Effects of machine learning models and spatial resolution on land cover classification accuracy in Dali County, Shaanxi, China

Yu Shi[1,2], Ning Jin[3*], Bingyan Wu[1,2], Zekun Wang[4], Shiwen Wang[2], Qiang Yu[2,5]

(1. *College of Natural Resources and Environment, Northwest A&F University, Yangling 712100, Shaanxi, China*;
2. *State Key Laboratory of Soil Erosion and Dryland Farming on the Loess Plateau, Institute of Soil and Water Conservation, Northwest A&F University, Yangling 712100, Shaanxi, China*;
3. *Department of Resources and Environment, Shanxi Institute of Energy, Jinzhong 030600, Shanxi, China*;
4. *Department of Mechanical Engineering, Auburn University, AL 36849, USA*;
5. *Key Laboratory of Water Cycle and Related Land Surface Processes, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China*)

Abstract: Land use and land cover (LULC) has undergone drastic changes with the rapid growth of the global population, economic development, and the expansion of agricultural activities. However, the uncertainty of classification algorithms and image resolution based on satellite data for land cover mapping, particularly cropland cover mapping, needs to be investigated sufficiently. In this study, the influence of different spatial-resolution images on classification results was explored by comparing the differences between four machine learning algorithms for LULC mapping. The classification results of this model were also compared with existing global land cover datasets to determine whether the model was capable of producing reliable results. According to the results of this study, the random forest (RF) classifier outperformed the support vector machine (SVM), decision tree (DT), and artificial neural network (ANN) with an overall accuracy (OA) and kappa coefficient of 81.99% and 0.78, respectively. However, SVM and ANN showed greater accuracy on the water class and unused land class, respectively. With increasing spatial resolution, RF's accuracy increased initially and then decreased when classifying images with five different spatial resolutions (30 m, 16 m, 10 m, 8 m, and 2 m). In particular, with an OA of 82.54% and a kappa coefficient of 0.78, RF performed the best on images with 8 m resolution. Additionally, the RF-based image with 8 m resolution produced a higher OA of 0.88 for cropland. Topography is the main factor that determines the classification performance of different-resolution images. The classification accuracies of RF10 m and RF30 m (10 m and 30 m resolution images, respectively, using RF) were higher (OAs of 93.59% and 94.59%, respectively) than those of the global land cover dataset (LC10 m and LC30 m, land cover images with 10 m and 30 m resolution, respectively), whose high-resolution images showed more details of the land cover. The results of this study highlight that classification algorithms and image resolution are the sources of uncertainty for land mapping. Obtaining reliable land cover mapping requires the use of appropriate classification algorithms and spatial resolution. With these results, it will be possible to develop a national land monitoring system and basic ecological climate models using LULC.

Keywords: satellite data, different spatial resolutions, machine learning, land use, land cover, random forest, uncertainty
DOI: 10.25165/j.ijabe.20251801.7252

Citation: Shi Y, Jin N, Wu B Y, Wang Z K, Wang S W, Yu Q. Effects of machine learning models and spatial resolution on land cover classification accuracy in Dali County, Shaanxi, China. Int J Agric & Biol Eng, 2025; 18(1): 245–259.

1 Introduction

Land use and land cover (LULC) has changed dramatically worldwide with the rapid growth of the global population, economic development, and the expansion of agricultural activities[1]. The effects of climate warming on the phenology of terrestrial plants change LULC[2,3]. These changes in LULC are also one of the main drivers of global environmental change[4]. Cropland is an important part of LULC and an important source of greenhouse gas emissions and carbon (e.g., rice paddy-released methane). The distribution and area of cropland directly affect grain production, economic fluctuations, and international trade, especially in agricultural countries. Moreover, cropland maps are usually crucial input data for semi-empirical and mechanism models that incorporate ecology, the environment, geography, and hydrology[5,6]. Cropland maps are further processed into multiple types of maps, such as specific crop maps[7], planting intensity maps[8], and agricultural zone maps[9]. Therefore, the cropland area and distribution can provide support for decision-making concerning food imports and exports nationwide. Detailed mapping can help guide agricultural practices to improve yield on the local scale.

