west to east, it spans four different vegetation types: Qinghai-Tibet Plateau vegetation, desert, grassland, and deciduous broad-leaved forest. The soil in the study area is dominated by primary loess and secondary loess (Figure 1). Most of the YRB is classified as arid, semi-arid and semi-humid climate types, with an

annual average precipitation of 466.6 mm and a mean temperature of 9.4°C (Wang et al., 2021).

The YRB is a substantial grain production base and a vital region for economic development in China. The regional GDP of the basin is approximately CNY 23.9 trillion, accounting for 26.5% of the country's total GDP. The YRB accounts for 21.8% of the country's total economic output with 27.3% of the national area. The basin has a better agricultural and livestock base, including major agricultural production areas such as the Hetao-Plain and the North China Plain (Figure 1), which account for approximately one-third of the country's grain and meat production. Rich in energy resources and metal reserves, it is a vital energy, chemical, raw material, and essential industrial base in China.

### 2.2 Main data sources and preprocessing

Five prevailing and current (2020) LULC datasets, i.e., LSV10, ESRI10, GLC\_FCS30, Globeland30, and CLCD, were used in this study (Table 1). The five data sets were all produced at pixel scale but differed in their specific classification strategies. Among them, LSV10 was generated based on a decision tree classifier; ESRI10 was produced using a deep learning method; GLC\_FCS30 and CLCD were generated using a random forest classification model, while Globeland30 was based on a Pixel-object-knowledge-based approach (POK) method to accomplish the classification (Supplementary Text 1). These datasets had a higher spatial resolution (10–30 m) and officially declared overall accuracy (at the global and/or national scale) that ranged from 72.27% (GLC\_FCS30) to 85.72% (Globeland30). Detailed information on these five datasets is summarized in Table 1. We generated the boundary of the YRB from a DEM with a 30 m resolution using the hydrologic analysis method (Khan et al., 2014; Tang, 2019).

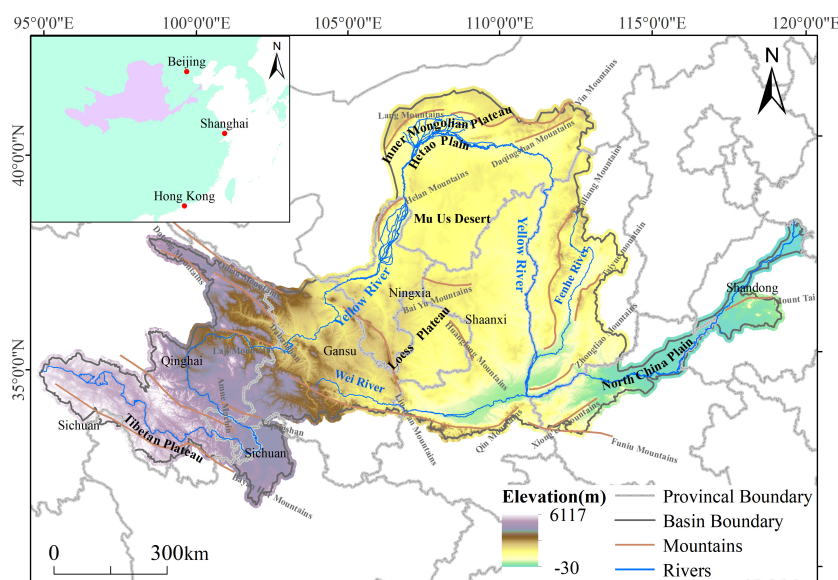


FIGURE 1  
Location of the study area.

TABLE 1 Main information of the five LULC datasets.

Datasets	Producer	Data sources	Images covering the YRB	Classification system	Number of LULC types	Classification method	Resolution (m)	Overall accuracy (%)	Data available from
LSV10	European Space Agency (ESA)	Sentinel-2 multispectral images, Sentinel-1 SAR images, DEM data	22	LCCS	11	Pixel-based Decision tree	10	74.40	<a href="https://viewer.esa-worldcover.org/worldcover/">https://viewer.esa-worldcover.org/worldcover/</a> Accessed on 4 October 2023
ESRI10	ESRI Corporation	Sentinel-2 images	8	/	10	Pixel-based Deep learning	10	85.96	<a href="https://esri.maps.arcgis.com/home/index.html">https://esri.maps.arcgis.com/home/index.html</a> Accessed on 4 October 2023
GLC_FCS30	Institute of Air and Space Information Innovation, Chinese Academy of Sciences	Landsat images, Sentinel-1 SAR images, DEM data	11	GLC_FCS30-2020 Classification System	30	Pixel-based random forest	30	72.27	<a href="https://zenodo.org/record/3986872">https://zenodo.org/record/3986872</a> Accessed on 4 October 2023
Globeland30	China National Basic Geographic Information Center	Landsat8-OLI, GF-1 multispectral imagery	11	/	10	Pixel-object-knowledge-based approach (POK)	30	85.72	<a href="http://globeland30.org/">http://globeland30.org/</a> Accessed on 9 August 2023
CLCD	Institute of Remote Sensing, Wuhan University	Landsat imagery, SRTM data	10	/	9	Pixel-based random forest	30	79.31	<a href="https://zenodo.org/record/5816591">https://zenodo.org/record/5816591</a> Accessed on 4 October 2023

Different academic institutions generate LULC datasets and thus differ in remote sensing sensor types, data sources, imaging times, spatial resolutions, classification systems, and classification methods. Such differences do not facilitate subsequent analysis of spatial consistency. This is primarily because different projections lead to the non-overlapping display of pixels with identical positions. In addition, disparities in spatial resolution result in varying total pixel counts among the five LULC datasets. Furthermore, the classification accuracy of higher-resolution remote sensing data may surpass that of lower-resolution remote sensing data owing to the advantages offered by remote sensing data sources (Sentinel-1,2 versus Landsat images, i.e., 10m versus 30m spatial resolution) (Table 1), thereby impacting the comparability of classification accuracy across different datasets. Therefore, preprocessing was necessary for the subsequent analysis, including image mosaicking and clipping, projection transformation, upscaling, and classification system merging. One note should be made that the global LULC datasets of LSA10, ESRI10, GLC\_FCS30 and Globeland30 were grouped by several regional tiles, except for the CLCD dataset. Therefore, first, several images that covered the YRB from each dataset should be mosaiced to acquire whole LULC images of the YRB. The number of images spanning the study area can be found in Table 1. Then, we transformed the projection to Krasovsky\_1940\_Albers, of which the central meridian is 105°E, and the longitude ranges from 27°N to 47°N. Third, the mosaiced images of each dataset were extracted by the boundary of YRB. Subsequently, the datasets with a finer resolution of LSV10 and ESRI10 were upscaled to 30 m using the resampling method to make the resolution consistent with the GLC\_FCS30, Globeland30, and CLCD datasets (Wang et al., 2022). Finally, we reclassified the classification system of the five datasets based on the comprehensive understanding of the actual LULC represented by each classification code and corresponding description (Table 2), and we merged and unified the classification systems with the support of previous research (Table 3) (Hu et al., 2015; Dai et al., 2017; Yang et al., 2017).

It should be noted that the shrubland in ESRI10 was removed from its classification system. Hence, we combined the shrubland into woodland to maintain the consistency of the classification system of the five datasets. Additionally, some missing pixels were found in the five obtained datasets after preprocessing, resulting in differences from the original data. However, the missing pixels could barely influence the research due to their small size. We set these pixels null; thus, they are not part of the subsequent analysis.

### 2.3 Methodologies

Three indicators proposed by Hu et al. (2015), i.e., LULC composition similarity, classification and spatial consistency and accuracy assessment, were used to evaluate the classification consistency and accuracy of the five datasets in the YRB (Hu et al., 2015).

LULC composition similarity assesses the consistency of the land composition of different datasets based on the area of different LULC types based on the principle of correlation analysis. The area of different LULC types could be calculated based on the total pixel numbers recognized as a land type and the spatial resolution. Then,

TABLE 2 Classification system of the five LUCC datasets.

