

## Article

# Wetlands Mapping and Monitoring with Long-Term Time Series Satellite Data Based on Google Earth Engine, Random Forest, and Feature Optimization: A Case Study in Gansu Province, China

Jian Zhang , Xiaoqian Liu, Yao Qin, Yaoyuan Fan and Shuqian Cheng

College of Geography and Environmental Science, Northwest Normal University, Lanzhou 730070, China; 2022213026@nwnu.edu.cn (X.L.); 2021212884@nwnu.edu.cn (Y.Q.); 2023212956@nwnu.edu.cn (Y.F.); 2022223056@nwnu.edu.cn (S.C.)

\* Correspondence: jianzhang@nwnu.edu.cn

**Abstract:** Given global climate change and rapid land cover changes due to human activities, accurately identifying, extracting, and monitoring the long-term evolution of wetland resources is profoundly significant, particularly in areas with fragile ecological conditions. Gansu Province, located in northwest China, contains all wetland types except coastal wetlands. The complexity of its wetland types has resulted in a lack of accurate and comprehensive information on wetland changes. Using Gansu Province as a case study, we employed the GEE platform and Landsat time-series satellite data, combining high-quality sample datasets with feature-optimized multi-source feature sets. The random forest algorithm was utilized to create wetland classification maps for Gansu Province across eight periods from 1987 to 2020 at a 30 m resolution and to quantify changes in wetland area and type. The results showed that the wetland mapping method achieved robust classification results, with an average overall accuracy (OA) of 96.0% and a kappa coefficient of 0.954 across all years. The marsh type exhibited the highest average user accuracy (UA) and producer accuracy (PA), at 96.4% and 95.2%, respectively. Multi-source feature aggregation and feature optimization effectively improve classification accuracy. Topographic and seasonal features were identified as the most important for wetland extraction, while textural features were the least important. By 2020, the total wetland area in Gansu Province was 10,575.49 km<sup>2</sup>, a decrease of 4536.86 km<sup>2</sup> compared to 1987. The area of marshes decreased the most, primarily converting into grasslands and forests. River, lake, and constructed wetland types generally exhibited an increasing trend with fluctuations. This study provides technical support for wetland ecological protection in Gansu Province and offers a reference for wetland mapping, monitoring, and sustainable development in arid and semi-arid regions.

**Keywords:** wetland mapping; random forests; feature optimization; long time series; wetland dynamics



**Citation:** Zhang, J.; Liu, X.; Qin, Y.; Fan, Y.; Cheng, S. Wetlands Mapping and Monitoring with Long-Term Time Series Satellite Data Based on Google Earth Engine, Random Forest, and Feature Optimization: A Case Study in Gansu Province, China. *Land* **2024**, *13*, 1527. <https://doi.org/10.3390/land13091527>

Academic Editors: Zhenguo Niu, Bin Zhao, Zhaoqing Luan and Bo Guan

Received: 29 August 2024

Revised: 18 September 2024

Accepted: 19 September 2024

Published: 20 September 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Wetlands, referred to as the “kidneys of the earth”, are recognized as one of the most crucial natural ecosystems globally due to their significant ecological functions and socio-economic value [1–3]. They play an indispensable role in climate regulation, food and fiber production, flood control and storage, carbon sequestration and storage, water treatment and purification, as well as in the maintenance of biodiversity, ecological security, and human well-being [4,5]. Although wetlands account for only 9% of the total land area, they contribute over 23% of the global ecosystem service value [6,7]. Despite their immense value, over half of the world’s wetlands have disappeared in recent decades due to drainage and filling [8]. The surrounding landscape of the remaining wetlands is undergoing significant changes, which may adversely affect their condition and functionality [9]. Therefore, it is essential to develop rapid methods for identifying, extracting, analyzing, and evaluating

wetland resources. Accurate inventorying of wetland distribution is crucial to safeguard against further changes and degradation of land cover while enhancing our understanding of wetland conditions and resource management.

Remote sensing technology has been widely used for large-scale wetland mapping and monitoring, due to the low cost of data acquisition and high spatial and temporal resolution [10]. In recent decades, several global wetland datasets have been generated from satellite images, categorized into three groups: land cover and land use change (LCLUC) datasets, single-type wetland datasets, and multitype wetland datasets [7]. The global LCLUC datasets include the MCD12Q1 product by NASA, updated yearly from 2000 to 2020; the 300 m CCI\_LC product by ESA updated yearly from 1992 to 2020; the GLC\_FCS30 product by Liu's team at the Institute of Space and Astronautical Information Innovation, Chinese Academy of Sciences, updated every five years from 1985 to 2020 [11]; the GlobeLand30 data released by the National Geographic Information Center [12]; and the CLCD product by Yang and Huang (2021) from Wuhan University, updated yearly from 1985 to 2021 [13]. Although these products maintain good temporal continuity, they include only two wetland-related land types: water and wetland. Consequently, these datasets do not distinguish between more detailed wetland types, such as rivers, lakes, marshes, constructed wetlands, and coastal wetlands [14]. Single-type wetland datasets focus on a single wetland type (e.g., water, mangroves) and can accurately delineate the extent of specific wetland types [15,16]. However, these datasets do not provide a comprehensive understanding of all wetland types. Multitype wetland datasets delineate the spatial distribution of different wetland types, but their resolution and thematic accuracy are too coarse and do not account for temporal changes, limiting their utility for wetland protection and management, as exemplified by the Global Lakes and Wetlands Database (GLWD) [17]. Therefore, mapping wetlands with detailed types and high spatiotemporal resolution is essential for monitoring and studying wetland resource dynamics.

Given the variations in geographic locations and wetland types, selecting appropriate classification methods for extracting wetland information is particularly important. Machine learning algorithms, such as decision trees (DTs), random forests (RFs), and support vector machines (SVMs), are widely used in wetland mapping due to their fast construction speed, high discrimination efficiency, and effective utilization of big Earth data [18]. DTs offer certain advantages over traditional maximum likelihood supervised classification, but they rely heavily on empirical knowledge and human intervention, which has a significant impact on classification accuracy [19]. Although some studies demonstrate that SVMs can produce highly accurate wetland maps [20–22], they struggle with large-scale training samples and multi-classification problems and are also very sensitive to parameter and kernel function selection [23]. Random forest (RF) has proven to be practical due to its advantages in processing correlated high-dimensional data [24], addressing nonlinearity and overfitting [25], as well as screening and prediction. Ranking the importance of selected factors enhances simulation precision [26]. To date, RF is considered one of the most effective algorithms for classifying remotely sensed images in topographically complex regions [27]. Lawrence and Moran (2015) systematically compared the performance of multiple machine learning classification algorithms using 30 different datasets and found that RF's average classification accuracy was significantly higher than other algorithms [28]. More importantly, RF has distinct advantages in wetland classification [29]. Studies, such as those by Millard and Richardson (2013), have identified RF's potential in marsh mapping, demonstrating its ability to prioritize features based on their classification accuracy [30]. Therefore, RF can be effectively utilized for wetland mapping by integrating existing wetland samples with environmental data as model inputs.

Due to the complexity of wetland feature types and significant spectral confusion, the extraction and selection of remote sensing classification features present a substantial challenge for current wetland mapping. Spectral reflectance features are the basis for wetland remote sensing extraction. Wetlands are intrinsically linked to vegetation, water bodies, soil, topography, and other environmental factors. Since wetlands are constantly

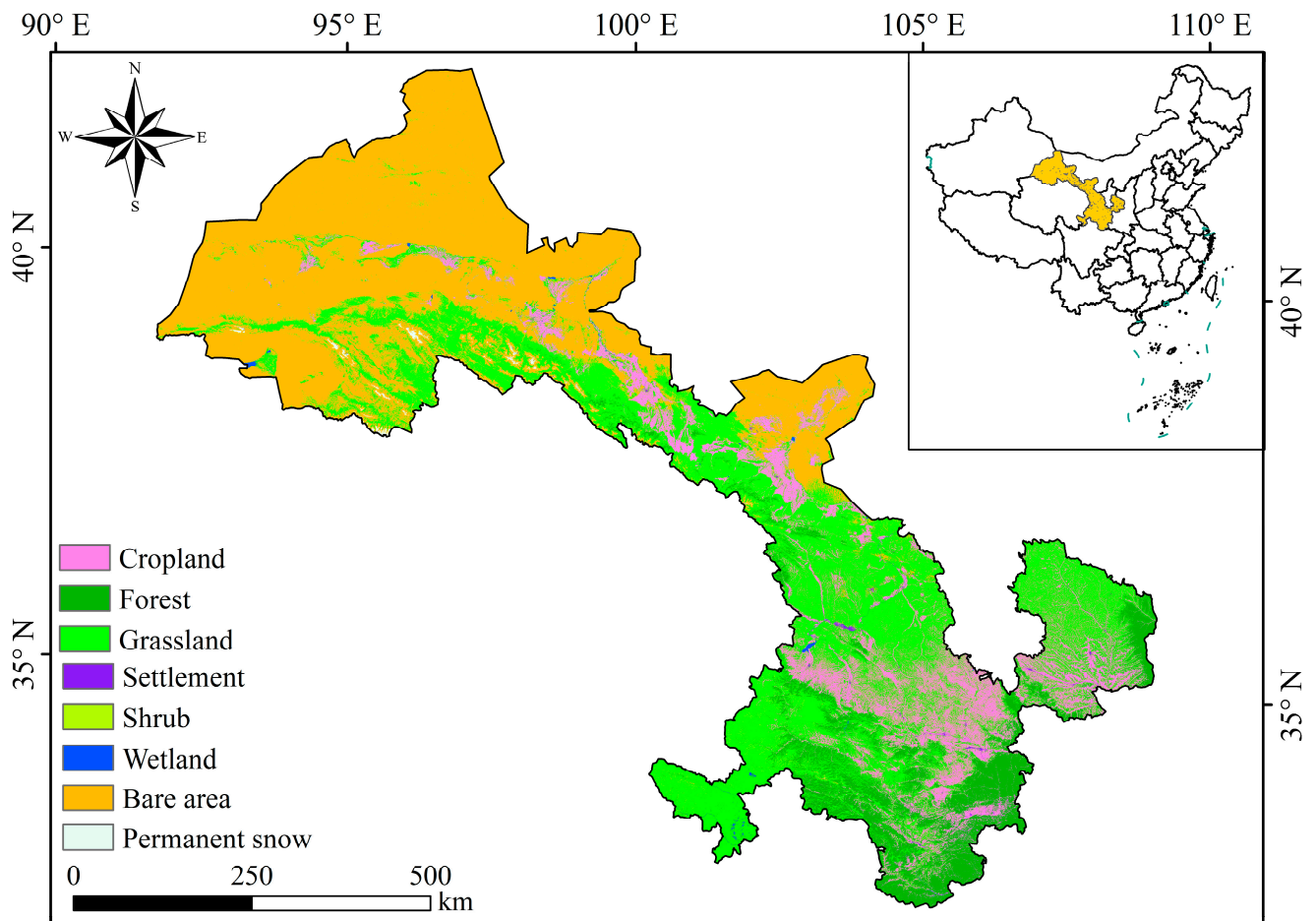
changing, multi-temporal images can capture the seasonal characteristics of wetland water bodies or vegetation, enhancing classification accuracy [31]. While it is undoubtedly beneficial to fully utilize multiple types of features to identify wetland types, an excess of features can lead to “dimensional catastrophe” and data redundancy [32]. Therefore, it is necessary to filter and optimize features before classification. Currently, common feature optimization methods include filter, embedded, and wrapper techniques [33]. Among these, recursive feature elimination (RFE) is a classical wrapper algorithm that can extract effective features in land cover classification studies of wetlands, cities, forests, and other landscapes [34–36].

Gansu Province, located in northwestern China, has a fragile ecological background and diverse types of wetlands, including rivers, lakes, marshes, and constructed wetlands, but excluding coastal wetlands (China’s national standard GB/T 24708-2009). Since the vast majority are in semi-arid and arid regions, these wetland resources are particularly precious and play a vital role in mitigating drought and flood disasters, maintaining watershed water balance, and promoting Gansu’s socioeconomic development [37]. However, in recent years, under the dual threat of natural factors (climate, vegetation, hydrology, etc.) and anthropogenic factors (water resource exploitation, agricultural and animal husbandry development, engineering construction, etc.) [38,39], the utilization, protection, and destruction of wetland resources in Gansu Province have been equally emphasized. Although some protection efforts have yielded results, Gansu Province still faces problems such as shrinking wetland areas, ecological function degradation, and water environment deterioration [40]. Monitoring wetland resources in Gansu Province and understanding the reality of the distribution and evolution trends of various wetland types are profoundly significant for further wetland protection and management. Therefore, this study uses Gansu Province as a case to explore the mapping method of wetland resource types using machine learning within a feature optimization framework and to identify the spatial distribution of different wetland types. The main objectives of this study are (1) to construct a classification method using random forests within a feature optimization framework and map the wetland categories of Gansu Province from 1987 to 2020 based on the GEE platform and Landsat data; (2) to assess the influence of different features on automatic wetland classification and extraction; and (3) to analyze the dynamic changes and underlying reasons for wetland resource changes in Gansu Province from 1987 to 2020. This study can provide data support for the monitoring and managing of wetland resources in Gansu Province and offer potential application references for evaluating wetland-related goals at national and global scales in the context of global change.