Cropland has become fragmented due to human activities, socioeconomic levels, and topography, especially in Africa and Asia. It is rare for low-medium resolution satellites, such as Moderate-resolution Imaging Spectroradiometer (MODIS), to capture fine plots, resulting in the increasing uncertainty of cropland

Received date: 2021-12-08 **Accepted date:** 2023-02-26
Biographies: Yu Shi, PhD, Assistant Professor, research interest: climate change and food security, Email: shiyu174274@163.com; **Ning Jin**, PhD, Associate Professor, research interest: precision agriculture, Email: jinn.13b@igsnrr.ac.cn; **Bingyan Wu**, PhD candidate, research interest: climate change and crop water use efficiency, Email: yan18434764144@163.com; **Zekun Wang**, PhD, Assistant Professor, research interest: whisker growth in extreme environments, Email: zkwang@shu.edu.cn; **Shiwen Wang**, PhD, Professor, research interest: water and soil conservation for dryland agriculture, Email: shiwenwang@nwsuaf.edu.cn.
***Corresponding author: Ning Jin**, PhD, Associate Professor, research interest: precision agriculture. Department of Resources and Environment, Shanxi Institute of Energy, Jinzhong 030600, Shanxi, China. Email: jinn.13b@igsnrr.ac.cn.

mapping. With the rapid development of satellite sensors, many available satellite images with high resolution bring new opportunities for the fine-scale mapping of cropland. Compared with remote sensing datasets with lower spatial resolution, those with high spatial resolution can maximize the reduction in mixed pixels and provide clearer geomorphic features[10]. Additionally, high-resolution images can capture more information about human activities. At present, multiscale agricultural cover with high resolution (less than 30 m) has been mapped to meet production needs[11,12]. For example, do Nascimento Bendini et al.[13] classified cropland with 30 m resolution in Brazil by using multitemporal Landsat images. Preidl et al.[14] produced 20 m-resolution agricultural maps containing 19 crop types in Germany by using Sentinel-2A images; the authors achieved an 88% overall accuracy. However, the uncertainty of mapping with high-resolution satellite images has rarely been discussed. In particular, this uncertainty may be further amplified in ecological modeling[15,16]. Hence, a major effort is required to investigate the reason for this uncertainty and to reduce it.

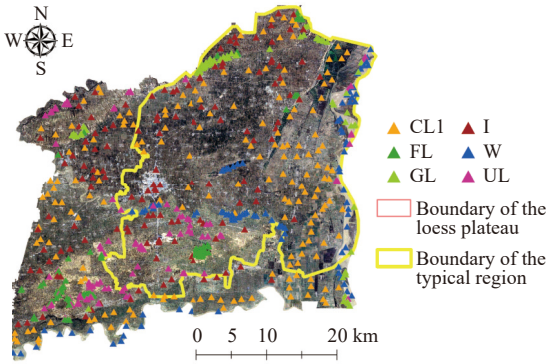
The feature combinations play various roles in classifying different satellite images[17,18]. Hu et al.[19] introduced a variety of vegetation indices to successfully classify Landsat 8 images and obtained an overall accuracy of 81.04% and more land-surface details. Sharma et al.[20] indicated that using texture features can effectively improve the classification accuracy for land cover mapping. However, due to the uneven spectral characteristics of land cover and a large amount of data, it is difficult to effectively and accurately classify cropland by using traditional classification methods[21]. Compared with traditional algorithms, machine learning (ML) techniques, such as artificial neural networks (ANNs)[22], support vector machines (SVMs)[23], and random forests (RFs)[24], provide a more accurate and efficient alternative for land cover classification based on big data and complex landscapes. Classification results calculated by ML models will be more accurate with sufficient training data[25,26]. Many scholars have classified research on cropland by using machine learning algorithms[25,26]. However, the performance of different algorithm-based machine-learning techniques varies depending on the cropland. For example, Hackman et al.[27] found that the maximum likelihood classifier outperformed the SVM and decision tree (DT) classifiers. Maxwell et al.[28] reviewed the application of machine learning classification in remote sensing mapping and indicated that the overall accuracies of support vector machines, decision trees, and random forests are higher than those of other machine learning classifiers. Consequently, it is necessary to obtain prior knowledge by comparing multiple models to explore the uncertainty of the classification.

A variety of global cropland maps have been produced, including the Global Map of Rainfed Cropland Areas[29], the Global Irrigated Area Map[30], and the Agricultural Map[31]. Low spatial resolution and limited accuracy impeded their business applications at the provincial and county levels[32]. The quality of these maps remains uncertain and questionable in the specific scenario[33]. Pérez-Hoyos et al.[34] compared nine global land cover datasets, and only 2.5% of the farmland was the same. Therefore, this study aimed to compare its results with the global cropland map to provide useful advice on the resolution of the map for small-scale (county-level) areas. China has implemented a large-scale "Gain-for-Green" (GFG) program since 1999. Human activities profoundly influenced LULC changes in the Loess Plateau. The ecologically fragile areas of the Loess Plateau are an important application target of the plan. Overall, Dali County on the Loess Plateau was selected as the study's case study, which comprises considerable cropland and is an important grain and cash crop-producing region. The objective of this study is to introduce multiple vegetation indices, texture features, and topographic factors based on prior knowledge combined with machine learning methods to classify cropland. The purpose of this study is to explore the effects of spatial resolution on cropland classification by using the best machine learning methods. This study addresses three key questions: 1) the effect of different classifiers on the accuracy of cropland classification; 2) the impact of remote sensing images with varying spatial resolutions on cropland classification accuracy; and 3) the effectiveness and reliability of the optimal classification method, evaluated through comparison with existing global land cover datasets.