LSV10		ESRI10		Globelnd30		GLC_FCS30				CLCD	
Code	LULC type	Code	LULC type	Code	LULC type	Code	LULC type	Code	LULC type	Code	LULC type
10	Tree cover	1	Water	10	Cropland	10	Rain-fed cropland	120	Shrubland	1	Cropland
20	Shrubland	2	Trees	20	Woodland	11	Herbaceous cover	121	Evergreen shrubland	2	Woodland
30	Grassland	4	Flooded vegetation	30	Grassland	12	Tree or shrub cover (orchard)	122	Deciduous shrubland	3	Shrub
40	Cropland	5	Crops	40	Shrubland	20	Irrigated cropland	130	Grassland	4	Grassland
50	Built-up	7	Built area	50	Wetland	50	Evergreen broadleaved woodland	180	Wetlands	5	Water
60	Bare/sparse vegetation	8	Bare ground	60	Water bodies	60	Deciduous broadleaved woodland	190	Impervious surfaces	6	Snow/ice
70	Snow and ice	9	Snow/ice	70	Tundra	61	Closed deciduous broadleaved woodland	140	Lichens and mosses	7	Barren
80	Permanent water bodies	10	Clouds	80	Artificial surfaces	62	Open deciduous broadleaved woodland	150	Sparse vegetation	8	Impervious
90	Herbaceous wetland	11	Rangeland	90	Bare land	70	Evergreen needle-leaved woodland	152	Sparse shrubland	9	Wetland
95	Mangroves			100	Permanent snow and ice	71	Closed evergreen needle-leaved woodland	153	Sparse herbaceous cover		
100	Moss and lichen					72	Open evergreen needle-leaved woodland	200	Bare areas		
						80	Deciduous needle-leaved woodland	201	Consolidated bare areas		
						81	Closed deciduous needle-leaved woodland	202	Unconsolidated bare areas		
						82	Open deciduous needle-leaved woodland	210	Water body		
						90	Mixed-leaf woodland	220	Permanent ice and snow		

the LULC composition similarity could be calculated based on the area of a land type from any two datasets (Equation 1).

$$R_{AB} = \frac{\sum_{k=1}^8 (A_k - \bar{A})(B_k - \bar{B})}{\sqrt{\sum_{k=1}^8 (A_k - \bar{A})^2 \sum_{k=1}^8 (B_k - \bar{B})^2}} \quad (1)$$

where  $R_{AB}$  is the land composition similarity for LULC datasets A and B.  $k$  denotes the LULC type,  $A_k$  and  $B_k$  are the LULC areas of datasets A and B, respectively.  $\bar{A}$  and  $\bar{B}$  are the mean areas of eight LULC types in datasets A and B, respectively.

The essence of composition similarity is the correlation coefficients between LULC areas of any two datasets, which can only quantitatively describe the land composition. It only evaluates the consistency of the

LULC composition of multiple datasets and cannot reveal the classification confusion among different datasets. For this reason, a consistency analysis method was used to carry out consistency analysis regarding the number of pixels and classification confusion by considering their location. The main idea of this classification consistency analysis is to compare any two datasets at the pixel scale. If the same pixel from any two datasets has the same code, it is considered a pure pixel. If not, then it is regarded as a mixed pixel. The numbers of pure and mixed pixels were counted, and category consistency was then analyzed based on Equation 2 and Equation 3.

$$DP_{AB}(k) = \frac{N(kk)}{N(k)} \quad (2)$$

TABLE 3 Correspondence between the five sets of data classification systems and the new classification system in our study.

Code	LULC types	LSV10	ESRI10	GLC_FCS30	Globeland30	CLCD
1	Cropland	40	5	10, 20	10	1
2	Woodland	10	2	51–92	20	2,3
3	Grassland	20, 30, 95, 100	3,6	11, 12, 120–122, 130–150, 152, 153	30, 40	4
4	Wetland	90	4	180	50	9
5	Water body	80	1	210	60	5
6	Construction land	50	7	190	80	8
7	Bare land	60	8	200–202	90	7
8	Snow and ice	70	9	220	100	6

where  $DP_{AB}(k)$  denotes the purity of LULC type  $k$  in datasets A and B,  $N(kk)$  is the total number of pure pixels for LULC type  $a$  in datasets A and B, and  $N(k)$  is the number of LULC type  $k$  in dataset A.

$$DC_{AB}(ab) = \frac{N(ab)}{N(a)} \tag{3}$$

where  $DC_{AB}(ab)$  denotes the degree of category confusion of any two LULC datasets, and  $a$  and  $b$  are the LULC type.  $N(ab)$  is the total number of mixed pixels of datasets A and B, and  $N(a)$  is the pixel number of LULC type  $a$  in dataset A.

Although the classification consistency analysis accounts for the positional information of pixels, the final result is presented in statistical form. Nevertheless, it cannot visualize the spatial consistency and confusion of individual land types. Geospatial analysis allowed the spatial consistency of the five datasets to be portrayed and visualized on maps. Specifically, we binarized the five sets of LULC data and then obtained the correspondence of pixels for these datasets employing spatial overlay analysis. Then, for pixels from the five LULC datasets, it was determined whether they had the same values, and the numbers of these pixels were counted. The spatial consistency of the five LULC datasets was then divided into five levels based on the order of the pixel numbers recognized with the same values.

If the pixels are identified as the same land type in the five datasets, the spatial consistency of this land type is considered entirely consistent (i.e., 100%). Similarly, suppose that a pixel is simultaneously identified as the same land type by four, three, two and only one dataset. In that case, the spatial consistency of a particular land type is considered to be highly consistent (i.e., 80%), basically consistent (i.e., 60%), low consistent (i.e., 40%) and completely inconsistent (i.e., less than 20%). A spatial consistency higher than or equal to 60% is collectively considered good consistency (Equation 4).

$$C_p(k) = \frac{\sum_{L=1}^5 (D_L == k)}{5} \tag{4}$$

where  $C(p)$  denotes the spatial consistency of land type  $k$  at pixel  $p$ .  $D_l$  is the land type of  $k$  identified by dataset  $L$ .

Composition similarity, classification, and spatial consistency were analyzed by cross-referencing the five datasets. However, these

indicators could not present their classification accuracy. To understand which dataset best represents the LULC of the YRB, we used four indicators, i.e., overall classification accuracy (OA), Kappa coefficients, user accuracy (UA) and producer accuracy (PA), to assess the quality of the five datasets (Lyons et al., 2018).

Building upon previous studies (Tilahun, 2015; Zhang et al., 2020), we assessed five datasets using the reference LULC information extracted from Google Earth as our benchmark data. First, we divided the study area into  $0.25^\circ \times 0.25^\circ$  grids and extracted the geometric center points of each grid. A total of 895 points were generated, among which 184 were for croplands, 104 points for woodlands, 499 points for grasslands, 6 points for wetlands, 10 points for water bodies, 32 points for construction land, 59 points for bare land, and 1 point for ice and snow. The area of wetland types in the study area is small (area proportion is less than 1.03%), and less than five fall within the wetland types. In this respect, only two sample points fall within the wetland types of the GLC\_FCS data, and no samples fall within the wetland types in the CLCD data.

The small sample size of wetlands was not conducive to determining accuracy in the GLC\_FCS30 and CLCD data. Therefore, we converted the wetlands of GLC\_FCS30 and CLCD (in raster) to points (4,220,009 points and 573,245 points, respectively). We randomly selected 84 points for GLC\_FCS30 (i.e., 0.002% of total wetlands) and 57 points for CLCD (i.e., 0.01% of total wetlands). In addition, the area of the ice and snow LULC type in the five datasets was smaller than 0.2%, and only one sample point fell within the ice and snow. The influence of the smaller area and the smaller ice and snow sample size on the overall accuracy was negligible. Thus, this sampling point was removed, and the ice and snow LULC type was no longer considered for accuracy validation. Subsequently, we combined the added 141 samples for wetlands with the 894 existing sampling points, giving 1035 sampling points (Figure 2).

Then, the pixel values (i.e., the codes of the LULC types) of the five datasets that corresponded with the center points were extracted, and they could serve as the classified datasets. Third, we loaded the classified datasets into Google Earth (the platform can provide high-resolution remote sensing images with 0.5 m online and no position offset). The actual LULC types were acquired through visual interpretation, and they could serve as a reference

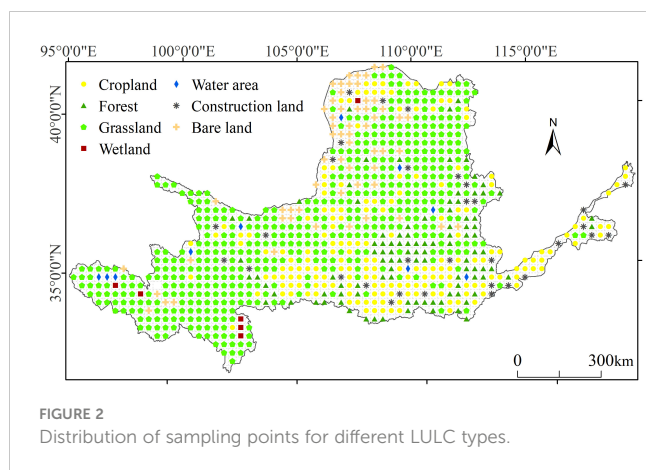


FIGURE 2  
Distribution of sampling points for different LULC types.

dataset. Fourth, we generated the classification confusion matrix based on the classified and reference datasets and further calculated the four-accuracy metrics: OA, the Kappa coefficient, UA, and PA.