## 2. Materials and Methods

### 2.1. Study Area

Gansu Province is located in northwest China, between 32°11′ N–42°57′ N and 92°13′ E–108°46′ E, and covers a total area of about 454,400 km<sup>2</sup> (Figure 1). The region is situated at the convergence of the Inner Mongolia, Loess, and Qinghai-Tibet Plateaus [41], featuring complex geomorphological types, including plateaus, mountains, basins, river valleys, deserts, and Gobi. The terrain is undulating and varied, with an overall high elevation, averaging 2260 m. Additionally, Gansu Province is situated at the confluence of the northwestern arid zone, the Qinghai-Tibet alpine region, and the eastern monsoon region [42], encompassing four major climate types: temperate continental (arid), plateau alpine, temperate monsoon, and subtropical monsoon climates. Precipitation varies considerably across the region, ranging from 37 mm to 750 mm, with a drying gradient from southeast to northwest. In addition, precipitation is unevenly distributed throughout the year, with most of it falling between the months of June and September [43]. Gansu Province is rich in solar energy resources, with annual sunshine hours increasing from southeast to northwest. The province’s rivers span from east to west and belong to the three major basins of the Yellow River, the Yangtze River, and the Inland River.



**Figure 1.** Overview of the study area. Note: The map is based on the Department of Natural Resources Standard Map Service website, Standard Map No. GS (2023) 2767, with no modifications to the base map boundaries.

## 2.2. Data Sources

In this study, Landsat 5 and Landsat 8 surface reflectance (SR) data with a resolution of 30 m, provided by GEE, were used. To ensure complete image coverage of the study area, images from the study period and the two years before and after were radiometrically calibrated, processed with a cloud mask, and then subjected to median filtering to obtain interannual composite images. The relevant image data are presented in Table 1. Additionally, digital elevation model (DEM) data from the Shuttle Radar Topography Mission (SRTM) with a resolution of 30 m were used in this study. These data were used to construct topographic features (elevation, slope, and slope direction) that can be directly utilized in the GEE cloud platform. The various auxiliary data used in this study are presented in Table 2. These data were utilized to aid in the creation of the sample dataset and the fine classification of wetlands.

**Table 1.** Number of images used in this study from 1987 to 2020.

Year	Satellite	Landsat Product Name	Total	Number
				Cloud $\leq$ 10%
1987	Landsat5 TM	LANDSAT/LT05/C02/T1_L2	960	663
1990	Landsat5 TM	LANDSAT/LT05/C02/T1_L2	1494	982
1995	Landsat5 TM	LANDSAT/LT05/C02/T1_L2	1395	958
2000	Landsat5 TM	LANDSAT/LT05/C02/T1_L2	1599	1134
2005	Landsat5 TM	LANDSAT/LT05/C02/T1_L2	1541	997
2010	Landsat5 TM	LANDSAT/LT05/C02/T1_L2	1375	933
2015	Landsat8 OLI	LANDSAT/LT08/C02/T1_L2	1873	1264
2020	Landsat8 OLI	LANDSAT/LT08/C02/T1_L2	1919	1251

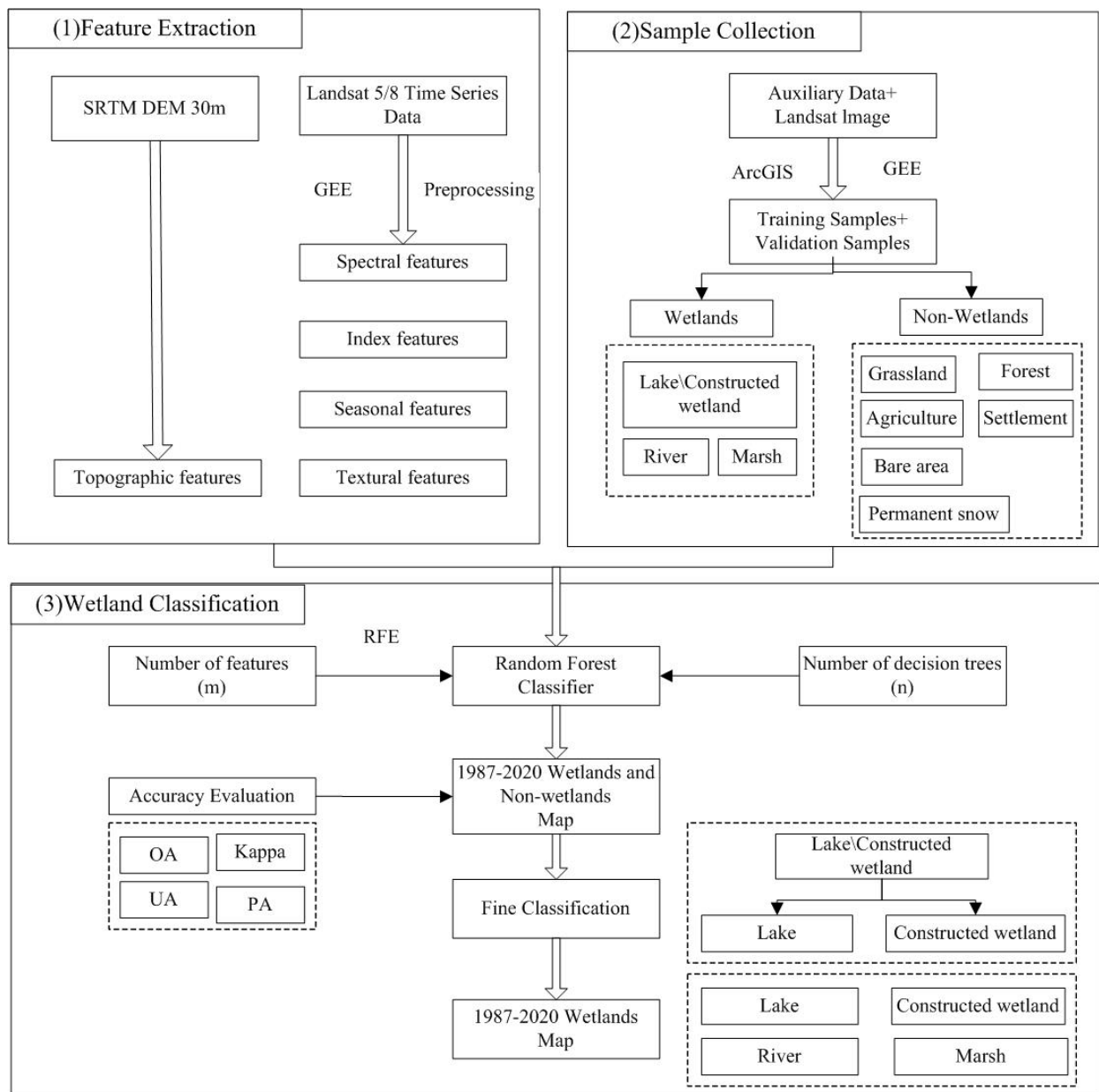
**Table 2.** Auxiliary data in this study.

Product Name	Date	Resolution	Data Resource	References
GlobalLand30	2000/2010/2020	30 m	<a href="http://globeland30.org/">http://globeland30.org/</a> (accessed on 7 July 2023)	[44]
GLC_FCS30	1985–2020	30 m	<a href="https://data.casearth.cn/">https://data.casearth.cn/</a> (accessed on 8 July 2023)	[11]
JRC-GSW	1984–2020	30 m	<a href="https://developers.google.com/earth-engine/datasets/">https://developers.google.com/earth-engine/datasets/</a> (accessed on 10 July 2023)	[15]
Hydro LAKES	—	1:24,000	<a href="http://www.hydrosheds.org">http://www.hydrosheds.org</a> (accessed on 7 July 2023)	[45]
GWRL	—	30 m	<a href="https://zenodo.org/">https://zenodo.org/</a> (accessed on 15 July 2023)	[46]
GLWD	—	1 km	<a href="https://www.worldwildlife.org/">https://www.worldwildlife.org/</a> (accessed on 20 June 2023)	[17]
SDSMW	2015	30 m	<a href="http://www.geodata.cn">http://www.geodata.cn</a> (accessed on 20 July 2023)	[47]
GOODD	2020	—	<a href="https://www.globaldamwatch.org/directory">https://www.globaldamwatch.org/directory</a> (accessed on 26 July 2023)	[48]
Reservoir statistics	—	—	<a href="http://www.stats.gov.cn/">http://www.stats.gov.cn/</a> (accessed on 27 June 2023)	—

Note: GLC\_FCS30: Global Land Cover with Fine Classification System at 30 m; JRC-GSW: JRC Global Surface Water; GWRL: Global River Widths from Landsat; GLWD: Global Lakes and Wetlands Database; SDSMW: spatial distribution dataset of marshes in China; GOODD: Global Geo-referenced Database of Dams.

### 2.3. Classification Method

The wetland classification process in this study is primarily divided into three steps (Figure 2): (1) Utilizing the GEE cloud platform to preprocess the Landsat series images and obtain five classification features. (2) Using ArcGIS 10.2 software, the GEE cloud platform, and auxiliary datasets to construct a sample collection process and obtain sample data. (3) Combining RF and RFE to select the optimal number of decision trees and the optimal combination of features. Three types of wetlands (lake/constructed wetland, river, and marsh) and six non-wetland types (grassland, agriculture, forest, settlement, bare area, and permanent snow) were obtained and evaluated for accuracy. Then, based on the wetland and non-wetland classification data, auxiliary data were used to make further distinction between lakes and constructed wetlands, ultimately obtaining the classification data of four types of wetlands in Gansu Province for eight periods from 1987 to 2020.



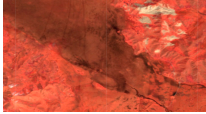

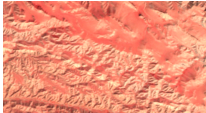







**Figure 2.** Flowchart of wetland classification in this study.

### 2.3.1. Classification System and Sample Construction

This study integrates the Convention on Wetlands, China’s national standard for wetland classification (GB/T 24708-2009), while considering the current status of wetland resources in the study area. Consequently, the wetland resources in Gansu Province were categorized into natural and constructed wetlands. Natural wetlands are subdivided into rivers, marshes, and lakes. Constructed wetlands include reservoirs and ponds, while non-wetland categories encompass grasslands, forests, agriculture, settlement, bare area, and permanent snow (Table 3).

**Table 3.** The remote sensing classification system used in this study.

Category I	Category II	Description	Landsat Image Example
Natural wetlands	Lake	Natural depressions in the ground, varying in size and form, are filled with bodies of water.	
	River	A naturally occurring linear body of water, such as a river or stream.	
	Marsh	The surface or subsurface soils are often excessively wet, and swampy plants grow on the surface	
Constructed wetlands	Reservoir/pond	Artificial water storage facilities are constructed for irrigation, hydropower, flood control, and other purposes.	
Non-wetlands	Grassland	Land dominated by herbaceous vegetation.	
	Forest	Biomes dominated by woody plants.	
	Agriculture	Land cultivated for growing crops.	
	Settlement	Surfaces created by artificial construction activities.	
	Bare area	Essentially bare ground with no plant cover.	
	Permanent snow	Land permanently covered by snow.	

Generally speaking, the number and quality of samples largely determine the accuracy of remote sensing classification [49]. This study integrates auxiliary data and manual visual discrimination to improve the accuracy of classification samples. Since the remote sensing features of lakes and constructed wetlands are too similar to be automatically extracted using machine learning methods, this study initially classifies these two wetlands as one category for automatic remote sensing classification and then uses auxiliary data of lakes and reservoirs for secondary fine extraction.

In this study, ArcGIS10.2 software was used to generate random points within the study area and then generate sample points of each type in conjunction with auxiliary data (Table 2). JRC-GSW, Hydro L AKES, GWRL, GLWD, GOODD, and reservoir statistics were utilized to assist in the extraction of sample points from lakes/constructed wetlands

and rivers. SDSMW was utilized to assist in the sample point extraction from marshes. GlobalLand30 and GLC\_FCS30 were utilized to assist in the sample point extraction for grasslands, forests, settlements, agriculture, bare areas, and permanent snow areas. Finally, on the GEE platform, further visual identification of the generated sample points was carried out using Landsat imagery from each of the years. To ensure a sufficient number of samples of each type, additional samples were included during the visual discrimination process to generate the final sample set. The samples were divided into training and validation sets in a 7:3 ratio for each year in the study area, with the statistics presented in Table 4 below.