2 Materials and methods

2.1 Study area

The Loess Plateau, located in the northern part of central China, experiences the most serious soil erosion in the world and has a fragile ecological environment[35]. The LULC of this region has changed greatly with rapid regional development and population growth. With a total area of 1776 km^2, Dali County in Shaanxi Province was the site of this study and is located in the south of the Loess Plateau (Figure 1). Regional features include the Yellow River in the east, the Wei River in the south, and mountainous hills in the northwest. The mean annual temperature (MAT) of Dali County is 13.6°C, and the mean annual precipitation (MAP) is 514 mm. The research area was divided into six categories: cultivated land (CL1), forestland (FL), grassland (GL), water (W), impervious (I), and unused land (UL) according to the characteristics of land use.



Note: The Google Earth image shows the entire study area of this study. The region with a yellow boundary, as the typical area, is used to further analyze the influence of the spatial resolution of images on land classification accuracy. CL1, FL, GL, W, I, and UL represent cultivated land, forestland, grassland, water, impervious, and unused land according to the characteristics of land use, respectively.

Figure 1 Location of the study area

2.2 Data source and image preprocessing

The purpose of this study was to explore the effects of different spatial resolution images on the classification results, especially in the case of high resolution. Five remote sensing images were selected from the following satellites (Table 1): Landsat 8 Operational Land Imager (OLI), GaoFen-1 (GF-1), and Sentinel-2. The range of the spatial resolution of these images is from 2 to 30 m. Although Google Earth provided more accurate images, this study did not consider using it as a part of the data classification study, because it is hard to compare Google Earth images, which

contain three bands (red, green, and blue bands), with other images. The remote sensing images were selected according to the following rules to display the six ground object types more clearly: 1) cloud cover of the remote sensing image is less than 5%; 2) remote sensing image features have obvious spectral and shape characteristics. Therefore, this study focused on the images from July to September 2015. The vegetation grew luxuriously during this period; consequently, farmland and forestland can be effectively distinguished from unused land and grassland, respectively, by texture and terrain. The difference was considered between the selected image year and month (changes in farmland vegetation planting), which will have different degrees of bias on the classification results, so this study paid special attention to this problem when selecting the validation data.

Table 1 Characteristics of satellite images used in this study

Satellite	Spatial-resolution/m	Image time	Cloud cover/%
GF-1-PMS	2	2015.06.04	0
GF-1-PMS	8	2015.06.04	0
Sentinel-2	10	2015.07-09	-
GF-1-WFV	16	2015.10.01	0
Landsat OLI	30	2015.07.26	3.17

Note: GF-1-PMS, GF-1-WFV, and Landsat OLI are the Gaofen-1 panchromatic/multispectral camera, the Gaofen-1 wide field-of-view (WFV) camera, and the Landsat Operational Land Imager camera, respectively.

Chen et al.[33] published the global 30 m land cover datasets in 2000 and 2010 by using pixel-based and object-based methods; these datasets described the global surface features with an overall accuracy of up to 80%. Furthermore, Gong et al.[36] published the global 10 m land cover datasets by using machine learning in the Google Earth Engine platform for the first time, thereby showing more spatial distribution details. However, the classification performance of these products is still unclear in small-scale (city-scale or county-scale) regions. The 30 m global land cover dataset was used for 2010 (LC30 m) and the 10 m global land cover dataset was used for 2015 (LC10 m) to analyze differences with the classification results of the small-scale region (i.e., Dali County).

2.2.1 Image preprocessing

All original images from different satellites were preprocessed for radiation calibration, atmospheric correction, mosaicking, and extracting the region of interest (ROI) to generate images of 2 m, 8 m, 10 m, 16 m, and 30 m spatial resolutions. More details of different satellite image preprocessing are provided below.

1) The Landsat OLI images (level-1) that were downloaded from an online website at https://earthexplorer.usgs.gov/ have a spatial resolution of 30 m[37]. The image contained 3.17% cloud cover and 126/36 path/row numbers on 26 July 2015. First, Landsat images were subjected to radiometric calibration and atmospheric correction by employing the Fast Line-of-Sight Atmospheric Analysis of Spectral Hypercubes (FLAASH) model by using ENVI 5.1 software. The radiometric calibration module was used to convert the DN value of the original image into radiance based on the radiation calibration parameters that came with the Landsat satellite. The formula is

$$L_\lambda = G \times \mathrm{DN} + O \quad (1)$$

where, L_λ represents radiance; G and O represent the gain and offset, respectively, from the image metafile; DN value is from the pixel value of the raw images. The FLAASH atmospheric correction module was used to convert radiance values into surface reflectance. More information on FLAASH atmospheric corrections is available in Reference [38].