### 3 Results

#### 3.1 The composition similarity of land use land cover

The correlation coefficients for the area composition of any two LULC datasets in the YRB were higher than 0.9 (Figure 3). The highest correlation was found in the CLCD/GLC\_FCS30 combination, of which the correlation coefficient reached 0.998. The lowest correlation was between the ESRI10 and Globeland30 datasets, whose correlation coefficient was 0.926 (Figure 3). The results suggest that although produced by different academic institutions and differing in LULC classification systems and methods, a high similarity among the LULC compositions was found in the five datasets (Figure 3), which means that the five datasets demonstrated the LULC consistency of the YRB. That is, the LULC type in the YRB in 2020 was dominated by grasslands, followed by croplands and woodlands, accounting for more than 88% of the total area. Construction land, bare land and water area

were the vital LULC types in the basin, occupying 10.84% of the study area. Wetlands and ice and snow have a tiny distribution, with a total area of less than 1% of the YRB (Table 4).

However, the areas of different LULC types given by the five datasets showed enormous differences. The area proportion of grasslands, the major LULC type, ranged from 45.22–64.42% in the five datasets (Table 4) and was widely distributed in the basin (Supplementary Figures 2A–E). The most considerable discrepancy in grasslands was found between ESRI10 and Globeland30, with a variance of nearly 1/5 (or 155,275.99 km<sup>2</sup>) of the total basin. Croplands covered 18.70–30.83% of the basin in the five datasets, constituting the second largest LULC type (Table 4), mainly in the central and eastern flatter areas (Supplementary Figures 3A–E). Similar to grasslands, significant differences in croplands were found among the five datasets. In particular, Globeland30 data identified a high proportion of croplands, 30.83%, which was much higher than the four LULC datasets. As described by the five LULC datasets, woodlands, covering 7.20–13.31% of the study area (Table 4), were the third largest LULC component in the basin, concentrated in the middle reaches of the Yellow River (Supplementary Figures 4A–E). The area of bare land in the five datasets ranged from 2.58–14.70%, mainly in the western and northern parts of the basin (Supplementary Figure 5A–E). A difference in the area of bare land of approximately 5.7 times was found among the five LULC datasets (Table 4).

Despite the smaller area of construction land, water bodies, wetlands, and ice and snow, they constituted the vital LULC composition of the basin. Except for construction land, there are apparent differences among the five datasets in water areas, wetlands, and ice and snow. For instance, wetlands account for 1.03% of the basin in Globeland30, compared to 0.06% in the CLCD dataset, with a difference of up to 17 times (Table 4). In summary, although the LULC composition of the YRB was consistent in the five datasets, there were significant differences in the areas of different LULC types.

#### 3.2 Classification consistency

The five datasets provide a higher classification consistency in the identification of croplands, woodlands, grasslands and water bodies than the other LULC types, with the classification purities of these four LULC types being higher than 60% in the combination of any two datasets (Table 5, Figure 4). Specifically, the highest classification consistencies of croplands (classification purity of 86.98%), woodlands (92.74%), grasslands (89.57%), and water bodies (94.04%) were found in the ESRI10/Globeland30, ESRI10/GLC\_FCS30, ESRI10/LSV10, and ESRI10/CLCD combinations, respectively (Table 5, Figure 4).

In contrast, the classification consistency of construction land, bare land, and ice and snow showed a considerable discrepancy (Table 5, Figure 4). For instance, the purity of construction land was 85.97% in the ESRI10/LSV10 combination and only 39.31% in the GLC\_FCS30/ESRI10 combination, with the remaining groups having a classification consistency ranging from 50% to 86% (Table 5, Figure 4). In particular, the classification consistency of

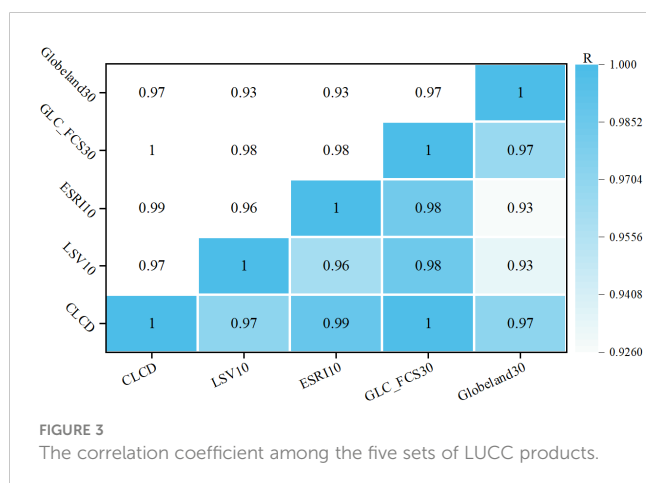


FIGURE 3  
The correlation coefficient among the five sets of LULC products.

TABLE 4 The LULC composition proportions of the five datasets in the YRB (%).

Land use/cover types	LSV10	ESRI10	GLC_FCS30	Globeland30	CLCD
Croplands	19.68	18.70	22.20	30.83	23.62
Woodlands	10.70	7.20	12.90	13.31	11.92
Grasslands	51.02	64.42	54.56	45.22	57.02
Wetlands	0.37	0.23	0.47	1.03	0.06
Water bodies	0.97	1.11	0.69	0.90	0.79
Construction land	2.55	5.68	2.45	3.86	3.11
Bare land	14.70	2.58	6.70	4.66	3.45
Ice and snow	0.02	0.08	0.02	0.19	0.03

bare land showed enormous differences among the combinations of the five LULC datasets. The highest classification consistency of bare land was found in the LSV10/CLCD combination, with the purity reaching 94.11%. In comparison, the lowest consistency was found in the combination of ESRI10/LSV10 combination, with a purity of only 12.14%. The classification purities ranged from 23–78% in the remaining LULC data groups.

The consistency of wetlands was the worst among the five LULC datasets. Only three combinations, i.e., Globeland30/GLCD, LSV/CLCD, and Globeland30/LSV, demonstrated higher classification consistencies with purities over 61.33% (Figure 4). In this respect, the highest purity, 96.83%, was found in Globeland30/GLCD. The rest of the dataset combinations all had low classification consistencies, with the lowest (0.01%) occurring in the Globeland30/GLC\_FCS30 combination, showing a vast difference between the Globeland30/GLCD and Globeland30/GLC\_FCS30 combinations (Table 5, Figure 4).

It is particularly noteworthy that despite the higher classification consistency of grasslands, woodlands and croplands, the classification confusion of grasslands in the five datasets should not be ignored (Table 5). Grasslands were confused with bare land (confusion degree of 67.86% in ESRI10/LSV10), woodlands (confusion degree of 67.86% in ESRI10/LSV10), and croplands (confusion degree of 30.74% GLC\_FCS30/ESRI10) (Table 5). In addition, there was a high degree of classification confusion found in wetlands. The highest rate of confusion was in the ESRI10/CLCD combination, in which the confusion degree reached 87.71%. In contrast, the confusion of the other LULC types in the five datasets was not as significant as that of grasslands in the five datasets.

### 3.3 Spatial consistency

Considering the LULC composition of the YRB, we chose four major LULC types, i.e., croplands, woodlands, grasslands, and bare land (the total area of which accounts for more than 95% of the basin) to assess the spatial consistencies of LULC in the five datasets. Similar to classification consistency, the spatial consistencies of these four LULC types were higher (Figure 5).

The spatial consistency of croplands showed a higher level (Figure 5A). The region with better consistency could occupy 53.53%

of the average cropland area in the five datasets. In comparison, the areas with complete inconsistency were 29.92% (Figure 5A). The areas that were entirely consistent in croplands were mainly distributed in the Hetao-Plain within Inner Mongolia, the southern part of the basin, and Henan and Shandong Provinces, the traditional agricultural areas in China. The areas with low or complete inconsistency of croplands were concentrated in the Ordos Plateau, the Mu Us Desert, and the southern and southwestern margins of the Loess Plateau (Figure 5B).

The areas with better consistency of woodlands accounted for more than 50% of the average woodland area in the five datasets but were lower than that of grasslands, showing a lower spatial consistency (Figure 5B). Furthermore, the proportion of the completely inconsistent area increased to 36.15% of the mean woodlands in the five datasets. Woodlands with high spatial consistency were in the mountainous regions distributed in the southeastern parts of the YRB, while less spatially consistent woodlands were in the central Loess Plateau and the northern and western parts of the basin (Figure 5B).

The spatial consistency of grasslands was the highest among the four types (Figure 5C). The area with better spatial consistency (consistency value was higher than 60%) accounted for 71.84% of the average grassland area in the five datasets and was concentrated in the western highlands and the central Loess Plateau. The region with complete inconsistency covered only 16.16% of the mean grassland area and was mainly distributed in the Hetao-Plain of Inner Mongolia and the lower reaches of the Yellow River in the southwestern parts of the basin (Figure 5C).

Bare land showed a lower spatial consistency than woodlands in the five LULC datasets. The areas with high spatial consistency were only 22.48% of the mean bare land of the five datasets, mainly in the west and south of the Hetao-Plain (Figures 1, 5D). In contrast, the area of poor spatial consistency for bare land accounts for up to 57.36% of the average bare land area of the five datasets. This area is mainly found north of the Liu-p'an Mountains, south of the Helan Mountains, on the Inner Mongolia Plateau and northwest of the Ordos Plateau (Figures 1, 5D).