**Table 4.** Statistics on the number of samples of each type in the study area from 1987 to 2020.

Year	Lake/Constructed Wetland	River	Marsh	Grassland	Forest	Agriculture	Settlement	Bare Area	Permanent Snow
1987	355/152	360/154	358/153	358/153	356/153	368/158	336/144	442/190	162/69
1990	356/152	358/154	358/154	357/153	354/152	369/158	346/148	440/188	161/69
1995	357/153	361/155	359/154	356/152	356/153	368/158	347/149	441/189	164/70
2000	358/153	361/155	357/153	356/153	356/152	370/159	344/148	444/190	163/70
2005	355/152	360/154	358/154	357/153	355/152	370/158	344/147	441/189	158/68
2010	350/150	359/154	357/153	359/154	356/153	369/158	347/149	444/190	168/72
2015	349/150	358/154	358/153	356/152	358/154	370/158	342/147	448/192	155/67
2020	363/155	358/153	360/154	357/153	361/155	368/158	340/146	440/188	169/72

Note: The number of samples in the table is expressed as training/validation.

### 2.3.2. Feature Construction and Optimization

A total of 63 classification features for 5 classes were constructed for wetland extraction using the GEE cloud platform in this study. (Table 5). Six major spectral bands (blue, green, red, near-infrared, shortwave infrared1, and shortwave infrared2) were selected for spectral features. Index features included normalized difference vegetation index (NDVI) [50], enhanced vegetation index (EVI) [51], ratio vegetation index (RVI) [52], soil-adjusted vegetation index (SAVI) [53], modified soil-adjusted vegetation index (MSAVI) [54], normalized difference water index (NDWI) [55], enhanced water index (EWI) [56], modified normalized difference water index (MNDWI) [57], land surface water index (LSWI) [58], automatic water extraction index (AWEI) [59], modified adjusted water extraction index (MAWEI) [60], normalized difference building index (NDBI) [61], and normalized difference snow index (NDSI) [62]. Physically meaningful brightness, greenness, and humidity are obtained through tasseled cap transformation, reflecting the bare soil rocks, vegetation cover, and moisture information of the surface [63]. Phenological characteristics can reflect the seasonal phenomena of water bodies or vegetation life activities. Therefore, this study calculated NDWI, NDVI, MNDWI, LSWI, and EVI in spring, summer, fall, and winter to better extract plants and water bodies [64]. Texture is the spatial distribution of gray levels of neighboring pixels in an image region that obeys the statistical distribution law. The Gray Level Co-occurrence Matrix (GLCM) is a commonly used method for texture computation that not only responds to the distributional characteristics of contrast but also responds to the distributional characteristics among pixels with equally bright or similarly bright pixels and is a second-order statistical feature regarding the changes graphic contrast. Referring to previous studies, we selected 18 quantitative metrics as texture features based on the Gray Level Co-occurrence Matrix (GLCM) provided by GEE [65]. Additionally, this study utilized SRTM data to derive elevation, slope, and slope direction as topographic features [66]. Feature optimization reduces potential noise in the feature space and improves classification accuracy [32]. In this study, RFE is used for feature optimization. The algorithm removes the least important features in each round of iterations, then performs a new round of iterations until all features have been iterated, and ultimately compares the accuracy to determine the optimal combination of features [67].



**Table 5.** The specific information of classification features.

Feature Types	Indicators	Description or Formula	References
Spectral features	BLUE, GREEN, RED, NIR, SWIR1, SWIR2	Spectral bands	—
	NDVI	$(\text{NIR} - \text{RED}) / (\text{NIR} + \text{RED})$	[50]
	EVI	$2.5 \times (\text{NIR} - \text{RED}) / ((\text{NIR} + 6 \times \text{RED} - 7.5 \times \text{BLUE}) + 1)$	[51]
	RVI	$(\text{NIR} - \text{RED}) / (\text{NIR} + \text{RED})$	[52]
	SAVI	$(\text{NIR} - \text{RED}) \times 1.5 / (\text{NIR} + \text{RED} + 0.5)$	[53]
	MSAVI	$(2 \times \text{NIR} + 1 - \sqrt{(2 \times \text{NIR} + 1) - 8 \times (\text{NIR} - \text{RED})}) / 2$	[54]
Index features	NDWI	$(\text{GREEN} - \text{NIR}) / (\text{GREEN} + \text{NIR})$	[55]
	EWI	$(\text{Green} + \text{NIR} + \text{SWIR1}) / (\text{Green} - \text{NIR} - \text{SWIR1})$	[56]
	MNDWI	$(\text{GREEN} - \text{SWIR}) / (\text{GREEN} + \text{SWIR})$	[57]
	LSWI	$(\text{NIR} - \text{SWIR}) / (\text{NIR} + \text{SWIR})$	[58]
	AWEI	$4 \times (\text{GREEN} - \text{SWIR1}) - (0.25 \times \text{NIR} + 2.75 \times \text{SWIR2})$	[59]
	MAWEI	$5 \times (\text{GREEN} - \text{NIR}) + \text{BLUE} + \text{RED} + 4 \times \text{SWIR2}$	[60]
	NDBI	$(\text{SWIR} - \text{NIR}) / (\text{SWIR} + \text{NIR})$	[61]
	NDSI	$(\text{GREEN} - \text{SWIR}) / (\text{GREEN} + \text{SWIR})$	[62]
	BRIGHTNESS, GREENNESS, WETNESS	Tasseled cap transformation	[63]
Seasonal features	Four-season averages (NDWI, NDVI, MNDWI, LSWI, EVI) _SP/SU/FA/WI	Time series characteristics	[64]
Textural features	ASM, CONTRAST, CORR, MAXCORR, VAR, IDM, SAVG, SVAR, SENT, ENT, DVAR, DENT, IMCORR1, IMCORR2, INTERIA, DISS, PROM, SHADE	GLCM	[65]
Topographic features	ELEVATION, SLOPE, ASPECT	SRTM	[66]

Note: (NDWI, NDVI, MNDWI, LSWI, EVI) \_SP/SU/FA/WI: “\_SP” for spring, “\_SU” for summer, “\_FA” for fall, and “\_WI” for winter; ASM: angular second-order matrix; CORR: autocorrelation; MAXCORR: maximum correlation coefficient; VAR: variance; IDM: inverse difference matrix; SAVG: summed mean; SVAR: summed variance; SENT: sum entropy; ENT: entropy; DVAR: difference variance; DENT: difference entropy; IMCORR1: correlation information measure 1; IMCORR2: correlation information measure 2; INERTIA: inertia; DISS: dissimilarity; PROM: cluster prominence; SHADE: cluster shading.

### 2.3.3. Random Forest Classification

Random forest is an integrated learning method that employs decision trees as the basic classifiers [68]. Compared to other machine learning methods, random forest (RF) algorithms are more suitable for wetland extraction [23]. They effectively incorporate the idea of randomly selecting samples and features. Their powerful computational speed and generalization ability enable them to adapt to sample datasets with heterogeneous categories and have been widely used in classification problems with multidimensional features [69]. Moreover, the RF algorithm can evaluate and rank the importance of the input features, facilitating feature optimization to improve accuracy. Additionally, there are two main parameters to establish in the random forest step: the number of decision trees (n) and the number of features (m) [70,71].

### 2.3.4. Classification Accuracy Assessment

A commonly used method to evaluate the classification accuracy of remote sensing images is to calculate the confusion matrix using validation sample points. In this study, 70% of the sample data is used as training samples, and the remaining 30% is used as validation samples (Table 4). The overall accuracy (OA), kappa coefficient, user accuracy (UA), and producer accuracy (PA) are calculated using the confusion matrix to evaluate the accuracy of the classification results.

$$OA = \left( \frac{P_c}{P_n} \right) \times 100 \quad (1)$$

where  $P_c$  denotes the count of accurately classified pixels for a specific class, while  $P_n$  signifies the total number of pixels.

$$Kappa = \frac{N \sum_{i=1}^r x_{ii} - \sum_{i=1}^r (x_{i+} \times x_{+i})}{N^2 - \sum_{i=1}^r (x_{i+} \times x_{+i})} \quad (2)$$

where  $N$  represents the total number of pixels. The sums of each row and column in the confusion matrix are denoted by  $x_{i+}$  and  $x_{+i}$ , respectively, with  $x_{ii}$  specifically referring to the diagonal elements of the confusion matrix.

$$UA_i = \frac{N_{ii}}{N_{i+}} \quad (3)$$

where  $UA_i$  denotes the user accuracy for the  $i$ -th land use type,  $N_{ii}$  indicates the count of correctly classified samples, and  $N_{i+}$  signifies the true number of samples for the  $i$ -th land use type.

$$PA_i = \frac{N_{ii}}{N_{+i}} \quad (4)$$

where  $PA_i$  denotes the producer accuracy for the  $i$ -th land use type,  $N_{ii}$  indicates the count of correctly classified samples, and  $N_{+i}$  signifies the predicted number of samples for the  $i$ -th land use type.

#### 2.4. Wetland Evolution Analysis

The land use transfer matrix is utilized to characterize the dynamics of various land use types based on the area transformation matrix of land cover over different time phases in the same area. This study employs this method to investigate the transfer and change characteristics between different wetlands, as well as between wetlands and non-wetlands [72]:

$$X = \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1j} \\ X_{21} & X_{22} & \dots & X_{2j} \\ \vdots & \vdots & \ddots & \vdots \\ X_{i1} & X_{i2} & \dots & X_{ij} \end{bmatrix} \quad (5)$$

where  $X_{ij}$  represents the area converted from land cover type  $i$  to  $j$ .

The single land use momentum ( $K$ ) reflects the rate of change in the area of a particular land use type in the study area over a certain time frame, as defined by the following formula [73]:

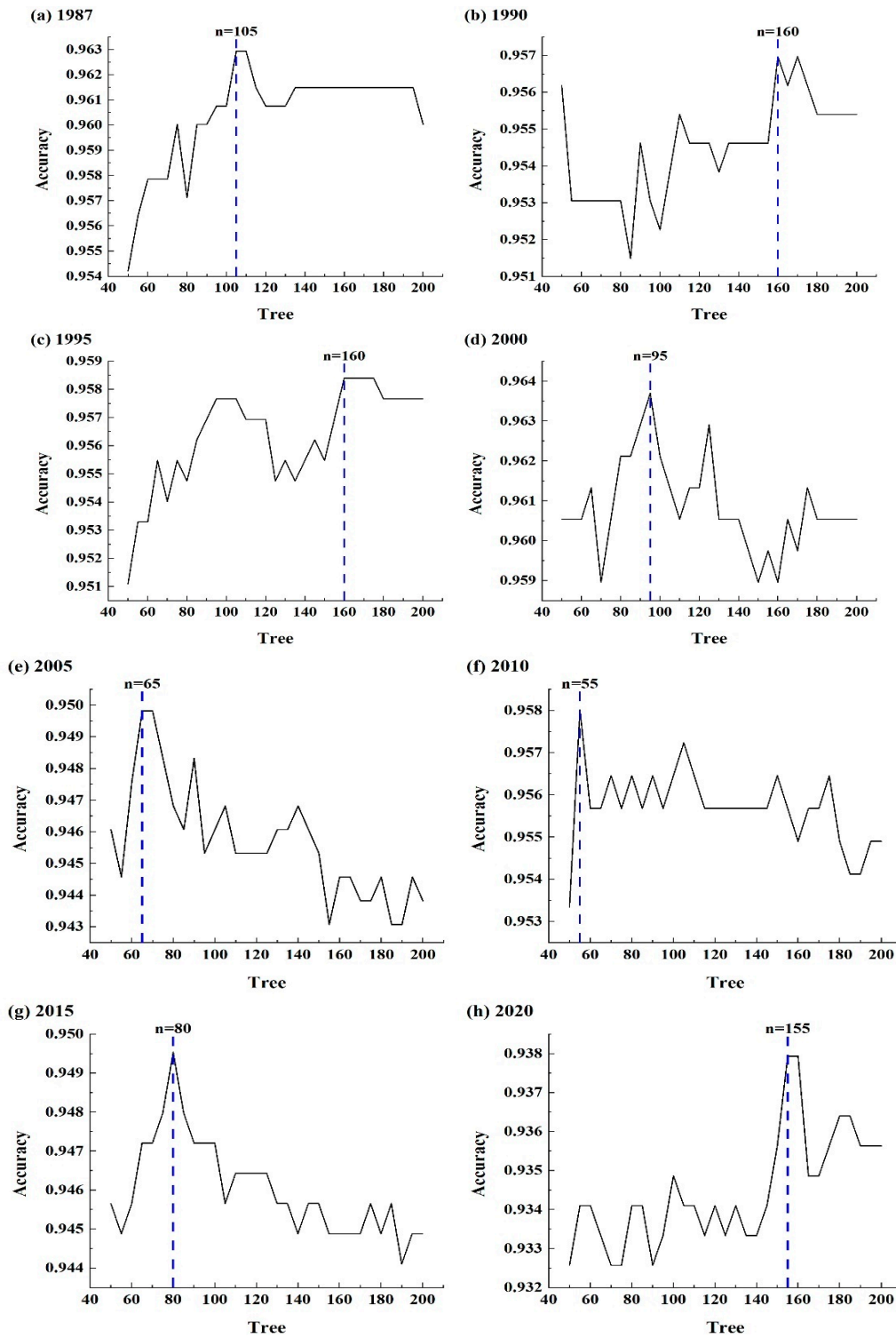
$$K = \frac{U_b - U_a}{U_a} \times \frac{1}{T} \times 100\% \quad (6)$$

where  $K$  represents the annual change rate (dynamic index) of a specific land use type over the study period;  $U_i$  and  $U_j$  denote the areas of the land type at the study's start and end, respectively; and  $T$  signifies the duration of the study period.