2) Sentinel-2 A/B is based on a constellation of two identical satellites in the same orbit[39]. Sentinel-2 A/B has higher spatial resolution surface reflectance data (10 m) and a shorter revisit period (5 d) and can thus supplement missing data in the time since the satellites became fully operational in 2017. Sentinel-2 A/B contains three red-edged spectral bands (20 m) specifically for monitoring vegetation[40]. The Level-1C (L1C) top-of-atmosphere (TOA) reflectance of Sentinel-2 images was obtained from the European Space Agency (https://scihub.copernicus.eu/dhus/#/home/). Furthermore, the Sen2cor module was employed within the Sentinel-2 toolbox for the L1C TOA reflectance of Sentinel-2 images to conduct atmospheric correction; L1C TOA reflectance was converted into Level-2A top-of-canopy (TOC) reflectance. The four spectral bands, i.e., the red, green, blue, and NIR bands, with 10 m spatial resolution were used in this study. Detailed band configuration can be found on the European Space Agency's website at https://earth.esa.int/web/sentinel/user-guides/sentinel-2-msi/resolutions/spatial. Given the absence of remote sensing images in a single month over the study area, all images from July to September 2015 were selected, and each optimal band of all images in this period was restructured as a new image to achieve classification.

3) Launched at the Jiuquan Satellite Launch Centre of China in April 2013, the Gaofen-1 satellite carried two panchromatic/multispectral (PMS) and four wide field-of-view (WFV) cameras. GF-1-WFV has four spectral channels (blue, green, red, and near-infrared bands) and a spectral range from 450 to 892 nm[41]. In this study, the GF-1-WFV image from the China Center for Resources Satellite Data and Application (http://www.cresda.com/) has a spatial resolution of 16 m and a four-day revisit period. The image was selected with 0% cloud cover and 9/97 path/row numbers on 1 October 2015. Then, the GF-1-WFV image was subjected to radiometric calibration and atmospheric correction by employing the Fast Line-of-Sight Atmospheric Analysis of Spectral Hypercubes (FLAASH) model by using ENVI 5.1 software. The specific implementation steps are in Equation (1) and Reference [38].

4) The Gaofen-1 panchromatic/multispectral (GF-1-PMS) has a spatial resolution of 2/8 m and a 41 d revisit period (Table S1). Two images were used with 0% cloud cover and 9/97 and 9/98 path/row numbers on 4 June 2015 from the China Center for Resources Satellite Data and Application (http://www.cresda.com/). The GF-1-PMS was processed in the same way as GF-1-WFV. Since only one water band was obtained from the 2 m panchromatic images of GF-1-PMS, the vegetation index cannot be calculated directly. Therefore, the 8 m multispectral and 2 m panchromatic images were used to generate a 2 m multispectral image by using the Gram-Schmidt (GS) pansharpening method in ENVI software[42]. The algorithm was based on vector orthogonalization to increase the spatial resolution and provide better visualization of a multiband image using the higher-resolution panchromatic bands[42].

5) LC30 m was generated by an approach based on the integration of pixel- and object-based methods with knowledge (POK-based). The classification of 10 land cover types adopted a split-and-merge strategy. The authors developed a knowledge-based interactive verification process to improve the quality of classification results. The overall classification accuracy of LC30 m was above 80%[33]. In this study, the LC30 m in 2010 was used for further analysis.

6) LC10 m was generated using stable classification with

limited samples and a random forest algorithm. Tens of thousands of remote sensing images from Sentinel-2 data were classified on the Google Earth Engine. The overall accuracy of LC10 m was 72.76%[36].

7) The ASTER GDEM V2 global DEM dataset with 30 m resolution was used from the Earth Remote Sensing Data Analysis Center (http://www.ersdac.gdem.aster.or.jp/). The ASTER GDEM dataset was jointly released by METI (Japan) and NASA (the United States) and is available free to the public. The ASTER GDEM V2 dataset was developed from the V1 version of the GDEM image by using an advanced algorithm to improve the spatial resolution and elevation accuracy of the data.