At the basin scale, 45.96% of the basin was entirely consistent (i.e., defined with the same LULC types by the five datasets), 28.14% of the basin was highly consistent (i.e., illustrated with the same LULC types by the four datasets), 21.49% had better consistency (i.e., defined with the same LULC types by the three datasets), 4.63%



TABLE 5 Classification consistency of LULC types in the five datasets in the Yellow River Basin.

LULC datasets	CLCD								
	LULC types	Cropland	Woodland	Grassland	Wetland	Water body	Construction land	Bare land	Ice and snow
LSV10	Cropland	69.91	0.41	4.48	0.00	2.78	17.16	0.11	0.00
	Woodland	3.69	71.76	2.07	0.00	0.16	2.92	0.20	0.00
	Grassland	15.37	26.96	77.08	8.17	0.56	2.79	3.65	18.10
	Wetland	0.19	0.21	0.39	91.53	0.98	0.14	0.01	0.00
	Water body	0.21	0.06	0.18	0.26	91.47	2.22	0.62	0.37
	Construction land	1.98	0.04	0.79	0.00	0.06	50.73	1.26	0.00
	Bare land	8.65	0.55	15.01	0.03	3.97	24.04	94.11	27.75
	Ice and snow	0.00	0.00	0.00	0.00	0.02	0.00	0.04	53.78
ESR110	Cropland	61.43	1.84	6.13	4.18	0.75	14.11	0.63	0.00
	Woodland	1.15	55.98	0.45	0.24	0.01	0.12	0.00	0.00
	Grassland	26.29	41.61	89.34	87.71	0.92	8.52	57.25	1.11
	Wetland	0.09	0.12	0.31	7.16	1.45	0.04	0.03	0.01
	Water body	0.55	0.10	0.19	0.49	94.04	3.50	0.45	0.13
	Construction land	10.21	0.26	1.62	0.00	0.17	72.36	1.83	0.00
	Bare land	0.28	0.04	1.89	0.22	2.57	1.35	39.47	37.12
	Ice and snow	0.00	0.04	0.07	0.00	0.08	0.00	0.34	61.63
GLC_FCS30	Cropland	66.18	1.69	10.02	0.07	2.53	19.79	0.43	0.02
	Woodland	1.93	84.32	3.98	18.58	0.45	0.49	2.87	0.26
	Grassland	28.51	13.83	78.87	81.29	2.18	17.53	16.90	14.19
	Wetland	0.34	0.08	0.17	0.01	18.08	3.26	1.11	3.36
	Water body	0.07	0.03	0.06	0.05	76.03	0.70	0.28	3.03
	Construction land	1.95	0.02	0.42	0.00	0.25	55.07	0.90	0.00
	Bare land	1.01	0.01	6.46	0.00	0.47	3.17	77.36	23.81
	Ice and snow	0.00	0.00	0.00	0.00	0.01	0.00	0.15	55.31
Globeland30	Cropland	82.32	5.96	17.04	0.73	7.59	27.44	1.29	0.00
	Woodland	3.69	78.21	5.37	0.03	0.51	0.89	0.76	0.05
	Grassland	8.55	15.29	70.47	2.23	3.89	5.58	28.48	3.10
	Wetland	0.27	0.26	1.33	96.83	6.42	1.46	0.42	0.03
	Water body	0.29	0.07	0.17	0.18	80.50	2.19	0.58	0.04
	Construction land	4.55	0.08	1.28	0.00	0.48	61.75	3.38	0.00
	Bare land	0.32	0.11	4.11	0.00	0.55	0.69	63.82	22.38
	Ice and snow	0.00	0.01	0.22	0.00	0.06	0.00	1.27	74.40

(Continued)

TABLE 5 Continued

LULC datasets	LSV10								
	LULC types	Cropland	Woodland	Grassland	Wetland	Water body	Construction land	Bare land	Ice and snow
ESRI10	Cropland	67.60	5.14	6.64	4.98	1.53	5.18	8.77	0.00
	Woodland	0.16	59.83	1.47	0.24	0.01	0.06	0.13	0.00
	Grassland	24.00	33.06	89.57	70.00	2.51	7.86	67.86	0.21
	Wetland	0.02	0.02	0.23	19.54	2.46	0.01	0.06	0.00
	Water body	0.37	0.11	0.06	4.45	88.23	0.23	0.84	0.01
	Construction land	7.62	1.78	0.60	0.22	0.70	85.97	10.10	0.02
	Bare land	0.23	0.05	1.33	0.54	4.54	0.70	12.14	17.27
	Ice and snow	0.00	0.02	0.09	0.04	0.02	0.00	0.09	82.50
GLC_FCS30	Cropland	69.61	6.63	10.14	4.96	3.51	16.20	14.68	0.03
	Woodland	0.43	79.62	7.14	8.86	0.70	0.38	4.14	0.19
	Grassland	27.25	13.12	81.06	77.72	6.70	16.08	38.59	3.61
	Wetland	0.21	0.04	0.05	6.30	19.86	0.51	1.15	3.23
	Water body	0.07	0.00	0.01	1.81	66.37	0.02	0.14	1.46
	Construction land	1.23	0.49	0.05	0.08	0.79	63.41	3.46	0.00
	Bare land	1.21	0.09	1.54	0.26	2.04	3.39	37.80	16.20
	Ice and snow	0.00	0.00	0.01	0.01	0.02	0.00	0.04	75.27
Globeland30	Cropland	86.98	10.16	17.35	8.06	9.36	21.36	21.16	0.00
	Woodland	1.36	74.27	9.33	0.64	0.62	0.72	2.09	0.06
	Grassland	6.94	14.06	69.34	26.99	6.52	5.45	45.34	2.26
	Wetland	0.21	0.04	1.19	61.33	8.27	0.12	0.44	0.00
	Water body	0.25	0.05	0.07	2.44	73.53	0.17	0.55	0.00
	Construction land	3.87	0.97	0.38	0.18	1.08	71.79	6.52	0.04
	Bare land	0.39	0.46	2.06	0.34	0.61	0.38	23.58	15.07
	Ice and snow	0.00	0.00	0.26	0.01	0.01	0.00	0.32	82.56
LULC datasets	ESRI10								
	LULC types	Cropland	Woodland	Grassland	Wetland	Water body	Construction land	Bare land	Ice and snow
GLC_FCS30	Cropland	65.58	2.48	11.54	1.69	9.06	38.06	2.42	1.75
	Woodland	1.50	92.74	9.06	6.07	1.02	0.73	1.09	17.39
	Grassland	30.74	4.77	71.52	75.99	9.41	18.71	39.30	53.75
	Wetland	0.27	0.00	0.10	8.78	19.25	0.44	3.65	0.60
	Water body	0.03	0.00	0.01	6.48	58.43	0.02	0.30	1.03
	Construction land	0.68	0.01	0.10	0.07	1.14	39.31	0.60	0.00
	Bare land	1.20	0.00	7.68	0.90	1.67	2.73	52.36	6.10
	Ice and snow	0.00	0.00	0.00	0.01	0.03	0.00	0.28	19.38

(Continued)

TABLE 5 Continued

LULC datasets	ESRI10								
	LULC types	Cropland	Woodland	Grassland	Wetland	Water body	Construction land	Bare land	Ice and snow
Globeland30	Cropland	83.77	5.07	19.03	2.18	17.24	39.38	4.33	0.41
	Woodland	2.38	87.43	10.01	0.25	1.37	1.33	0.97	4.43
	Grassland	11.09	7.31	63.86	42.01	7.60	5.08	37.71	48.77
	Wetland	0.19	0.02	1.19	49.32	6.66	0.06	1.42	0.05
	Water body	0.21	0.02	0.10	5.50	65.06	0.17	1.68	0.10
	Construction land	1.93	0.05	0.55	0.14	1.27	53.76	2.93	0.01
	Bare land	0.44	0.11	5.09	0.60	0.78	0.21	48.65	14.63
	Ice and snow	0.00	0.00	0.17	0.00	0.00	0.00	2.33	31.61
LULC datasets	GLC_FCS30								
	LULC types	Cropland	Woodland	Grassland	Wetland	Water body	Construction land	Bare land	Ice and snow
Globeland30	Cropland	55.39	2.82	37.68	0.44	0.12	1.69	1.86	0.00
	Woodland	5.32	68.85	24.90	0.07	0.03	0.10	0.72	0.00
	Grassland	7.02	5.85	79.70	0.12	0.07	0.11	7.12	0.00
	Wetland	4.42	3.09	80.13	6.90	3.51	0.47	1.48	0.00
	Water body	6.61	0.77	9.36	16.35	63.69	0.99	2.22	0.01
	Construction land	25.56	0.70	19.45	0.67	0.09	48.02	5.50	0.00
	Bare land	3.25	3.14	38.23	0.47	0.09	0.04	54.65	0.14
	Ice and snow	0.90	5.71	76.06	0.99	0.51	0.00	7.87	7.97

had low consistency (i.e., defined with the same LULC types by the two datasets), and only 0.05% was completely inconsistent (i.e., only occurred in one dataset but not consistent with any of the other four datasets) (Figure 6). The regions with high consistency included the