### 3. Results

#### 3.1. Number of Decisions Tree and Features Tuning

The number of decision trees ( $n$ ) has a direct relationship with the accuracy of the random forest classification model. To determine the optimal value of the random forest parameter decision tree ( $n$ ) for each year's classification, this study traversed all the results with ( $n$ ) taking values from 50 to 200 in increments of 5. The results show that the number of decision trees that give the highest accuracy varies from one year to the next (Figure 3). Following the accuracy statistics, the numbers 105, 160, 160, 95, 65, 55, 80, and 155 were eventually selected as the number of decision trees for the eight study periods ranging from 1987 to 2020.



**Figure 3.** Number of decision trees and corresponding accuracy of random forest models from 1987 to 2020.

To achieve better classification results, this study also calculated the importance corresponding to each feature in the random forest and used the RFE method to tune a total of 63 features across five types, ensuring that the number of features selected each year maximizes the accuracy of its classification results. From the graph depicting the number of features and the importance ranking of each study period (Figure 4), several observations can be made: (1) After feature optimization, the number of features in each year was lower than 35, with the lowest being 18 in

1990, 2000, and 2020, respectively. (2) Elevation features consistently scored highest in importance across study periods except for 2015. (3) Each study period showed a preference for five or more feature types, indicating that the inclusion of multiple feature types aids in classification. Moreover, topographic and seasonal features seemed to play a more significant role in classifying compared to other features. (4) Among the preferred features for all eight periods were ELEVATION, SLOPE, LSWI\_SU, MNDWI\_SU, NDGI, SWIR2, BRIGHTNESS, and SAVG, with MAWEI being selected in seven periods. (5) In feature optimization, texture features were frequently deleted, with only three features selected out of 18 in 2015, and none selected in 1987 and 2000.

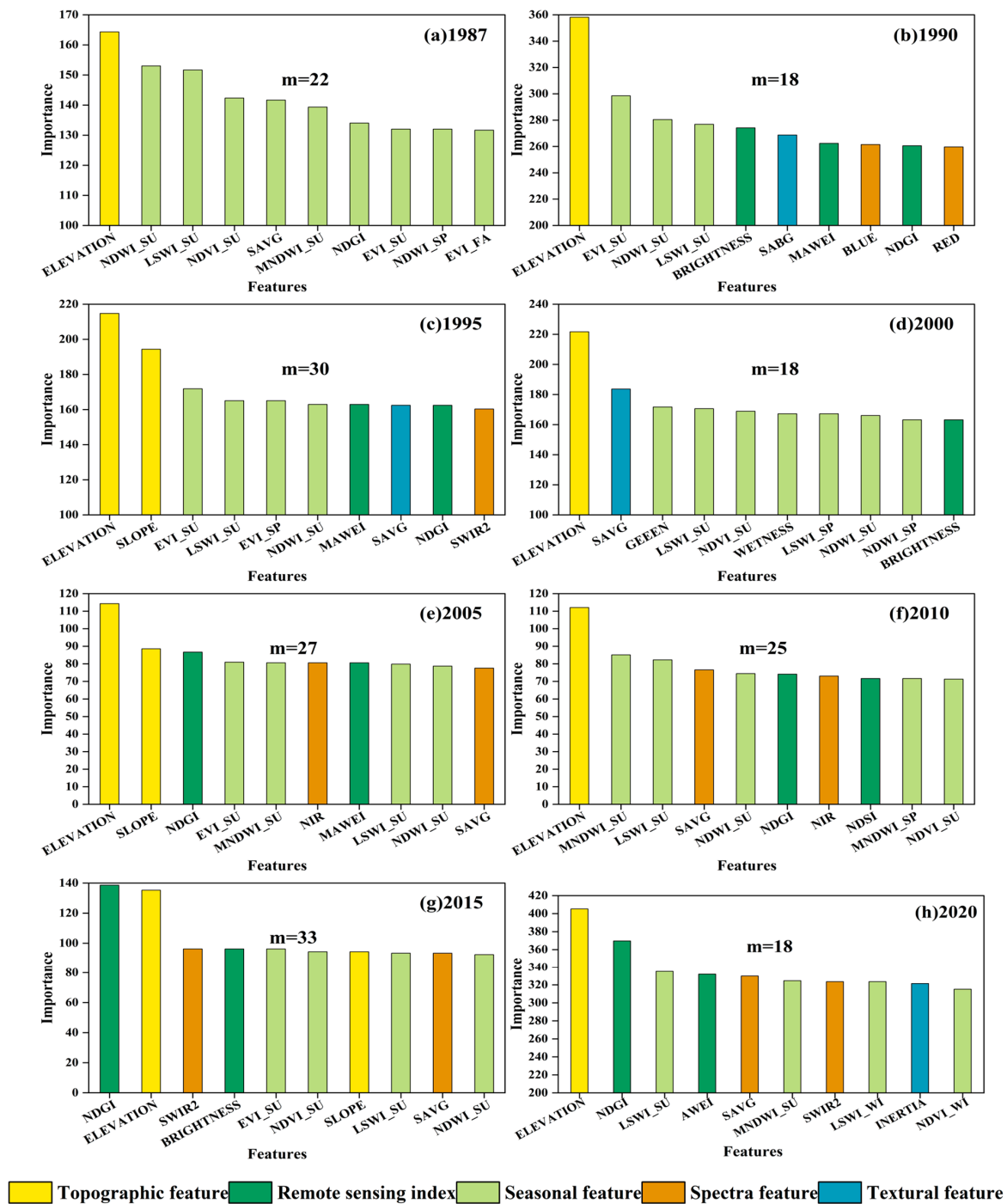


Figure 4. Top 10 feature preference results and importance rankings. m is the number of features at the highest accuracy.

### 3.2. Classification and Accuracy Evaluation

The results revealed that the average OA of the classification accuracy from 1987 to 2020 after feature optimization was 96.0%, with a Kappa coefficient of 0.954 (Figure 5). After feature optimization, among different wetland types, the average UA and PA of lake/constructed wetlands reached 96.4% and 92.8%, respectively. For rivers, the optimized mean UA and PA increased to 92.2% and 93.2%, respectively. Marshes demonstrated the highest classification accuracy, with their optimized mean UA and PA reaching 96.4% and 95.2%, respectively (Figure 6). Taken collectively, these results demonstrate that the classification method exhibits high accuracy and fulfills the requirements for wetland classification in the study area.

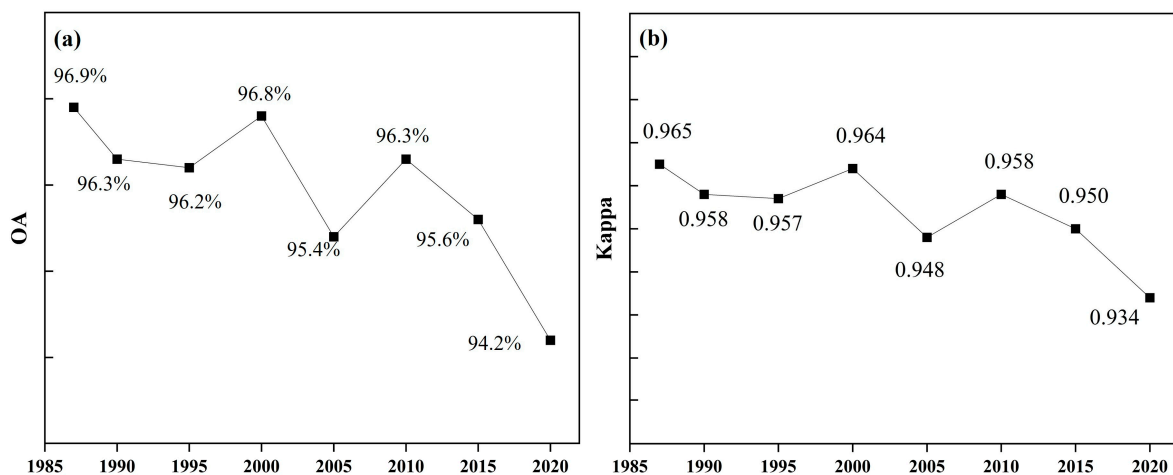


Figure 5. The figures of OA (a) and Kappa coefficient (b) after feature optimization from 1987 to 2020.

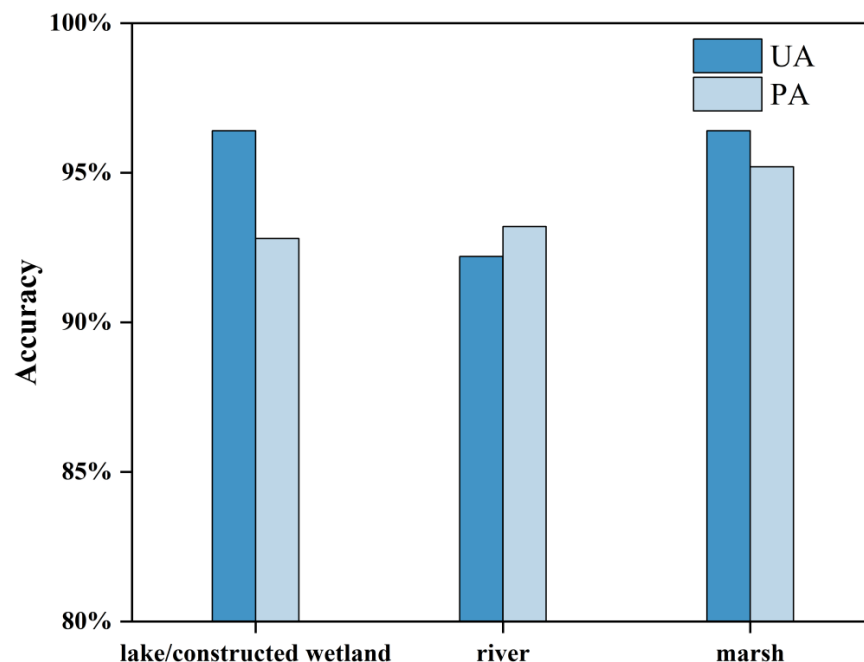


Figure 6. Average user accuracy (UA) and producer accuracy (PA) of different wetland types from 1987 to 2020.

Utilizing the automatically extracted wetland and non-wetland data in Gansu Province, this study further differentiated constructed wetlands and lakes using auxiliary datasets. Ultimately, wetlands were categorized into four major types: lakes, rivers, constructed wetlands (including reservoirs and ponds), and marshes (Figure 7). Overall, wetlands

occupy a relatively small area in Gansu Province, and their distribution is scattered. Among the wetland types, marshes dominate in terms of both area and distribution, with the highest concentration observed in southern Gansu Province (Figure 7).

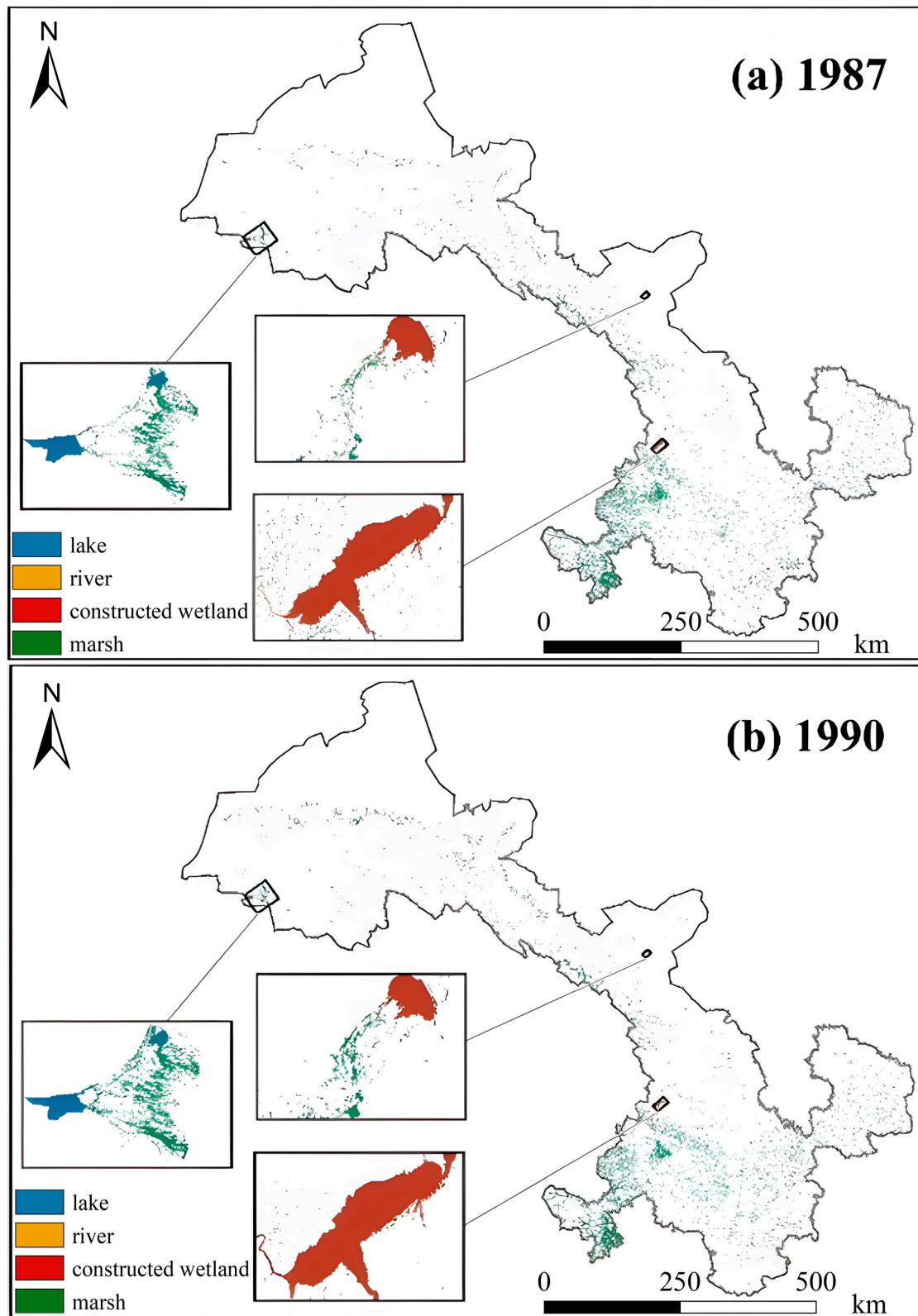


Figure 7. Cont.