2.2.2 Training features

The complexity of the actual ground feature spectrum affected the classification results. Previous research explored the selection of classification features. Therefore, the vegetation indices (NDVI, EVI, and NDWI) and texture features of all images were calculated that were stacked with the DEM and four preprocessed radiation bands (red, green, blue, and near-infrared bands) to form a new dataset for LULC classification (Table 2). All the feature variables in Table 2 are adopted in the classification of satellite images, i.e., Landsat OLI, Sentinel-2, and GF-1-WFV/PMS images. The calculation formulas of the vegetation indices are as follows:

$$EVI = 2.5 \times \frac{\rho_{NIR} - \rho_{Red}}{\rho_{NIR} + 6 \times \rho_{Red} - 7.5 \times \rho_{Blue} + 1} \quad (2)$$

$$NDVI = \frac{\rho_{NIR} - \rho_{Red}}{\rho_{NIR} + \rho_{Red}} \quad (3)$$

$$NDWI = \frac{\rho_{Green} - \rho_{NIR}}{\rho_{Green} + \rho_{NIR}} \quad (4)$$

where, ρ_{Red}, ρ_{NIR}, ρ_{SWIR}, ρ_{Green}, and ρ_{Blue} are the surface reflectance values of the red band, near-infrared band, shortwave infrared band, green band, and blue band, respectively.

Table 2 Features of all images with different spatial resolutions

Bands	Description	References
VIS	Red, Green, and Blue bands	[43]
NIR	Near-Infrared band	[44]
DEM	30 m ASTER GDEM V2	[45]
NDVI	Equation (3)	[46]
EVI	Equation (2)	[46]
NDWI	Equation (4)	[47]
texture	mean, variance, homogeneity, contrast, dissimilarity, entropy, second moment, correlation for each band of different satellite images	[20]

Note: VIS, NIR, DEM, NDVI, EVI, and NDWI are abbreviations for visible radiation, near-infrared, digital elevation model, normalized difference vegetation index, enhanced vegetation index, and normalized difference water index, respectively.

The texture features were calculated by using the following three steps. 1) First, all bands of remote sensing were subject to principal component changes by using ENVI software. The first principal component contained more than 85% of the information of remote sensing images. 2) Second, the texture characteristics of the first principal component extracted in ENVI software include the mean, variance, homogeneity, contrast, dissimilarity, entropy, second moment, and correlation (Table 2). 3) Finally, the visible light, near-infrared, terrain factors, vegetation indices, and texture features were stacked to form a dataset to be classified.

2.3 Methods of classifying images

The datasets generated in Section 2.2 were classified using random forest (RF), support vector machine (SVM), decision tree (DT), and artificial neural network (ANN) models. The remote sensing bands (Table 2) were used as the input variable and the land use classes were used as the target variable to build the model and obtain the prediction dataset. The optimal classifier was selected to identify remote sensing images with different spatial resolutions to explore the influence of image resolution on classification accuracy. Additionally, the classification results were compared with existing global land cover datasets to further evaluate the effectiveness of the classification methods and the reliability of the results.

1) RF is an integrated classifier based on decision trees[24,48], with a very fast generation rate and short training time for hundreds of decision trees, compared with a single classifier[49]. The RF algorithm requires only a few parameters but has the advantages of removing abnormalities and optimizing data[50]. RF improves generalization accuracy by increasing the complexity of the training data to avoid over-classification[24,51]. In this study, the random forest model was built by using the "randomForest" package in the R program (https://www.r-project.org/). Random forest models are usually influenced by two parameters, namely, "ntree" and "mtry", which represent the number of decision trees and features, respectively, selected at each node of the decision tree. Based on previous research, the value of mtry is set as the square root of the number of features for the RF model[52]. Hence, this study sets "ntree" and "mtry" as 1500 and 5, respectively, to obtain the best performance.

2) SVM: A support vector machine is a machine learning algorithm based on the principle of structural risk minimization by using statistical theory[23]. SVM converts a nonlinear problem into a linear one by using a kernel function, which can map the problem to a high-dimensional space. In this space, SVM finds the optimal segmentation plane that maximizes the dispersion between each sample. Hence, SVM can solve small samples and high-dimensional problems well and has a strong generalization ability[53,54]. In this study, the "e1071" package in the R program was used to establish an SVM model of the radial basis function for finding a nonlinear implicit relationship between the classes and features (Table 2). The grid-search method was also used to determine the best penalty coefficient (cost) and kernel parameter (gamma) for reducing model classification errors. Specifically, the cost and gamma were set as 10 and 0.1 for 30 m images, respectively, while the remaining parameters were set to their default values.

3) DT: Classification and regression trees (CARTs) are a binary tree classification method that uses fully spatial-auxiliary information and combines multiple feature variables to achieve the classification and prediction of images. CARTs determine the optimal segmentation point by comparing the Gini coefficient of attributes in the data preprocessing stage and finally establish a decision tree under different segmentation methods. CARTs prevent model overfitting problems by adopting the pruning approach[55]. The DT model is usually divided into two main steps: 1) a spanning tree; and 2) a pruning tree[56]. In most cases, pruning may be necessary to avoid inappropriate nodes in the tree. In this study, the DT model was built using the "rpart" package in the R program, and the DT model was adjusted to achieve the best performance through pruning. Specifically, for pruning, a 10-fold cross-validation was applied to filter the tree with the lowest prediction error. Moreover, the complexity parameter was set to 0.013 to penalize tree size based on the prune function in the R program. Finally, a pruned tree was used to predict land cover class.