Qinghai-Tibet Plateau, the North China Plain and the Hetao-Plain, while the Loess Plain and the central part of the YRB showed lower classification spatial consistency in the five LULC datasets (Figures 1, 6). Approximately 73.83% of the land cover

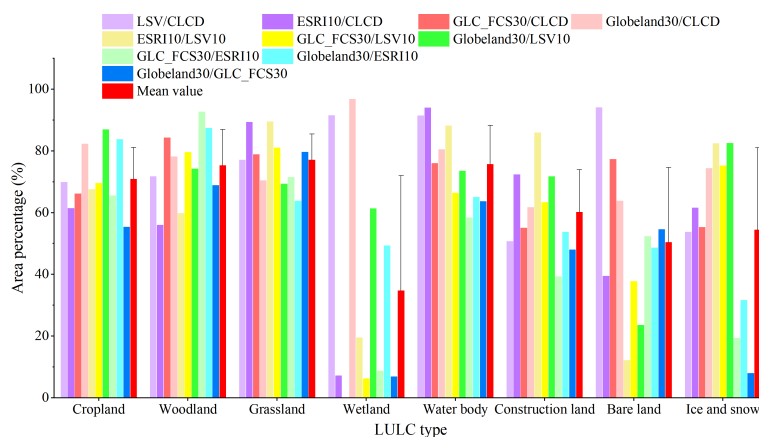
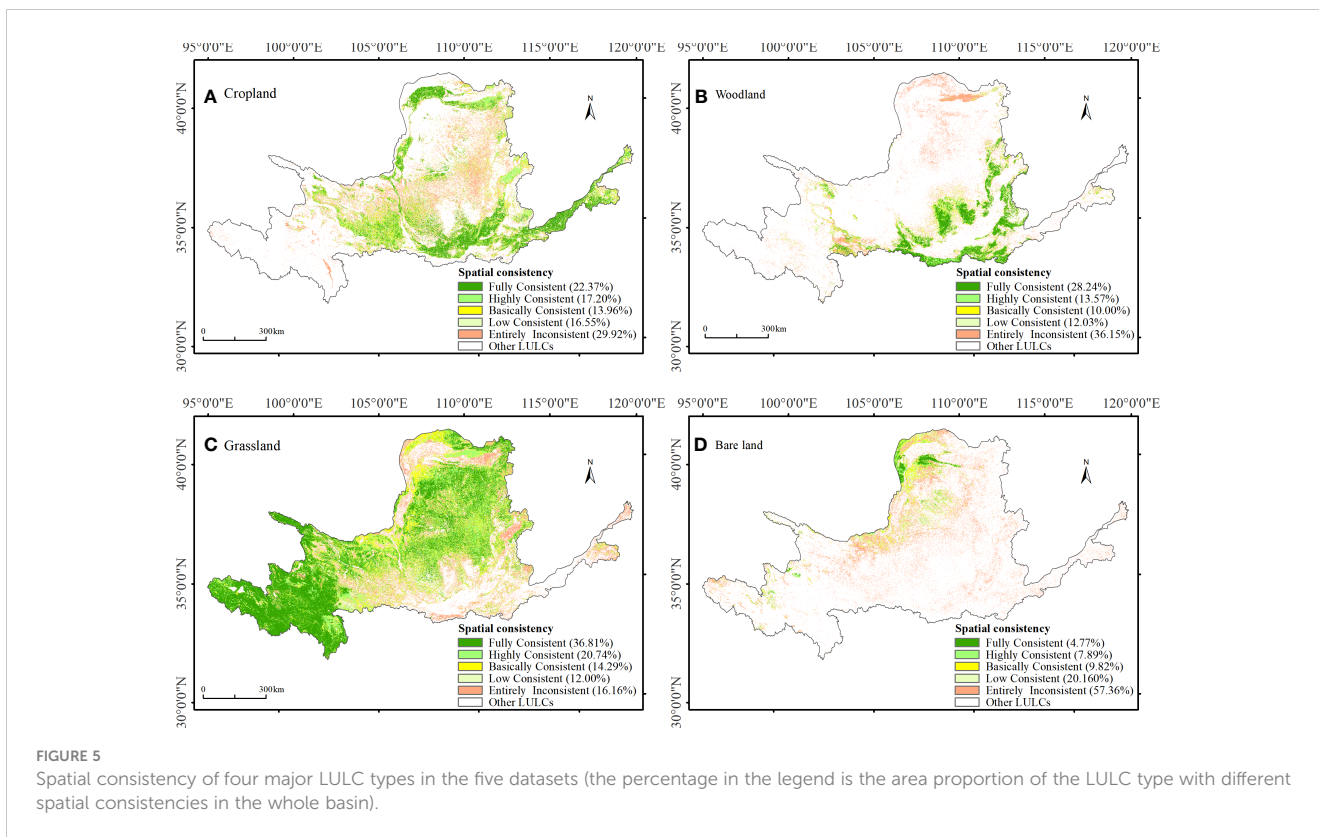


FIGURE 4 The overall classification consistency of LULC types among five datasets.



information in the Yellow River Basin was credible, while 26.17% of the regions were less spatially consistent and credible.

### 3.4 Accuracy assessments

The overall accuracy (OA) of the five LUCL datasets in the YRB exceeded 70%. However, the classification accuracy of all datasets in the YRB was below the officially declared accuracy, except for LSV10 (Table 1). Nevertheless, LSV10 showed the best performance in the LULC classification of the YRB among the five datasets, with an overall classification accuracy of 79.32% (Table 6, Supplementary Table 2),

which is in contrast to GLC\_FCS30, with an overall accuracy of 70.14%, which is the lowest among the five datasets (Table 6, Supplementary Table 2). The five LULC datasets have higher accuracy for croplands, woodlands, water bodies, and grasslands (of which the producer and user accuracy ranged from 60 to 90%), moderate accuracy for construction land and wetlands (with the two kinds of accuracy ranging from 50 to 70%), and lower accuracy for bare land (with the two kinds of accuracy ranging from 30 to 60%) (Table 6, Supplementary Table 2).

LSV10 has the best performance in describing the LULC in the YRB. The producer accuracy (PA) and user accuracy (UA) of the LULC types are generally higher than 74%, except for the lower PA of construction land (66.67%) and wetlands (63.24%) and the lower UA of bare land (36.97%) (Table 6, Supplementary Table 2).

In ESRI10, even though the PA of woodlands and wetlands exceeds 90%, enormous differences were found between the PA and UA of the remaining LULC types in this dataset. For instance, the UA of wetlands in this dataset reached 100%, but the PA was only 8.82%. Similar situations were found in woodlands and wetlands in ESRI10 (Table 6, Supplementary Table 2).

In Globeland30, grasslands and woodlands were classified better than other LULC types (Table 6, Supplementary Table 2). The UA of grassland reached 87.76%, and the PA of woodland was 83.65%. In addition, the UA of croplands was higher than 90% but differed considerably from the UA (59.66%) (Table 6, Supplementary Table 2). This dataset had general accuracy in identifying other LULC types.

For GLC\_FCS30, this dataset had the highest classification accuracy for grasslands, with UA and PA values of 77.64% and 83.14%, respectively. Construction land in GLC\_FCS30 had a higher

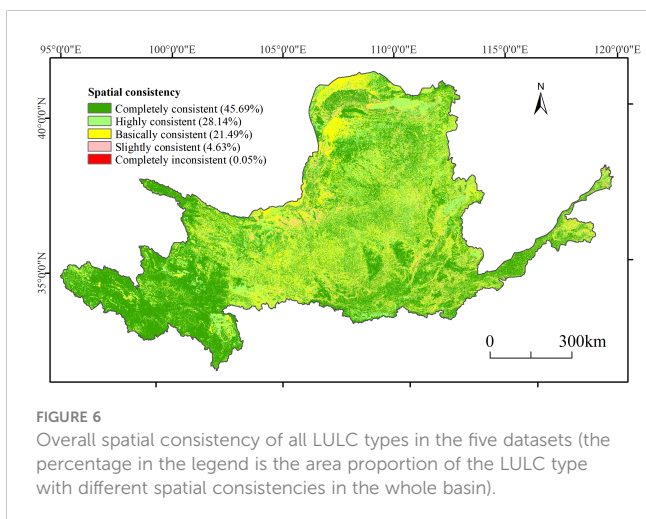


TABLE 6 Accuracy assessments of the five LULC datasets in the YRB (%).