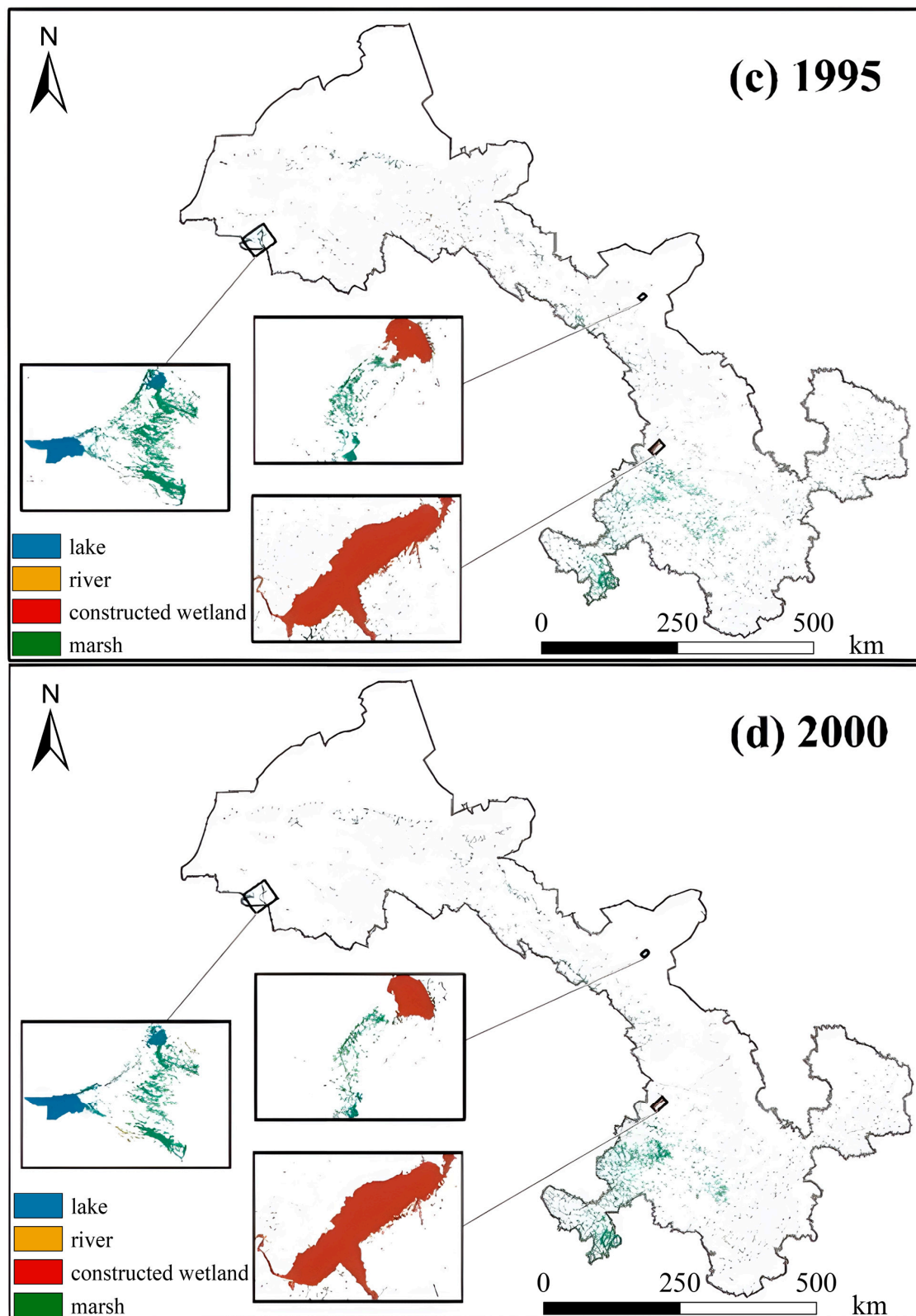


Figure 7. Cont.

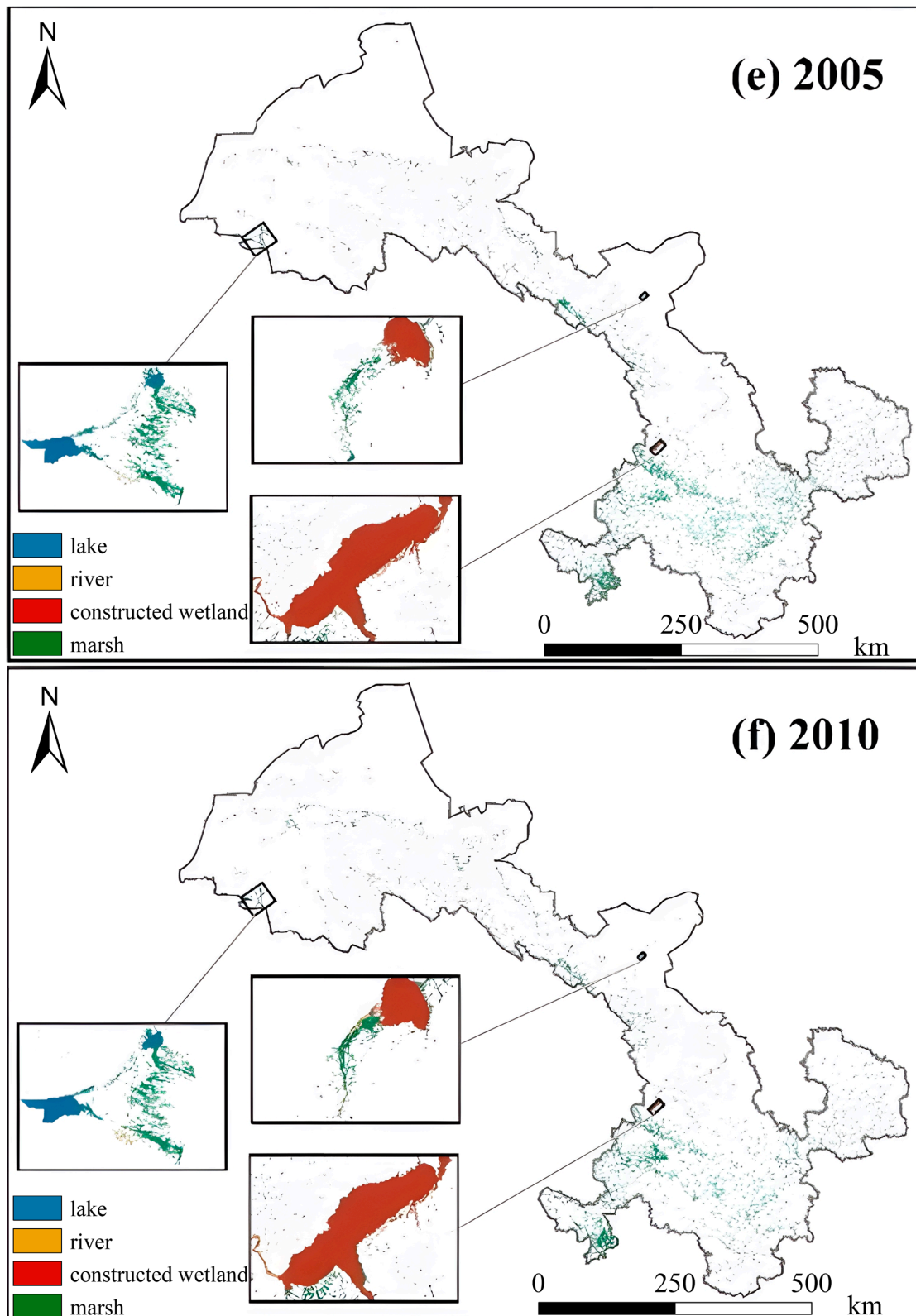
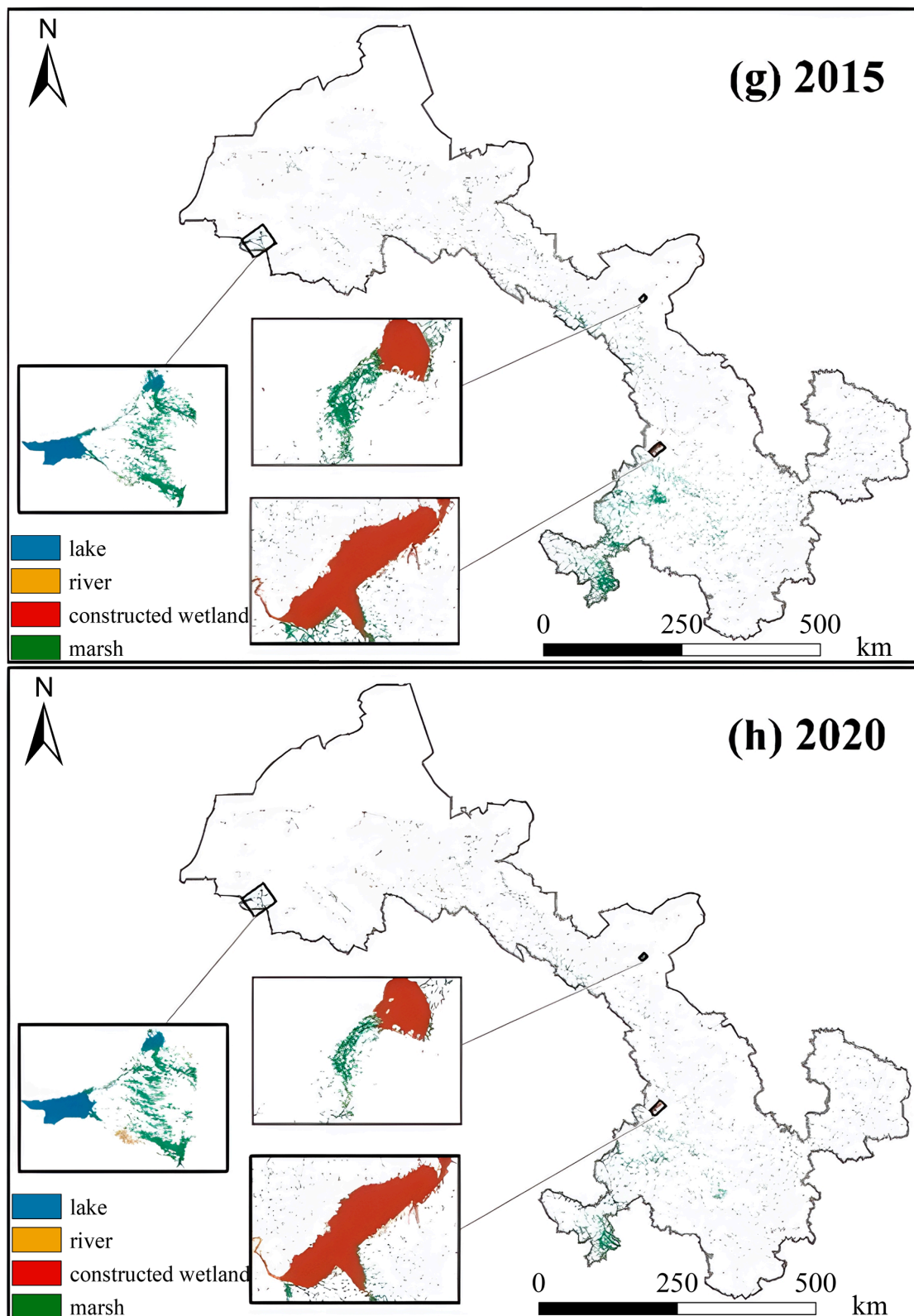


Figure 7. Cont.