4) ANN: An artificial neural network with strong learning and

collective computing ability can automatically realize the nonlinear mapping relationship between two groups of variables and requires few initial samples[57]. ANN does not need to display explicit relationships and samples to follow the independent or normal distribution[58]. Training a neural network plays a critical role in classification accuracy. Many algorithms, such as backpropagation (BP), have been applied to train neural networks to update weights[59]. BP neural networks, as the classical algorithms of deep learning, have been exploited in multiple agricultural activities, such as land use classification[59], crop growth monitoring[60], predicting the water consumption of vegetation[61], and yield assessment[62]. The remote sensing data were standardized before classification. Hence, a backpropagation neural network model was established based on the "nnet" package in the R program. A three-layer network was conducted with an input, hidden, and output layer. It was determined that the number of nodes (neurons) for the input, hidden, and output layers would be 16 (response variables), 10, and 6 (predicted class), respectively. The decay and max values in the ANN model were set to 0.01 and 1000, respectively, while the remaining parameters were set to their default values.

2.4 Verifying the accuracy of classification results

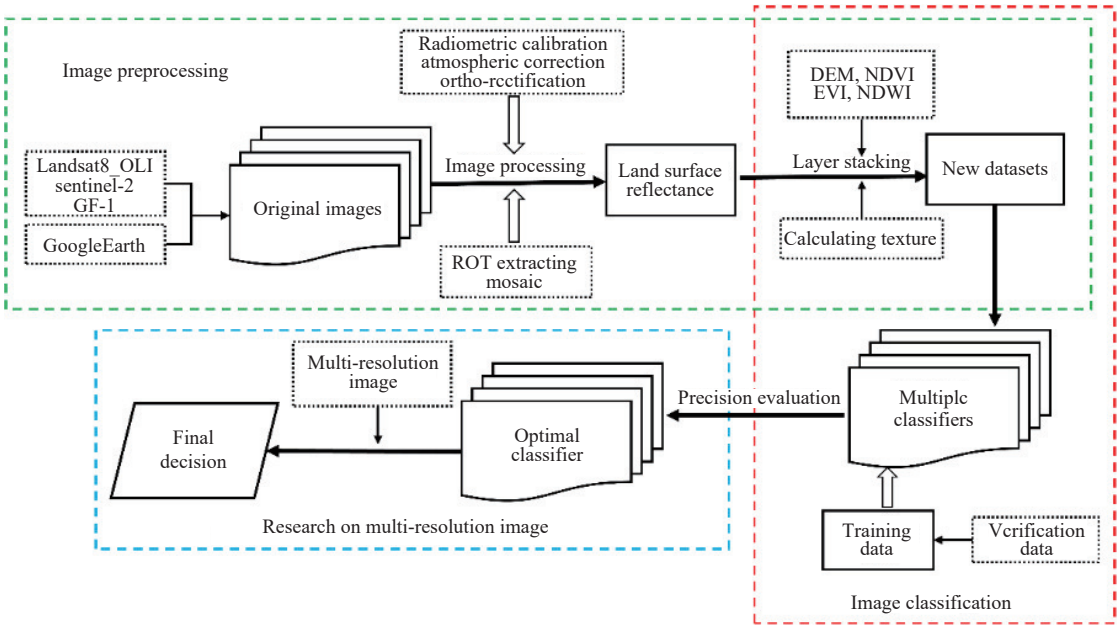
The verification data were obtained from very-high-resolution (VHR) Google Earth images to construct a classification model and to test its accuracy in this study. The 702 verification data points contained different types of features (Figure 1 and Table 3), of which 70% were used to build the classification model and 30% were used to verify the accuracy of the classification results (Figure 2). The verification points based on stratified random sampling were used to reasonably cover the study area. First, the study area was divided into six layers (cultivated land, forestland, grassland, water, impervious, and unused land) according to the 2015 China Land Use Dataset and MCD12Q1 land use data. Second, random points were generated in each layer. Third, historical VHR images were used to check each random point and selected and marked points covered by pure soil according to the VHR image; points without clear land cover information were excluded[63]. If the 2015 VHR image is missing in this area, images from adjacent years were used to filter points. Specifically, two adjacent years in the target region were selected for 2015. If the land cover did not change in these two years, we assumed that the land cover in 2015 would be consistent with that of the adjacent years. The validation points were reserved; otherwise, they would be removed. This methodology maximally reduces the sampling errors because of the repeated change in land cover during the adjacent years with a small probability. Therefore, the samples of different land cover types at the same time were obtained as the satellite images (i.e., 2015). The few samples from adjacent years do not interfere with outcomes due to the unremarkable land-use change in a short time.