Datasets	LSV10		ESRI10		Globeland30		GLC_FCS30		CLCD	
	UA (%)	PA (%)	UA (%)	PA (%)	UA (%)	PA (%)	UA (%)	PA (%)	UA (%)	PA (%)
Croplands	87.93	80.95	73.23	76.72	59.66	93.12	68.82	67.72	68.47	80.42
Woodlands	88.76	75.96	94.83	52.88	70.73	83.65	69.49	78.85	81	77.88
Grasslands	90.38	81.03	75.91	91.76	87.76	72.8	77.64	83.14	81.04	83.52
Wetlands	74.14	63.24	100	8.82	67.14	69.12	26.74	33.82	68.42	57.35
Water bodies	75.47	80	68.97	80	72.73	48	100	14	77.14	54
Construction land	78.57	66.67	61.7	87.88	63.89	69.7	78.95	45.45	37.21	48.48
Bare land	36.97	88.41	62.16	33.33	66.67	43.48	61.67	53.62	75	43.48
OA (%)	79.32		75.07		74.11		70.14		75.46	
Kappa	0.71		0.61		0.64		0.56		0.64	

UA of 78.95% but had a lower PA of only 45.45%, showing enormous differences between the UA and PA of construction land. Furthermore, this dataset had lower classification accuracy for the remaining LULC types, even having a lower UA of 26.74% in wetland classification (Table 6, Supplementary Table 2).

CLCD draws the grasslands and woodlands of the YRB more accurately than other LULC types and more reliably than most datasets (except LSV10). The differences between the UA and PA of grasslands (81.04% and 83.52%, respectively) and woodlands (81.00% and 77.88%, respectively) were also slighter than the other datasets (i.e., ESRI10, GLC\_FCS30, etc.) (Table 6, Supplementary Table 2). For croplands, although the UA was 68.84%, which was lower than that of grasslands and woodlands, it was still comparable to the GLC\_FCS30 and Globeland30 datasets, which had the same spatial resolution as CLCD (Table 6, Supplementary Table 2). Compared to GLC\_FCS30 and Globeland30, the classification accuracy of CLCD for the other LULC types was higher, but the accuracy for construction land needs to be improved.

In summary, the LSV10 dataset performed best in terms of the OA, Kappa coefficients, UA and PA of a single LULC type among the five datasets. In contrast, the ESRI10 dataset with the same spatial resolution of 10 m had poorer accuracy in portraying the LULC of the YRB, even lower than the CLCD data product with a lower resolution (Table 3, Supplementary Table 2). Concurrently, among the three Globeland30, GLC\_FCS30 and CLCD datasets with a 30 m spatial resolution, CLCD had a higher OA and Kappa coefficient and a mostly higher UA and PA for individual LULC types (excluding construction land) than the other two datasets.

## 4 Discussion

### 4.1 Evaluation of the five land use land cover datasets from the classification accuracy

Our results show that the most accurate identification of LULC in the YRB among the five datasets was made by LSV10, which has

the highest overall accuracy (79.32%), followed by CLCD (75.46%), ESRI10 (75.07%), and Globeland30 (74.11%), while the worst performance was by GLC\_FCS30 (70.14%) (Table 6). In contrast, the classification accuracy that was officially declared for LSV10 was 74.40% (Zanaga et al., 2021), 79.31% for CLCD (Yang and Huang, 2021), 85.96% for ESRI10 (Karra et al., 2021), 85.72% for Globeland30 (Chen et al., 2014a), and 72.27% for GLC\_FCS30 (Zhang et al., 2020). Except for the LSV10 data, the OA of all the datasets is lower than the officially declared accuracy by approximately 2.23–11.61% (Table 1). The probable reason for this discrepancy is the uneven spatial distribution of the training samples for the LULC classification, resulting in the lower accuracy of the multiple LULC types within the basin. Additionally, the officially claimed classification accuracy for the five datasets is assessed at the global (LSV10, ESRI10, Globeland30, GLC\_FCS30) or national scale (CLCD), whereas our study was conducted on the basin scale. Therefore, it is possible that our results differ somewhat from the official results. This also further suggests that accuracy assessments of LULC datasets in a specific region are necessary before conducting land cover change and related studies.

Grasslands constitute the dominant land cover type in the YRB, accounting for more than 75% of the basin (Table 2), and they are widely distributed in the basin (Supplementary Figure 2). For grasslands, the overall accuracy in the five datasets is generally high (higher than 70%) (Table 3). Meanwhile, the classification consistency of grasslands is also higher than that of the other LULC types in the five datasets (Figures 4, 5A). Similarly, woodlands are identified comparably accurately in the five datasets. As major LULC types in the basin, grasslands and woodlands are distributed spatially continuously and have a unique spectrum that can easily be recognized with the assistance of digital elevation model (DEM) data and image texture (Lu et al., 2014). Likewise, water bodies are easily distinguishable from other LULC types due to their strong absorption of wavelengths other than the blue-green band, which presents a lower reflectance spectrum (Huang et al., 2018).

The average UA and OA of 71.62% and 79.79% for croplands, respectively, which are lower than those of woodlands (80.96% and

73.84%, respectively) and grasslands (82.55% and 82.45%, respectively) but higher than those of the remaining LULC types in the basin (Table 6). Croplands have the same spectral characteristics as grasslands and woodlands. Additionally, croplands are a type of land that has been disturbed to a greater extent by humans and that has textural characteristics in images that differentiate croplands from grasslands and woodlands, making it possible to be identified more accurately (Phalke et al., 2020). However, seasonal farming activities and the time remote sensing data are acquired may cause some croplands to be confused with grasslands and bare land, thus reducing their classification accuracy (Tariq et al., 2022) (Figure 4A, Supplementary Figure 3), which is evident in the GLC\_FCS30 data. In addition, the classification of croplands is easily confused with woodlands, i.e., some fruit trees, and the current classification system of several LULC data does not always distinguish this LULC type in detail (Xu et al., 2018). Thus, it is essential to clarify the distinction between croplands and terraces, economic fruit forests and other land types in the future.

The classification accuracy of bare land shows a vast difference in the five datasets (Table 6, Supplementary Table 2). For instance, the numbers of pixels recognized as bare land in LSV10 and ESRI10 show enormous differences of approximately  $10^8$  times. Despite the vast differences in the pixel numbers, bare land is mainly in the northwestern part of the YRB (Supplementary Figure 5). Bare land interacts most closely with the other LULC types, such as grasslands, croplands, and construction land. The significant ecological restoration projects of large-scale afforestation, reforestation, and rapid urbanization in the basin could result in rapid displacements of bare land with other LULC types in a short period (Zhang et al., 2019). In addition, the temporal differences between the remote sensing data sources of different datasets may lead to poor consistency in bare land assessment (Li et al., 2017; Nguyen et al., 2021). Furthermore, the small patches (Table 4) and discrete distribution of bare land (Supplementary Figure 5) in the study area increase the uncertainty of the classification results due to the small sample size when selecting the training sample, resulting in considerable differences in the classification results among the five different datasets (Lu et al., 2014).

The classification accuracy of wetlands in the five datasets was generally low, with UA and PA values of 26.74% and 33.82%, respectively, in GLC\_FCS (Table 6). In particular, a significant difference of 91.81% was found in the UA and PA of ESRI10 (Table 6). The lower classification accuracy of the wetland stems from the confusion with grassland, woodlands and croplands (Table 5). For example, classification confusion can reach 81.29% in the GLC\_FCS30/CLCD combination (Table 5). The different definitions of wetlands in different datasets could influence the classification result. For instance, wetlands in LSV10 were defined as areas dominated by natural herbaceous vegetation (with vegetation coverage of no less than 10%) and inundated permanently or periodically by fresh, brackish or salt water, excluding mangroves (Zanaga et al., 2021). In contrast, wetlands in the Globeland30 data were defined as land located in the border zone between land and water, with shallow standing water or excessively wet soils, mainly growing with marsh or aquatic vegetation, including mangroves (Chen et al., 2014a). In addition, the conversion of wetlands to

other LULC types, such as croplands, grasslands and construction land, is faster due to a combination of natural and anthropogenic drivers, all of which affect the classification results of wetlands (Zong et al., 2009). Moreover, wetlands occupy a small proportion of the basin. Only six points for accuracy assessment fall within wetlands in the ESRI data, while most of these points were misclassified in ESRI10, all of which led to considerable differences between the UA and PA of ESRI10. In conclusion, all five datasets, i.e., LSV10, ESRI10, CLCD, GLC\_FCS30, and Globeland30, could be more suitable when studying spatiotemporal changes in grasslands and associated ecological and environmental effects. Corrections are required when revealing woodlands, croplands and bare land based on these datasets, and caution should be taken when studying spatiotemporal changes in wetlands, a vital LULC type of the basin.