**Figure 7.** Wetland classification map and local enlargement map of Gansu Province from 1987 to 2020.

### 3.3. Analysis of Changes in Wetland Resources in Gansu Province

#### 3.3.1. Characteristics of Changes in Wetland Area

From 1987 to 2020, the wetland area in Gansu Province exhibited a fluctuating decreasing trend overall (Figure 8), with a total decrease of 4536.86 km<sup>2</sup> by the end of 2020. The average

annual rate of change (K) was  $-0.91\%$ . However, on a phased basis, there was an increase in the area between 2000–2005 and 2010–2015, particularly notable in 2000–2005, which experienced the highest rate of increase ( $5.83\%$ ) and reached a maximum area of  $16,181.25 \text{ km}^2$ .

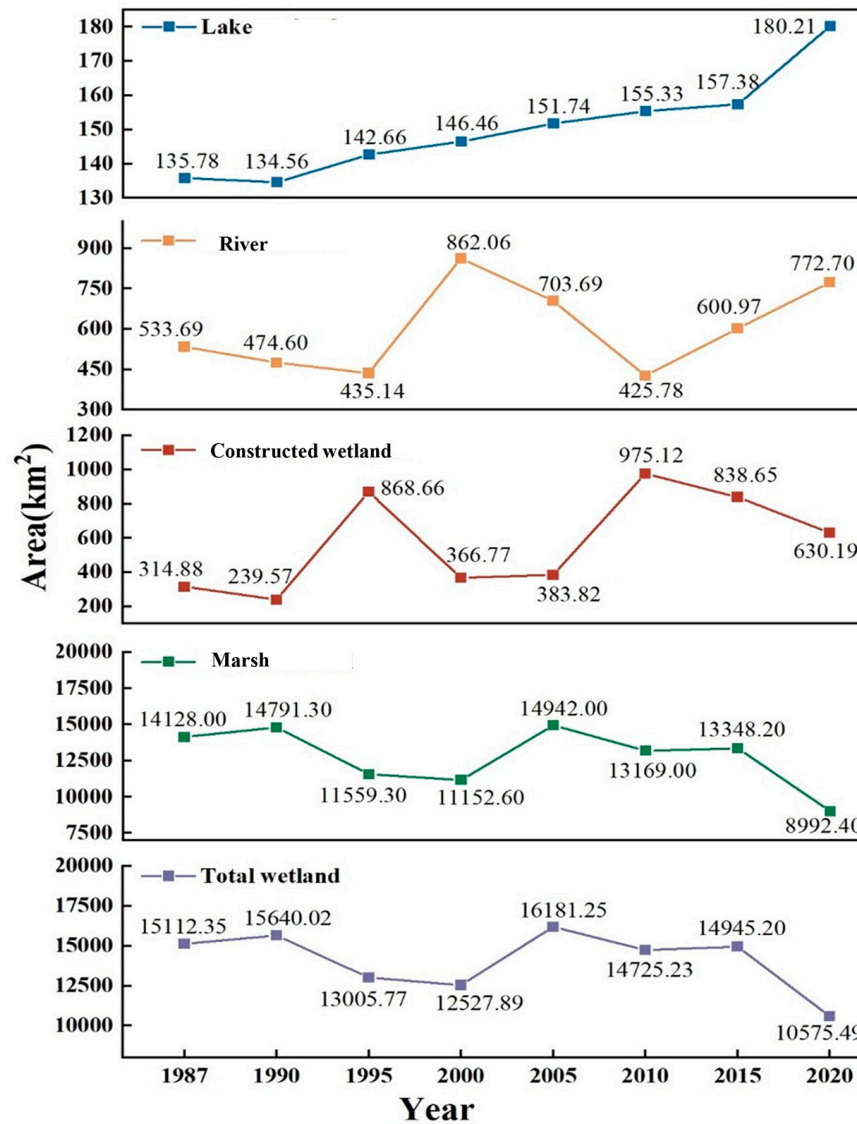


Figure 8. Change in area of different wetland types from 1987 to 2020.

The area of lakes showed a slight increase except for a minor decrease in 1990, rising from  $135.78 \text{ km}^2$  to  $180.21 \text{ km}^2$ , with the highest annual rate of change (K) recorded during the 2015–2020 period at  $2.90\%$ . By the end of 2020, the area of rivers and constructed wetlands had expanded by  $239.02 \text{ km}^2$  and  $315.31 \text{ km}^2$ , respectively, with growth rates of  $44.79\%$  and  $100.14\%$ . However, the trends of these two types of wetlands diverged, with rivers experiencing the highest K from 1995 to 2000 ( $19.62\%$ ) and reaching a peak area of  $862.06 \text{ km}^2$  in 2000. Constructed wetlands had the highest K of  $52.52\%$  during the 1990–1995 phase, but their area peaked at  $975.12 \text{ km}^2$  in 2010 and has been decreasing since then. By the end of 2020, the area of marshes had been reduced by a total of  $5135.60 \text{ km}^2$ , a reduction rate of  $36.35\%$ . The largest decrease occurred between 2015 and 2020, amounting to  $4355.80 \text{ km}^2$ , with an average rate of change of  $-6.53\%$ . Notably, the average rate of change in the area of increased marshes from 2000 to 2005 was  $6.80\%$ , exceeding the average rate of decrease from 2015 to 2020.

### 3.3.2. Characteristics of Changes in Wetland Types

Between 1987 and 2020, transfers in and out between lakes and all other land use types were minimal. Their main sources of the increase were from marshes, grasslands, and bare lands, resulting in a maximum net increase of only 20.71 km<sup>2</sup>. The reduction in rivers was primarily attributed to losses to grasslands and forests, resulting in net reductions of 130.66 km<sup>2</sup> and 121.97 km<sup>2</sup>, respectively. The increase in river areas was primarily due to the conversion of grassland, bare area, and permanent snow. Additionally, the directions of transfer into and out of constructed wetlands paralleled that of rivers. Their main sources of increase were predominantly grassland and bare area, with bare area undergoing the most conversion, resulting in a net increase in area of 208 km<sup>2</sup>. The decrease in the area of constructed wetlands is primarily converted to grassland (77.60 km<sup>2</sup>). Compared with the other three wetland types, the marsh type underwent the most significant change. Specifically, a substantial portion of its diminished area was converted into grassland and forest, resulting in net reduction rates of 31.17% and 32.44%, respectively. In addition, grassland has the largest contributing role in the increase in the area of marshes, amounting to 2876.34 km<sup>2</sup>, followed by agriculture, which transformed 1151.19 km<sup>2</sup> (Table 6).

**Table 6.** Land use area transfer matrix for Gansu Province from 1987 to 2020 (units: km<sup>2</sup>).

1987 \ 2020	Lake	River	Constructed Wetland	Marsh	Grassland	Forest	Agriculture	Settlement	Bare Area	Permanent Snow
Lake	120.26	0.00	0.00	20.71	12.95	0.25	1.55	0.06	15.04	9.36
River	0.02	120.13	6.80	26.23	216.40	64.29	10.42	1.02	200.70	124.09
Constructed wetland	0.00	39.09	183.38	16.84	103.90	14.08	22.54	0.10	208.00	42.09
Marsh	5.17	22.80	5.85	4057.94	2876.34	337.40	1151.19	12.70	427.12	94.72
Grassland	5.13	130.66	77.60	4404.40	97,126.30	2447.67	4552.63	57.24	35,915.00	3262.84
Forest	0.25	121.97	6.37	4583.36	203,86.80	48,353.40	194.68	0.82	147.84	56.89
Agriculture	0.81	12.68	1.94	795.74	3893.70	46.62	12,496.80	29.29	3952.90	24.34
Settlement	1.16	14.28	0.89	168.97	634.93	17.89	928.67	822.54	448.04	0.09
Bare area	2.71	22.62	10.40	11.61	1583.74	3.13	68.84	1.12	158,971.00	1999.61
Permanent snow	0.21	48.85	21.58	39.94	541.89	26.77	0.62	0.00	1018.63	4323.35

## 4. Discussion

### 4.1. Influencing Factors of Wetland Classification

In this study, the RFE method reduced the number of features by half or even more while simultaneously improving OA, kappa, UA, or PA. This is sufficient evidence to demonstrate that feature optimization can aid researchers in integrating effective information from multi-source data and enhancing classification accuracy. This finding aligns with results from previous studies in urban, wetland, and other domains [36].

Elevation, among the topographic features, exerted the most significant influence on wetland classification in the study area. These findings align with numerous previous studies [74,75], suggesting that topography largely determines wetland presence, with low-lying or flat conditions favoring wetland development. Furthermore, the selection of interannual synthetic images in this study limited the consideration of differences in vegetation or water body phenology. Hence, the inclusion of phenological features aided in wetland vegetation differentiation while effectively reducing classification uncertainty [66]. However, most prior studies focused solely on commonly used indicators like NDVI and NDWI, neglecting others like MNDWI and LSWI, which proved more influential in this study's classification. Compared to other features, texture features exhibited the least significant effect, likely due to the absence of distinctive geometries and regularities among wetland types in the study area. Overall, while each feature's contribution to classification was varied, the efficacy of utilizing multi-source information for wetland classification and extraction remains indisputable. Moreover, researchers increasingly apply this method to extract other features [76,77].

In addition to features, image and sample quality also play a critical role in the accuracy of wetland extraction [78]. It is well known that Landsat series images are affected by the temporal resolution and weather interference, among other factors, and the available Landsat data often lack continuous high-quality images, which poses challenges, particularly for large study areas [79]. Therefore, this study partially addressed this issue by synthesizing median images from three consecutive years, utilizing high-quality images, to some extent mitigating the abovementioned issues. Extracting land information from long time series data through purely visual interpretation not only proves time-consuming and labor-intensive but also involves complex phenomena such as homogeneous spectra, which are challenging to differentiate on 30 m resolution images [80]. In this study, the method of combining numerous auxiliary data and satellite images to construct samples improved the accuracy of the sample data, thereby resulting in more precise classification outcomes.

#### 4.2. Changes in Wetlands

From 1987 to 2020, the overall trend of wetland area in Gansu Province has been decreasing, primarily attributed to population growth, long-term overconsumption of resources, and wetland development, along with increased pollution [37,41–43]. However, it is noteworthy that the wetland area experienced a significant increase from 2000 to 2005, likely due to a series of ecological conservation measures, such as the addition of a large and medium-sized reservoir (Table 7). In 2000, the Chinese government initiated substantial investments in the protection and restoration of natural capital and formulated the National Action Plan for Wetland Conservation [81]. Additionally, Gansu Province enacted the Gansu Provincial Wetland Protection Regulations in 2003 to legally safeguard wetland ecology. While the number of reservoirs has increased during the 2015–2020 period, the total area of wetlands has decreased significantly as a result of strong economic development, particularly in marshes. This is due to the extensive conversion of marshes to grasslands and forests. This finding aligns with the results of an analysis comparing wetland data from the Second National Wetland Resources Survey and the Third National Land Survey in Gansu Province [82]. The area of lakes continues to increase but experiences little fluctuation, as most of the increase results from the conversion of marshes surrounding the lakes. With the development of the western region in 2000, water scarcity issues have worsened, especially in the northern region [83], consistent with the significant reduction in rivers since 2000 observed in this study. Additionally, historical data indicate that from 1976 to 1995, the water conservancy department of Gansu Province conducted surveys and constructed large-scale water conservancy projects, for example, the establishment of the Bikou Reservoir and the Huangcheng Reservoir, leading to a continuous increase in the area of constructed wetlands during this period. However, later fluctuations in the area of constructed wetlands occurred due to population growth and economic development, rendering the trend unstable.

**Table 7.** Auxiliary data related to reservoirs in Gansu Province.

Year	Average Annual Precipitation (mm)	GDP (Billion RMB)	Reservoir Capacity (Billion m <sup>3</sup> )	Large and Medium-Sized Reservoirs
1991	233.5	271.39	31.2852	16
1995	250.3	557.76	40.32751	28
2000	247.1	1052.88	35.2401	28
2005	281.2	1864.63	39.203	29
2010	263.9	3943.73	39.048	29
2015	251.4	6556.55	36.105	29
2020	317.6	9016.7	46.33	33

Notes: A total reservoir capacity  $\geq 1$  billion m<sup>3</sup> is considered large, and a total reservoir capacity of 1 to 0.1 billion m<sup>3</sup> is considered medium.

### 4.3. Implications and Improvements of Current Research

Wetland change unfolds over the long term, influenced by various factors leading to distinct trends in different wetlands [84]. Due to the heterogeneity and fragmentation of wetland landscapes and the spectral similarities between wetland classes, there are significant challenges to achieving accurate wetland mapping [85]. Nonetheless, wetland data with long-term frequency and high spatial resolution can elucidate the enduring dynamic changes of wetlands and play a pivotal role in delineating ecological protection red lines and promoting the sustainable development of wetlands [86]. Despite wetland studies being common in certain watersheds or protected areas within Gansu Province [41–43], there is still a significant lack of comprehensive research on multitype wetlands across the entire Gansu Province. The long-time series wetland dataset generated in this study, with a spatial resolution of 30 m, offers a robust depiction of the evolution of various wetland types in Gansu Province over the past years. This dataset constitutes a valuable resource, providing data support and a scientific basis for the sustainable management of wetland ecosystems in northwest China.