The verification of the results was based on the confusion matrix generated by the verification data. Specifically, the overall accuracy (OA, percentage of correctly classified sites) and the Kappa coefficient (OA normalized by the baseline random chance for the dataset) were used to represent the difference between predicted and observed values[64]. Producer accuracy (PA) and user accuracy (UA) were used to assess the classification accuracy of a single feature class[65]. The formulas for OA, the Kappa coefficient, PA, and UA are as follows:

Table 3 Number of verification points for six LULC classes across the study area

Abbreviation	Land use and Land cover classes	Number of samples		
		Total	Training	Validation
CL1	Cultivated land	173	119	54
FL	Forestland	102	75	27
GL	Grassland	101	68	33
W	Water	94	65	29
I	Impervious	132	94	38
UL	Unused land	100	70	30

Note: CL1, FL, GL, W, I, and UL represent cultivated land, forestland, grassland, water, impervious, and unused land according to the characteristics of land use, respectively. Same below.



Note: DEM, NDVI, EVI, and NDWI represent the digital elevation model, normalized difference vegetation index, enhanced vegetation index, and normalized difference water index, respectively. The ROI represents the region of interest.

Figure 2 Flow chart presenting the overall classification procedure

$$\text{OA} = \frac{\sum_{k=1}^{q} n_{kk}}{n} \quad (5)$$

$$\text{Kappa coefficient} = \frac{n\sum_{k=1}^{q} n_{kk} - \sum_{k=1}^{q} n_{k+}n_{+k}}{n^2 - \sum_{k=1}^{q} n_{k+}n_{+k}} \quad (6)$$

$$\text{UA} = \frac{n_{kk}}{n_{k+}} \quad (7)$$

$$\text{PA} = \frac{n_{kk}}{n_{+k}} \quad (8)$$

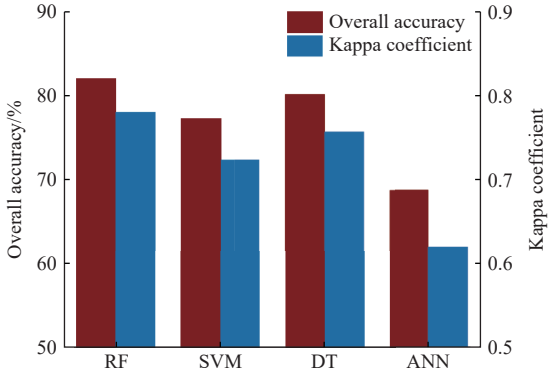
where, n_{k+}, n_{+k}, and n_{kk} are row sum value, column sum value, and main diagonal value of the matrix. n and q represent the number of validation samples and the number of classes, respectively.

3 Results

3.1 Classification results of different models

The four machine learning models described above were used to classify the processed Landsat 8 OLI (30 m) images. The overall accuracies of the models were as follows (in the order of highest to lowest): 81.99% (RF), 80.09% (DT), 77.25% (SVM), and 68.72% (ANN). Likewise, the Kappa coefficients of the four models were as follows (in the order of highest to lowest): 0.78 (RF), 0.76 (DT), 0.72 (SVM), and 0.62 (ANN) (Figure 3).