## 4.2 Evaluation of the five land use land cover dataset from classification consistency

The five datasets consistently identify grasslands (Figure 5). However, the classification accuracy of some datasets for grassland types is lower (i.e., ESRI10, GLC\_FCS30). Grasslands in the GLC\_FCS30 data are abnormal in the Mu Us Desert, and there is an apparent boundary between croplands and grasslands within the Mu Us Desert. Croplands are more densely distributed on the left side of the boundary than on the right side (Supplementary Figure 6). In contrast, the distribution of grassland on the right side of the boundary is significantly denser than that on the left side (Supplementary Figure 6). Given the large extent of the anomalies and the absence of similar anomalies in the other data, it is possible that GLC\_FCS30 may have been confused in determining grasslands versus croplands and that the spatial consistency of croplands and grasslands was also affected. The main reason for this confusion originates from the similar spectrum of grasslands and croplands, which leads to the misclassification or omission of these two LULC types (Qin, 2000), thus forming a boundary effect, which is especially apparent when there were fewer satellite data before 2000 because most LULC datasets were generated using a method that first obtained the base data of a period and then detected the dynamics (Supplementary Figure 6).

The five datasets showed a lower classification consistency for woodlands and croplands and poorer consistency in identifying bare land in the YRB compared with grasslands (Figures 5B, C, Table 5, Figure 4). The lower classification consistency of the five datasets in woodlands and croplands is mainly due to the wide distribution of terraces and fruit trees in the southeastern part of the basin, the spectrum of which is similar in the multispectral data (Ji et al., 2021) and presents similar textural patterns that usually confuse LULC classifications. For bare land, this category is a high albedo LULC type, the spectrum of which is similar to that of construction land and ice and snow. In particular, bare land is easily confused with impervious surfaces, which makes accurate classification of bare land difficult (Li et al., 2017). In addition, we upscaled the original ESRI10 and LSV10 data. Although this data processing ensured the consistency of the five LULC datasets, data

precision might be lost. Considering the small size of patches and the more dispersed spatial distribution, bare land in the basin is easily overshadowed by other features after upscaling, which is one of the critical reasons for the poor consistency of bare land.

### 4.3 Implication and limitations

Our study identified the strengths and weaknesses of five LULC datasets based on accuracy assessments and classification consistency analysis. According to our study, the classification accuracy of grasslands in the LSV10 and CLCD datasets is higher than that in the other datasets, while it is lower in classifying wetlands. In contrast, Globeland30 provides high accuracy for wetland identification. Mining this helpful information can be a reliable basis for selecting different data for subsequent research on land cover change and related studies in the YRB (Ran et al., 2009). This study can also provide adequate support for the fusion of information from multiple datasets to improve the classification accuracy of current LULC data products (Bai and Feng, 2018). Based on this study, the LSV10 data outperformed the other four datasets in terms of both overall accuracy and classification accuracy for a single LULC type. This dataset can be the preferred data for future land cover prediction in the YRB. However, given that the dataset has been released for only two periods (2020, 2021), it is insufficient for the long-term time series characterization of land cover change in the basin over the historical period. However, the area composition correlation of the LSV10 data with the other four datasets could be used as one of the criteria for selecting other data. For example, the two datasets with high correlation coefficients with the LSV10 data are GLC\_FCS30 ( $R = 0.984$ ) and CLCD ( $R = 0.971$ ) (Figure 3). Considering the temporal resolution and data accuracy, CLCD and Globeland30 may be more suitable for long-term time series land cover change studies in the basin, with CLCD having the advantage due to its annual scale resolution (Supplementary Table 1).

Some limitations of our study should not be overlooked. Only one period, 2020, for the five LULC datasets was selected for accuracy assessment and classification consistency analysis in this study. The results provide the basis for data selection for future land-use modeling studies in the YRB. However, the results cannot be used for data selection for research on long-term time series LULC changes and the resultant ecological effects. This study is a preliminary study to conduct subsequent long-term time series research on LULC changes and future simulations in the YRB and to assess the ecological and environmental effects of LULC changes. Therefore, it is essential to evaluate multisource LULC datasets at different periods to better demonstrate the long-term spatiotemporal changes in LULC or to perform data fusion to generate more highly accurate LULC data to support subsequent research.

Furthermore, the upscaling method is a prerequisite for accuracy assessment and consistency analysis of multisource LULC data. However, this method may introduce some uncertainties. The method usually has limited effects on LULC types with large areas and a continuous distribution, such as grasslands in the YRB, while it could trigger finer changes in the spatial information and quantitative characteristics of LULC types with small patches and complex mosaic

distributions (Yang et al., 2001; Hu et al., 2013). In addition, the reclassification of LULC types might influence the study. However, the uncertainty introduced by reclassifying classification systems could be minimized to LSV10, ESRI10, CLCD, and Globeland30 since the old LULC classification systems match well with the reclassification system (Tables 2, 3). In comparison, the uncertainty originating from the reclassification of LULC types in GLC\_FCS might be higher than that in the other four LULC types. It is also noteworthy that some uncertainty may exist in accuracy assessments based on the high-resolution data provided by Google Earth. The high-resolution images provided online by Google Earth are composed of remote sensing images from multiple sensors and different seasons in the same year, and there are likely to be some temporal differences. For example, grasslands and water bodies are susceptible to interconversion in different seasons, which may also impact the results of accuracy assessments.

## 5 Conclusion

The consistency (in terms of classification and spatial consistency) of multisource LUCC data in the Yellow River Basin varies widely, as does the overall classification accuracy and the classification accuracy of a single LULC type. Given the importance of the YRB in ecological security and economic development in China, an accurate understanding of the spatial and temporal changes in LULC is vital to grasp the situation of resources and the environment and to ensure the sustainability of regional development. To determine the most valid LULC data for the YRB, five mainstream LULC datasets were selected and tested for classification consistency and accuracy in terms of land composition similarity, classification and spatial consistency, and classification accuracy. Our results indicate that although the five datasets showed good consistency in the land cover composition of the YRB, significant differences in the area of each land cover type were identified by the different datasets. The five datasets have good classification consistency for the main land cover types in the YRB, such as grasslands, croplands, and woodlands, and approximately 74% of the basin can be considered to be accurately identified. The LSV10 dataset exhibited the best classification accuracy (both overall and for a single LULC type) and Kappa coefficient among the five datasets. However, considering the time series and temporal resolution, the CLCD data, which showed substantial similarity to LSV10, may be ideal for conducting studies related to LULC in the YRB.

## Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

## Author contributions

WZ: Conceptualization, Funding acquisition, Writing – original draft. JT: Formal Analysis, Investigation, Methodology, Writing –

original draft. XZ: Funding acquisition, Writing – review & editing. JC: Writing – review & editing. YY: Conceptualization, Funding acquisition, Methodology, Writing – original draft, Writing – review & editing.

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the Natural Science Foundation of Henan Province of China (grant number 212300410212), the Key Scientific Research Project Plan of Colleges and Universities in Henan Province (grant number 22A170014), and Key Research Projects of Humanities and Social Sciences of Henan Provincial Department of Education in 2022 (2022-ZZJH-088).

## Acknowledgments

We would like to thank Shihua Zhu and Yuanzheng Li for providing valuable advice and comments while writing this article.

## References

- Bai, Y., and Feng, M. (2018). Data fusion and accuracy evaluation of multi-source global land cover datasets (in Chinese). *Acta Geogr. Sin.* 73, 2223–2235. doi: 10.11821/dlxb201811013
- Chen, J., Ban, Y., and Li, S. (2014a). Open access to Earth land-cover map (in Chinese). *Nature* 514 (7523), 434–434. doi: 10.1038/514434c
- Chen, J., Liao, A., Cao, X., Chen, L., Chen, X., Peng, S., et al. (2014b). Concepts and key techniques for 30 m global land cover mapping (in Chinese). *Acta Geod. Cartogr. Sin.* 43, 551–557. doi: 10.13485/j.cnki.11-2089.2014.0089
- China (2021). Outline of the Yellow River Basin's ecological protection and high-quality development plan (in Chinese). *Gazette State Council People's Republic China* 30, 001.
- Dai, Z., Hu, Y., and Zhang, Q. (2017). Agreement analysis of multi-source land cover products derived from remote sensing in South America (in Chinese). *Remote Sens. Inf* 32, 137–148. doi: 10.3969/j.issn.1000-3177.2017.02.021
- Friedl, M. A., Mciver, D. K., Hodges, J., Zhang, X. Y., Muchoney, D., Strahler, A. H., et al. (2002). Global land cover mapping from MODIS: algorithms and early results. *Remote Sens. Environ.* 83 (1-2), 287–302. doi: 10.1016/S0034-4257(02)00078-0
- Hansen, M. C., Defries, R. S., Townshend, J. R. G., and Sohlberg, R. (2000). Global land cover classification at 1 km spatial resolution using a classification tree approach. *Int. J. Remote Sens.* 21 (6-7), 1331–1364. doi: 10.1080/014311600210209
- Hu, Y., Xu, Z., Liu, Y., and Yan, Y. (2013). A review of the scaling issues of geospatial data (in Chinese). *Adv. Earth Sci.* 28 (3), 297.
- Hu, Y., Zhang, Q., Dai, Z., Huang, M., and Yan, H. (2015). Agreement analysis of multi-sensor satellite remote sensing derived land cover products in the Europe Continent (in Chinese). *Geographical Res.* 34 (10), 1839–1852.
- Huang, C., Chen, Y., Zhang, S., and Wu, J. (2018). Detecting, extracting, and monitoring surface water from space using optical sensors: A review. *Rev. Geophysic* 56 (2), 333–360. doi: 10.1029/2018RG000598
- Ji, Q., Liang, W., Fu, B., Zhang, W., Yan, J., Lv, Y., et al. (2021). Mapping land use/cover dynamics of the Yellow River Basin from 1986 to 2018 supported by google earth engine. *Remote Sens.* 13 (7). doi: 10.3390/rs13071299
- Jin, F. (2019). Coordinated promotion strategy of ecological protection and high-quality development in the Yellow River Basin (in Chinese). *Reform* 11, 33–39.
- Jin, F., Ma, L., and Xu, D. (2020). Environmental stress and optimized path of industrial development in the Yellow River Basin. *Resour. Sci.* 42 (1), 127–136. doi: 10.18402/resci.2020.01.13
- Karra, K., Kontgis, C., Statman-Weil, Z., Mazzariello, J. C., Mathis, M. M., and Brumby, S. P. (2021). Global land use / land cover with Sentinel 2 and deep learning.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fevo.2023.1275054/full#supplementary-material>