Landsat data were selected for their extensive time series, which is essential for the extraction of wetland information over extended periods. The 30 m resolution of Landsat data, however, may lead to fragmented wetland information, thereby restricting the extraction of small water bodies and slender streams [80,81]. Additionally, there is still a persistent issue where certain roads, buildings, and water bodies can be easily mistaken for wetlands due to spectral similarity [87]. To address these challenges, an increasing number of studies have turned towards multi-source remote sensing image fusion techniques to enhance the accuracy of wetland mapping [88]. While this approach partially addresses the temporal and spatial resolution discrepancies of satellite sensors, its fusion accuracy is subject to notable errors due to technological complexity, radiometric differences, geometrical misalignment, and mixed pixels [89]. Thus far, employing multi-source remote sensing image fusion for wetland information extraction remains challenging, particularly for large-scale and multi-temporal analyses. Additionally, the automated classification method employed in this study struggled to differentiate lakes from reservoirs, necessitating manual secondary classification; however, the capacity for manual classification is limited. In conclusion, future research endeavors will focus on enhancing both image accuracy and classification methods to improve wetland classification accuracy. With advancements in remote sensing and computer technology, achieving automated, long-term, large-scale, high-precision, and fine wetland classification is imminent, offering significant implications for wetland ecological protection and restoration.

## 5. Conclusions

High-quality wetland data are indispensable for Gansu Province, given its scarce water resources and fragile ecological environment. In this study, leveraging the GEE cloud platform, we combined the recursive feature elimination method with the random forest model and integrated sample points with multi-source data features to extract wetland information in Gansu Province from 1987 to 2020. The average OA achieved 96.0%, with a kappa coefficient of 0.954. Additionally, we derived quantitative data on the changes in wetland area and type transformations and discussed the factors influencing classification uncertainty and the drivers of wetland changes. Our findings indicate that integrating multi-source data features enhances the distinguishability of wetland categories. Nevertheless, an excess of features can introduce noise, underscoring the importance of feature optimization for the enhancement of classification accuracy. Topographic features, particularly elevation, had the most impact, while textural features had the least impact. From 1987 to 2020, the overall wetland area in Gansu Province decreased by 4543.86 km<sup>2</sup>, with different wetland types exhibiting varied changes at each time stage. There has been a slight increase in the area of lakes, primarily sourced from marshes. The areas of river and constructed wetlands fluctuated, with increases primarily due to the conversion of grasslands and bare land. The most notable change was in marshes, which decreased by 5135.60 km<sup>2</sup> between 1987 and 2020, primarily due to conversion to grassland and forest. The wetland classification method employed in this study yielded

positive results in Gansu Province, offering a framework for extracting large-scale, long-term wetland data in other regions and providing scientific support for wetland protection and sustainable management.

**Author Contributions:** Conceptualization, J.Z., X.L. and Y.Q.; methodology, X.L. and Y.Q.; software, X.L., Y.Q., Y.F. and S.C.; data curation J.Z. and Y.Q.; formal analysis, X.L. and Y.Q.; writing—original draft preparation, J.Z., X.L. and Y.Q.; writing—review and editing, J.Z.; visualization, Y.Q., Y.F. and S.C.; funding acquisition, J.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** The raw data supporting the conclusions of this article will be made available by the authors on request.

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

## References

- Niu, Z.; Zhang, H.; Wang, X.; Yao, W.; Zhou, D.; Zhao, K.; Gong, P. Mapping wetland changes in China between 1978 and 2008. *Chin. Sci. Bull.* **2012**, *57*, 2813–2823. [[CrossRef](#)]
- Kovács, G.M.; Horion, S.; Fensholt, R. Characterizing ecosystem change in wetlands using dense earth observation time series. *Remote Sens. Environ.* **2022**, *281*, 113267. [[CrossRef](#)]
- Zhang, M.; Lin, H. Wetland classification using parcel-level ensemble algorithm based on Gaofen-6 multispectral imagery and Sentinel-1 dataset. *J. Hydrol.* **2022**, *606*, 127462. [[CrossRef](#)]
- Zhang, W.; Hu, B.X.; Brown, G.; Meyer, S. Beaver pond identification from multi-temporal and multi-sourced remote sensing data. *Geo-Spat. Inf. Sci.* **2023**, *27*, 953–967. [[CrossRef](#)]
- Sahraei, R.; Ghorbanian, A.; Kanani-Sadat, Y.; Jamali, S.; Homayouni, S. Mangrove plantation suitability mapping by integrating multi criteria decision making geospatial approach and remote sensing data. *Geo-Spat. Inf. Sci.* **2023**, *27*, 1290–1308. [[CrossRef](#)]
- Costanza, R.; de Groot, R.; Sutton, P.; van der Ploeg, S.; Anderson, S.J.; Kubiszewski, I.; Farber, S.; Turner, R.K. Changes in the global value of ecosystem services. *Glob. Environ. Change* **2014**, *26*, 152–158. [[CrossRef](#)]
- Peng, K.; Jiang, W.; Hou, P.; Wu, Z.; Ling, Z.; Wang, X.; Mao, D. Continental-scale wetland mapping: A novel algorithm for detailed wetland types classification based on time series Sentinel-1/2 images. *Ecol. Indic.* **2023**, *148*, 110113. [[CrossRef](#)]
- Davidson, N.C.; Finlayson, C.M. Extent, regional distribution and changes in area of different classes of wetland. *Mar. Freshw. Res.* **2018**, *69*, 1525. [[CrossRef](#)]
- Calhoun, A.J.K.; Mushet, D.M.; Bell, K.P.; Boix, D.; Fitzsimons, J.A.; Isselin-Nondedeu, F. Temporary wetlands: Challenges and solutions to conserving a “disappearing” ecosystem. *Biol. Conserv.* **2017**, *211*, 3–11. [[CrossRef](#)]
- Slagter, B.; Tsendbazar, N.E.; Vollrath, A.; Reiche, J. Mapping wetland characteristics using temporally dense Sentinel-1 and Sentinel-2 data: A case study in the St. Lucia wetlands, South Africa. *Int. J. Appl. Earth Obs. Geoinf.* **2020**, *86*, 102009. [[CrossRef](#)]
- 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](#)]
- Chen, J.; Chen, J.; Liao, A.; Cao, X.; Chen, L.; Chen, X.; He, C.; Han, G.; Shu, P.; Lu, M.; et al. Global land cover mapping at 30 m resolution: A POK-based operational approach. *ISPRS J. Photogramm.* **2015**, *103*, 7–27. [[CrossRef](#)]
- 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.
- Jiang, W.; Zhang, Z.; Ling, Z.; Deng, Y. Experience and future research trends of wetland protection and restoration in China. *J. Geogr. Sci.* **2024**, *34*, 229–251. [[CrossRef](#)]
- Pekel, J.F.; Cottam, A.; Gorelick, N.; Belward, A.S. High-resolution mapping of global surface water and its long-term changes. *Nature* **2016**, *540*, 418–422. [[CrossRef](#)]
- Xia, Q.; Jia, M.; He, T.; Xing, X.; Zhu, L. Effect of tide level on submerged mangrove recognition index using multi-temporal remotely-sensed data. *Ecol. Ind.* **2021**, *131*, 108169. [[CrossRef](#)]
- Lehner, B.; Döll, P. Development and validation of a global database of lakes, reservoirs and wetlands. *J. Hydrol.* **2004**, *296*, 1–22. [[CrossRef](#)]
- L’Heureux, A.; Grolinger, K.; Elyamany, H.F.; Capretz, M.A.M. Machine learning with big data: Challenges and approaches. *IEEE Access* **2017**, *5*, 7776–7797. [[CrossRef](#)]
- Zhao, F.; Feng, S.; Xie, F.; Zhu, S.; Zhang, S. Extraction of long time series wetland information based on Google Earth Engine and random forest algorithm for a plateau lake basin—A case study of Dianchi Lake, Yunnan Province, China. *Ecol. Indic.* **2023**, *146*, 109813. [[CrossRef](#)]
- Han, X.; Pan, J.; Devlin, A.T. Remote sensing study of wetlands in the Pearl River Delta during 1995–2015 with the support vector machine method. *Front. Earth Sci.* **2018**, *12*, 521–531. [[CrossRef](#)]

21. Kesikoglu, M.H.; Atasever, U.H.; Dadaser-Celik, F.; Ozkan, C. Performance of ANN, SVM and MLH techniques for land use/cover change detection at Sultan Marshes wetland, Turkey. *Water Sci. Technol.* **2019**, *80*, 466–477. [[CrossRef](#)]
22. Zhang, J.; Chu, L.; Zhang, Z.; Zhu, B.; Liu, X.; Yang, Q. Evolution of Small and Micro Wetlands and Their Driving Factors in the Yangtze River Delta—A Case Study of Wuxi Area. *Remote Sens.* **2023**, *15*, 1152. [[CrossRef](#)]
23. Mahdavi, S.; Salehi, B.; Granger, J.; Amani, M.; Brisco, B.; Huang, W. Remote sensing for wetland classification: A comprehensive review. *Gisci. Remote Sens.* **2018**, *55*, 623–658. [[CrossRef](#)]
24. Gounaridis, D.; Chorianopoulos, I.; Symeonakis, E.; Koukoulas, S. A random forest-cellular automata modelling approach to explore future land use/cover change in Attica (Greece), under different socio-economic realities and scales. *Sci. Total Environ.* **2019**, *646*, 320–335. [[CrossRef](#)]
25. Peng, K.F.; Jiang, W.G.; Deng, Y.; Liu, Y.H.; Wu, Z.F.; Chen, Z. Simulating wetland changes under different scenarios based on integrating the random forest and CLUE-S models: A case study of Wuhan Urban Agglomeration. *Ecol. Indic.* **2020**, *117*, 106671. [[CrossRef](#)]
26. Guo, X.Y.; Bian, Z.X.; Wang, S.; Wang, Q.B.; Zhang, Y.F.; Zhou, J.; Lin, L. Prediction of the spatial distribution of soil arthropods using a random forest model: A case study in Changtu County, Northeast China. *Agric. Ecosyst. Environ.* **2020**, *292*, 106818. [[CrossRef](#)]
27. Ma, H.; Gao, X.; Gu, X. Random forest classification of Landsat 8 imagery for the complex terrain area based on the combination of spectral, topographic and texture information. *Int. J. Geogr. Inf. Sci.* **2019**, *21*, 359–371.
28. Lawrence, R.L.; Moran, C.J. The America View classification methods accuracy comparison project: A rigorous approach for model selection. *Remote Sens. Environ.* **2015**, *170*, 115–120. [[CrossRef](#)]
29. Amani, M.; Salehi, B.; Mahdavi, S.; Granger, J.; Brisco, B. Wetland classification in Newfoundland and Labrador using multi-source SAR and optical data integration. *Gisci. Remote Sens.* **2017**, *54*, 779–796. [[CrossRef](#)]
30. Millard, K.; Richardson, M. Wetland mapping with LiDAR derivatives, SAR polarimetric decompositions, and LiDAR-SAR fusion using a random forest classifier, Canadian. *Remote Sens.* **2013**, *39*, 290–307.
31. Yang, H.; Liu, X.; Chen, Q.; Cao, Y. Mapping Dongting Lake Wetland Utilizing Time Series Similarity, Statistical Texture, and Superpixels with Sentinel-1 SAR Data. *IEEE J. Sel. Top.* **2022**, *15*, 8235–8244. [[CrossRef](#)]
32. Fluet-Chouinard, E.; Lehner, B.; Rebelo, L.M.; Papa, F.; Hamilton, S.K. Development of a global inundation map at high spatial resolution from topographic downscaling of coarse-scale remote sensing data. *Remote Sens. Environ.* **2015**, *158*, 348–361. [[CrossRef](#)]
33. Yamada, S.; Neshatian, K. Comparison of embedded and wrapper approaches for feature selection in support vector machines. In *PRICAI 2019: Trends in Artificial Intelligence: 16th Pacific Rim International Conference on Artificial Intelligence, Cuvu, Ya-nuca Island, Fiji, 26–30 August 2019*; Proceedings, Part II 16; Springer International Publishing: Berlin/Heidelberg, Germany, 2019; pp. 149–161.
34. Cheng, K.; Wang, J.; Yan, X. Mapping forest types in China with 10 m resolution based on spectral–spatial–temporal features. *Remote Sens.* **2021**, *13*, 973. [[CrossRef](#)]
35. Fu, B.; Xie, S.; He, H.; Zuo, P.; Sun, J.; Liu, L.; Gao, E. Synergy of multi-temporal polarimetric SAR and optical image satellite for mapping of marsh vegetation using object-based random forest algorithm. *Ecol. Indic.* **2021**, *131*, 108173. [[CrossRef](#)]
36. Xie, S.; Fu, B.; Li, Y.; Liu, Z.; Zuo, P.; Lan, F.; Fan, D. Classification method on marsh wetlands in Honghe National Nature Reserve based on multi-dimensional remote sensing images. *Wetl. Sci.* **2021**, *19*, 1–16.
37. Wang, J.; Liu, H.; Liu, P.; Chang, G. An evaluation of the spatial rationality of provincial territory: A case study of Gansu Province. *Environ. Sci. Pollut.* **2023**, *30*, 18268–18284. [[CrossRef](#)]
38. Xu, W.; Fan, X.; Ma, J.; Pimm, S.L.; Kong, L.; Zeng, Y.; Ouyang, Z. Hidden loss of wetlands in China. *Curr. Biol.* **2019**, *29*, 3065–3071. [[CrossRef](#)]
39. Zhu, X.; Jiao, L.; Wu, X.; Du, D.; Wu, J.; Zhang, P. Ecosystem health assessment and comparison of natural and constructed wetlands in the arid zone of northwest China. *Ecol. Indic.* **2023**, *154*, 110576. [[CrossRef](#)]
40. Jiang, P.; Cheng, L.; Li, M.; Zhao, R.; Huang, Q. Analysis of landscape fragmentation processes and driving forces in wetlands in arid areas: A case study of the middle reaches of the Heihe River, China. *Ecol. Indic.* **2014**, *46*, 240–252. [[CrossRef](#)]
41. Wen, X.; Wu, X.; Gao, M. Spatiotemporal variability of temperature and precipitation in Gansu Province (Northwest China) during 1951–2015. *Atmos. Res.* **2017**, *197*, 132–149. [[CrossRef](#)]
42. Zhao, X.; Wang, J.; Su, J.; Sun, W.; Meng, H. Research on a Biodiversity Conservation Value Assessment Method Based on Habitat Suitability of Species: A Case Study in Gansu Province, China. *Sustainability* **2021**, *13*, 3007. [[CrossRef](#)]
43. Wang, H.X.; Liu, X.N.; Li, C.B.; Ren, Z.C.; Pan, D.R.; Wei, J.Q. Spatial temporal Distribution of Precipitation in Gansu Province Last 42 Years. *Chin. J. Agrometeorol.* **2013**, *34*, 384–389.
44. Chen, J.; Chen, J. Globe Land 30: Operational global land cover mapping and big-data analysis. *Sci. China Earth Sci.* **2018**, *61*, 1533–1534. [[CrossRef](#)]
45. Messenger, M.L.; Lehner, B.; Grill, G.; Nedeva, I.; Schmitt, O. Estimating the volume and age of water stored in global lakes using a geo-statistical approach. *Nat. Commun.* **2016**, *7*, 13603. [[CrossRef](#)]
46. Allen, G.H.; Pavelsky, T.M. Global extent of rivers and streams. *Science* **2018**, *361*, 585–588. [[CrossRef](#)]
47. Mao, D.; Wang, Z.; Du, B.; Li, L.; Tian, Y.; Jia, M. National wetland mapping in China: A new product resulting from object-based and hierarchical classification of Landsat 8 OLI images. *ISPRS J. Photogramm.* **2020**, *164*, 11–25. [[CrossRef](#)]

48. Mulligan, M.; Soesbergen, A.; Sáenz, L. GOODD, a global dataset of more than 38,000 georeferenced dams. *Sci. Data* **2020**, *7*, 31. [[CrossRef](#)]
49. Lippitt, C.D.; Rogan, J.; Li, Z.; Eastman, J.R.; Jones, T.G. Mapping selective logging in mixed deciduous forest. *Photogramm. Eng. Remote Sens.* **2008**, *74*, 1201–1211. [[CrossRef](#)]
50. Rouse, J.W.; Haas, R.H.; Deering, D.W.; Schell, J.A.; Harlan, J.C. Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation. [Great Plains Corridor]. *Environ. Sci.* **1973**. Available online: <https://ntrs.nasa.gov/api/citations/19730017588/downloads/19730017588.pdf> (accessed on 18 September 2024).
51. Huete, A.; Didan, K.; Miura, T.; Rodriguez, E.P.; Gao, X.; Ferreira, L.G. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ.* **2002**, *83*, 195–213. [[CrossRef](#)]
52. Jackson, R.D.; Huete, A.R. Interpreting vegetation indices. *Prev. Vet. Med.* **1991**, *11*, 185–200. [[CrossRef](#)]
53. Huete, A.R. A soil-adjusted vegetation index (SAVI). *Remote Sens. Environ.* **1988**, *25*, 295–309. [[CrossRef](#)]
54. Qi, J.; Chehbouni, A.; Huete, A.R.; Kerr, Y.H.; Sorooshian, S. A modified soil adjusted vegetation index. *Remote Sens. Environ.* **1994**, *48*, 119–126. [[CrossRef](#)]
55. McFeeters, S.K. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *Int. J. Remote Sens.* **1996**, *17*, 1425–1432. [[CrossRef](#)]
56. Wang, S.; Baig, M.H.A.; Zhang, L.; Jiang, H.; Ji, Y.; Zhao, H.; Tian, J. A simple enhanced water index (EWI) for percent surface water estimation using Landsat data. *IEEE J. Sel. Top. Appl. Earth-Observ. Remote Sens.* **2015**, *8*, 90–97. [[CrossRef](#)]
57. Xu, H. A study on information extraction of water body with the modified normalized difference water index (MNDWI). *Natl. Remote Sens. Bull.* **2005**, *9*, 589–595.
58. Xiao, X.; Boles, S.; Liu, J.; Zhuang, D.; Froking, S.; Li, C.; Moore III, B. Mapping paddy rice agriculture in southern China using multi-temporal MODIS images. *Remote Sens. Environ.* **2005**, *95*, 480–492. [[CrossRef](#)]
59. Feyisa, G.L.; Meilby, H.; Fensholt, R.; Proud, S.R. Automated Water Extraction Index: A new technique for surface water mapping using Landsat imagery. *Remote Sens. Environ.* **2014**, *140*, 23–35. [[CrossRef](#)]
60. Ji, L.; Geng, X.; Sun, K.; Zhao, Y.; Gong, P. Target detection method for water mapping using Landsat 8 OLI/TIRS imagery. *Water* **2015**, *7*, 794–817. [[CrossRef](#)]
61. Zha, Y.; Gao, J.; Ni, S. Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *Int. J. Remote Sens.* **2003**, *24*, 583–594. [[CrossRef](#)]
62. Hall, D.K.; Riggs, G.A.; Salomonson, V.V. Development of methods for mapping global snow cover using moderate resolution imaging spectroradiometer data. *Remote Sens. Environ.* **1995**, *54*, 127–140. [[CrossRef](#)]
63. Huang, C.; Wylie, B.; Yang, L.; Homer, C.; Zylstra, G. Derivation of a tasseled cap transformation based on Landsat 7 at-satellite reflectance. *Int. J. Remote Sens.* **2002**, *23*, 1741–1748. [[CrossRef](#)]
64. Ashok, A.; Rani, H.P.; Jayakumar, K.V. Monitoring of dynamic wetland changes using NDVI and NDWI based landsat imagery. *Remote Sens. Appl. Soc. Environ.* **2021**, *23*, 100547.
65. Tassi, A.; Vizzari, M. Object-oriented lulc classification in google earth engine combining snic, glcm, and machine learning algorithms. *Remote Sens.* **2020**, *12*, 3776. [[CrossRef](#)]
66. Bwangoy, J.R.B.; Hansen, M.C.; Roy, D.P.; De Grandi, G.; Justice, C.O. Wetland mapping in the Congo Basin using optical and radar remotely sensed data and derived topographical indices. *Remote Sens. Environ.* **2010**, *114*, 73–86. [[CrossRef](#)]
67. Gregorutti, B.; Michel, B.; Saint-Pierre, P. Correlation and variable importance in random forests. *Stat. Comput.* **2017**, *27*, 659–678. [[CrossRef](#)]
68. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
69. Tatsumi, K.; Yamashiki, Y.; Torres, M.A.C.; Taipei, C.L.R. Crop classification of upland fields using Random forest of time-series Landsat 7 ETM+ data. *Comput. Electron. Agric.* **2015**, *115*, 171–179. [[CrossRef](#)]
70. Immitzer, M.; Atzberger, C.; Koukal, T. Tree species classification with random forest using very high spatial resolution 8-band worldview-2 satellite data. *Remote Sens.* **2012**, *4*, 2661–2693. [[CrossRef](#)]
71. Berhane, T.M.; Lane, C.R.; Wu, Q.; Anenkhonov, O.A.; Chepinoga, V.V.; Autrey, B.C.; Liu, H. Comparing pixel and object-based approaches in effectively classifying wetland-dominated landscapes. *Remote Sens.* **2017**, *10*, 46. [[CrossRef](#)]
72. Li, H.; Wang, J.; Zhang, J.; Qin, F.; Hu, J.; Zhou, Z. Analysis of characteristics and driving factors of wetland landscape pattern change in Henan Province from 1980 to 2015. *Land* **2021**, *10*, 564. [[CrossRef](#)]
73. Hu, W.; Li, G.; Li, Z. Spatial and temporal evolution characteristics of the water conservation function and its driving factors in regional lake wetlands—Two types of homogeneous lakes as examples. *Ecol. Indic.* **2021**, *130*, 108069. [[CrossRef](#)]
74. Chasmer, L.; Hopkinson, C.; Montgomery, J.; Petrone, R. A physically based terrain morphology and vegetation structural classification for wetlands of the Boreal Plains, Alberta, Canada. *Can. J. Remote Sens.* **2016**, *42*, 521–540. [[CrossRef](#)]
75. Ding, Q.; Liu, J.; Yu, J.; Tang, L.; Guo, B. Combination of LiDAR's Multiple Attributes for Wetland Classification: A Case Study of Yellow River Delta. *Can. J. Remote Sens.* **2020**, *46*, 753–764. [[CrossRef](#)]
76. Hoffman-Hall, A.; Loboda, T.V.; Hall, J.V.; Carroll, M.L.; Chen, D. Mapping remote rural settlements at 30 m spatial resolution using geospatial data-fusion. *Remote Sens. Environ.* **2019**, *233*, 111386. [[CrossRef](#)]
77. Mohammed, I.; Marshall, M.; de Bie, K.; Estes, L.; Nelson, A. A blended census and multiscale remote sensing approach to probabilistic cropland mapping in complex landscapes. *ISPRS J. Photogramm.* **2020**, *161*, 233–245. [[CrossRef](#)]



78. Liu, J.; Chen, H.; Wang, Y. Multi-source remote sensing image fusion for ship target detection and recognition. *Remote Sens.* **2021**, *13*, 4852. [[CrossRef](#)]
79. Wu, M.; Huang, W.; Niu, Z.; Wang, C. Generating daily synthetic Landsat imagery by combining Landsat and MODIS data. *Sensors* **2015**, *15*, 24002–24025. [[CrossRef](#)]
80. Xu, Q.; Yuan, X.; Ouyang, C. Class-Aware Domain Adaptation for Semantic Segmentation of Remote Sensing Images. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–17. [[CrossRef](#)]
81. Zhang, X. Study on the Relationship between Human Capital Investment and Economic Growth in Gansu. Ph.D. Dissertation, Xiamen University, Xiamen, China, 2011. (In Chinese).
82. Zhang, H.G.; Xu, F.Y.; Chen, G.Q. Analysis on Differences between Data of Wetlands in Gansu Province from the Second National Wetland Resources Survey and Those from the Third National Land Survey. *Wetl. Sci.* **2021**, *20*, 311–316. (In Chinese)
83. Li, W.; Liu, Y.J.; Yang, Z. Preliminary strategic environmental assessment of the Great Western Development Strategy: Safeguarding ecological security for a new western China. *Environ. Manag.* **2012**, *49*, 483–501. [[CrossRef](#)]
84. Demarquet, Q.; Rapinel, S.; Dufour, S.; Hubert-Moy, L. Long-term wetland monitoring using the landsat archive: A review. *Remote Sens.* **2023**, *15*, 820. [[CrossRef](#)]
85. Wang, M.; Mao, D.; Wang, Y.; Li, H.; Zhen, J.; Xiang, H.; Wang, Z. Interannual changes of urban wetlands in China's major cities from 1985 to 2022. *ISPRS J. Photogramm.* **2024**, *209*, 383–397. [[CrossRef](#)]
86. Deng, Y.; Shao, Z.; Dang, C.; Huang, X.; Wu, W.; Zhuang, Q.; Ding, Q. Assessing urban wetlands dynamics in Wuhan and Nanchang, China. *Sci. Total Environ.* **2023**, *901*, 165777. [[CrossRef](#)] [[PubMed](#)]
87. Geng, Z.; Jiang, W.; Peng, K.; Deng, Y.; Wang, X. Wetland mapping and landscape analysis for supporting international wet-land cities: Case studies in Nanchang City and Wuhan City. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2023**, *16*, 8858–8870.
88. Zhang, M.; Zhang, H.; Yao, B.; Lin, H.; An, X.; Liu, Y. Spatiotemporal changes of wetlands in China during 2000–2015 using Landsat imagery. *J. Hydrol.* **2023**, *621*, 129590. [[CrossRef](#)]
89. Qian, H.; Bao, N.; Meng, D.; Zhou, B.; Lei, H.; Li, H. Mapping and classification of Liao River Delta coastal wetland based on time series and multi-source Gao Fen images using stacking ensemble model. *Ecol. Inform.* **2024**, *80*, 102488. [[CrossRef](#)]

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