2021 IEEE Int. Geosci. Remote Sens. Symposium IGARSS, 4704–4707. doi: 10.1109/IGARSS47720.2021.9553499

Khan, A., Richards, K. S., Parker, G. T., McRobie, A., and Mukhopadhyay, B. (2014). How large is the Upper Indus Basin? The pitfalls of auto-delineation using DEMs. *J. Hydrology* 509, 442–453. doi: 10.1016/j.jhydrol.2013.11.028

Li, H., Wang, C., Zhong, C., Su, A., Xiong, C., Wang, J., et al. (2017). Mapping urban bare land automatically from landsat imagery with a simple index. *Remote Sens.* 9 (3). doi: 10.3390/rs9030249

Li, Z., He, W., Cheng, M., Hu, J., Yang, G., and Zhang, H. (2023). SinoLC-1: the first 1-meter resolution national-scale land-cover map of China created with the deep learning framework and open-access data. *Earth Syst. Sci. Data Discuss* 2023, 1–38. doi: 10.5194/essd-2023-87

Lu, D., Li, G., and Moran, E. (2014). Current situation and needs of change detection techniques. *Int. J. Image Data Fusion* 5 (1), 13–38. doi: 10.1080/19479832.2013.868372

Lyons, M. B., Keith, D. A., Phinn, S. R., Mason, T. J., and Elith, J. (2018). A comparison of resampling methods for remote sensing classification and accuracy assessment. *Remote Sens. Environ.* 208, 145–153. doi: 10.1016/j.rse.2018.02.026

Nguyen, C. T., Chidthaisong, A., Kieu Diem, P., and Huo, L.-Z. (2021). A modified bare soil index to identify bare land features during agricultural fallow-period in Southeast Asia using landsat 8. *Land* 10 (3), 231. doi: 10.3390/land10030231

Phalke, A. R., Özdoğan, M., Thenkabail, P. S., Erickson, T., Gorelick, N., Yadav, K., et al. (2020). Mapping croplands of Europe, Middle East, Russia, and Central Asia using Landsat, Random Forest, and Google Earth Engine. *Isprs J. Photogrammetry Remote Sens.* 167, 104–122. doi: 10.1016/j.isprsjprs.2020.06.022

Qin, Q. (2000). The problem and approach in the auto-interpretation of remote sensing imagery (in Chinese). *Sci. Surveying Mapp.* 25 (2), 21–24.

Ran, Y., Li, X., and Lu, L. (2009). Accuracy evaluation of the four remote sensing based land cover products over China (in Chinese). *J. Glaciology Geocryology* 31 (03), 490–500.

Tang, G. (2019). "Digital elevation model of China (1KM)," in *Big Earth Data Platform for Three Poles*. BEDPFT Poles. <https://poles.tpdc.ac.cn/en/data/12e91073-0181-44bf-8308-c50e5bd9a734/>

Tariq, A., Yan, J., Gagnon, A. S., Riaz Khan, M., and Mumtaz, F. (2022). Mapping of cropland, cropping patterns and crop types by combining optical remote sensing images with decision tree classifier and random forest. *Geo-spatial Inf. Sci.* 26, 1–19. doi: 10.1080/10095020.2022.2100287

Tilahun, A. (2015). Accuracy assessment of land use land cover classification using google earth. *Am. J. Environ. Prot.* 4, 193. doi: 10.11648/j.ajep.20150404.14

Wang, H., Yan, H., Hu, Y., Xi, Y., and Yang, Y. (2022). Consistency and accuracy of four high-resolution LULC datasets—IndoChina peninsula case study. *Land* 11 (5), 758. doi: 10.3390/land11050758



- Wang, Y., Tan, D., Han, L., Li, D., Wang, X., Lu, G., et al. (2021). Review of climate change in the Yellow River Basin. *J. Desert Res.* 41, 235–246. doi: 10.7522/j.issn.1000-694X.2021.00086
- Watson, R. T., Noble, I. R., Bolin, B., Ravindranath, N. H., Verardo, D. J., and Dokken, D. J. (2001). Land use, land-use change and forestry: a special report of the Intergovernmental Panel on Climate Change. Cambridge, UK: Cambridge University Press. doi: 0.1017/S0376892901280308
- Xi, J. (2019). Speech at the symposium on ecological protection and high-quality development of the Yellow River Basin. *China Water Resour* 20, 1–3.
- Xu, X., Liu, J., Zhang, S., Li, R., Yan, C., and Wu, S. (2018). *China multi-period land use and land cover remote sensing monitoring data set (CNLUCC)*, RaESaD Center. <https://www.resdc.cn/DOI/DOL.aspx?DOIID=54>
- Yang, J., and Huang, X. (2021). The 30 m annual land cover dataset and its dynamics in China from 1990 to 2019. *Earth System Sci. Data* 13 (8), 3907–3925. doi: 10.5194/essd-13-3907-2021
- Yang, C., Liu, J., Zhang, Z. X., and Wang, C. (2001). Analysis of accuracy loss during rasterizing vector data with different grid size (in Chinese). *J. Mountain Res.* 19 (3), 258–264. doi: 10.16089/j.cnki.1008-2786.2001.03.01
- Yang, Y., Liu, L., Zhang, P., Wu, F., Zhou, Y., Song, Y., et al. (2023). Advances and driving factors in soil organic carbon storage during vegetation restoration in the Loess Plateau, China (in Chinese). *J. Earth Environ.*, 1–24. doi: 10.7515/JEE221009
- Yang, Y., Xiao, P., Feng, X., and Li, H. (2017). Accuracy assessment of seven global land cover datasets over China. *Isprs J. Photogrammetry Remote Sens.* 125. doi: 10.1016/j.isprs.2017.01.016
- Zanaga, D., Kerchove, R. V. D., Keersmaecker, W. D., Souverijns, N., Brockmann, C., Quast, R., et al. (2021). *ESA WorldCover 10 m 2020 v100*.
- Zhang, X., Liu, L., Chen, X., Gao, Y., and Mi, J. (2020). GLC\_FCS30: Global land-cover product with fine classification system at 30 m using time-series Landsat imagery. *Earth Syst. Sci. Data* 13 (6), 2753–2776. doi: 10.5194/essd-2020-182
- Zhang, X., Liu, L., Fang, S., Jiang, W., and Wang, J. (2014). Research advances on the relationship between land use/cover change and environmental change (in Chinese). *Ecol. Environ. Sci.* 23 (12), 2013–2021. doi: 10.16258/j.cnki.1674-5906.2014.12.018
- Zhang, Y., Liu, L., Wang, Z., Bai, W., Ding, M., Wang, X., et al. (2019). Spatial and temporal characteristics of land use and cover changes in the Tibetan Plateau. *Chin. Sci. Bull.* 64 (27), 2865–2875. doi: 10.1360/TB-2019-0046
- Zhang, Z., Liu, H., Zuo, Q., Yu, J., and Li, Y. (2021). Spatiotemporal change of fractional vegetation cover in the Yellow River Basin during 2000–2019 (in Chinese). *Resour. Sci.* 4, 849–858. doi: 10.18402/resci.2021.04.18
- Zhao, Y., Hu, C., Zhang, X., Wang, Y., Cheng, C., Yin, X., et al. (2018). Analysis on runoff and sediment regimes and its causes of the Yellow River in recent 70 years (in Chinese). *Trans. Chin. Soc. Agric. Eng.* 34, 112–119. doi: 10.11975/j.issn.1002-6819.2018.21.014
- Zong, X., Liu, G., Qiao, Y., and Lin, S. (2009). Study on dynamic changes of wetland landscape pattern in Yellow River Delta (in Chinese). *J. Geo-information Sci.* 11 (1), 91–97.