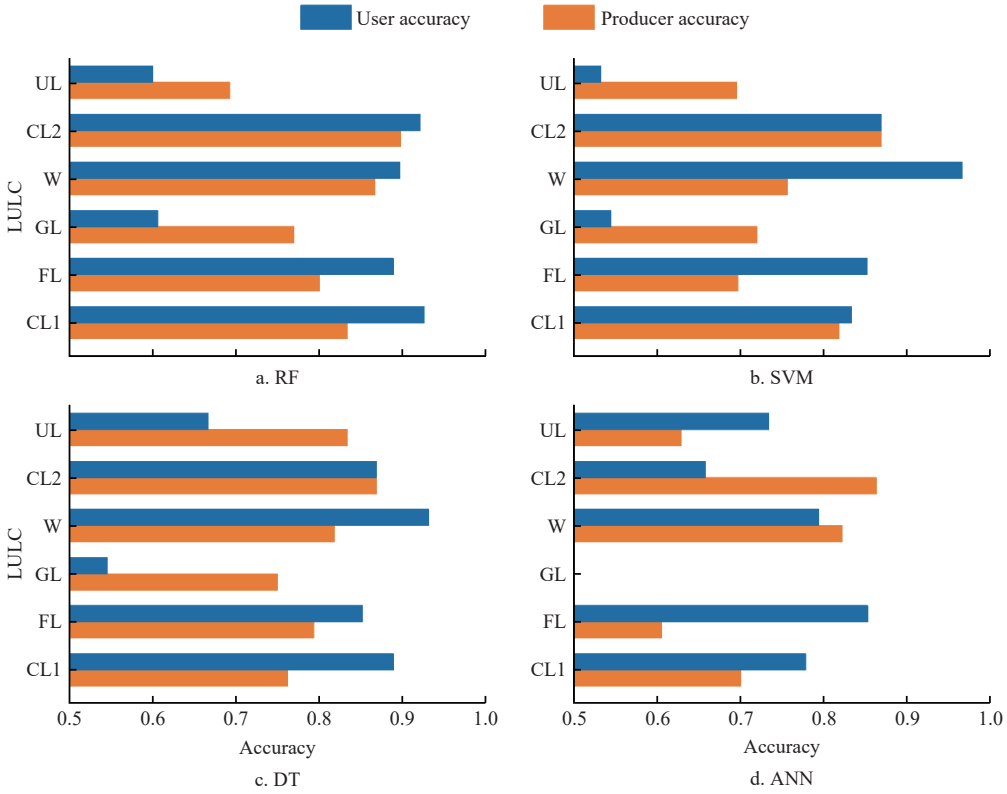
The RF, SVM, and DT models had higher user and producer accuracies for LULC classifications CL1, FL, W, and CL2 and performed better in distinguishing the classes of land features (Figure 4). None of the four classification models performed well in distinguishing GL (both the user and producer accuracies were relatively low). Based on confusion matrix, the study found that the four algorithms all confused the grassland as cultivated land and forest (Table S3). According to RF, SVM, DT, and ANN, cultivated land and forest accounted for 77%, 67%, 80%, and 70%, respectively, of the total misclassified (Table S3). RF and SVM had lower user and producer accuracies in distinguishing UL and thus did not achieve the best classification. ANN outperformed the other models in terms of user accuracy in distinguishing UL. However, the producer accuracy values of ANN for FL, GL, and UL were low, indicating that three classes were misclassified (Figure 4).



Note: RF, random forest; SVM, support vector machine; DT, decision tree; ANN, artificial neural network. Same below.

Figure 3 Overall accuracies and Kappa coefficients of different classification models

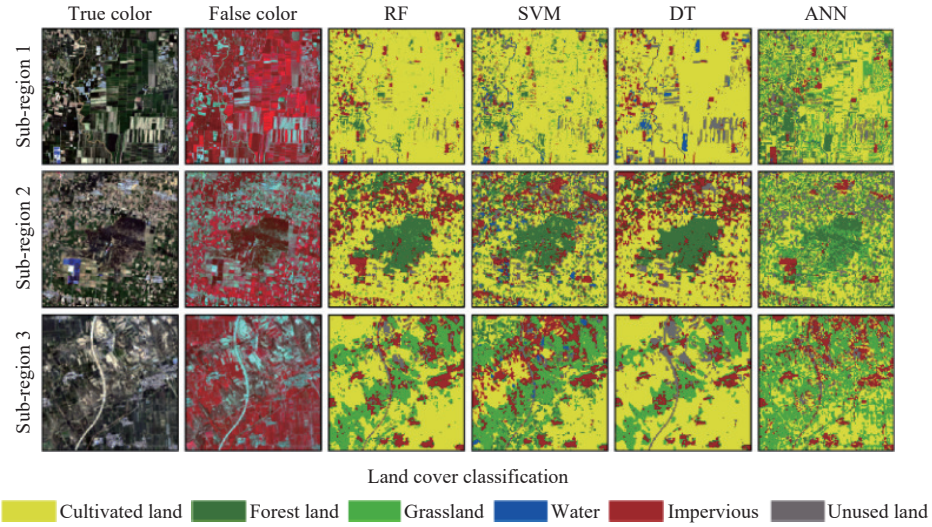
The performances of the four machine learning models on land classification (especially cropland, forest, and grassland) are shown in Figure 5. Specifically, RF, SVM, and DT can accurately reflect the spatial characteristics of the main ground classes in different subregions, while ANN has serious confusion and cannot distinguish CL1, GL, and UL. This research result is consistent with



Note: CL1, cultivated land; FL, forestland; GL, grassland; W, water; I, impervious; UL, unused land. Same below.

Figure 4 User and producer accuracies of different machine learning models in distinguishing six LULC classes in Dali County, Shaanxi Province, China

Figure S1, which shows the spatial distribution of classification results for the four machine learning models and indicates that RF, SVM, and DT can accurately capture the spatial traits of the main ground classes. There was also some confusion with the SVM model, which misclassified the middle and south subregions of the study area as UL and the northwest subregions as FL and GL (Figure S1). These results were likely caused by model overfitting. DT misclassified much of the CL2 near the river. RF performed well in classification and prediction, and the accuracy of classification was high for multiple ground classes (Figure S1).



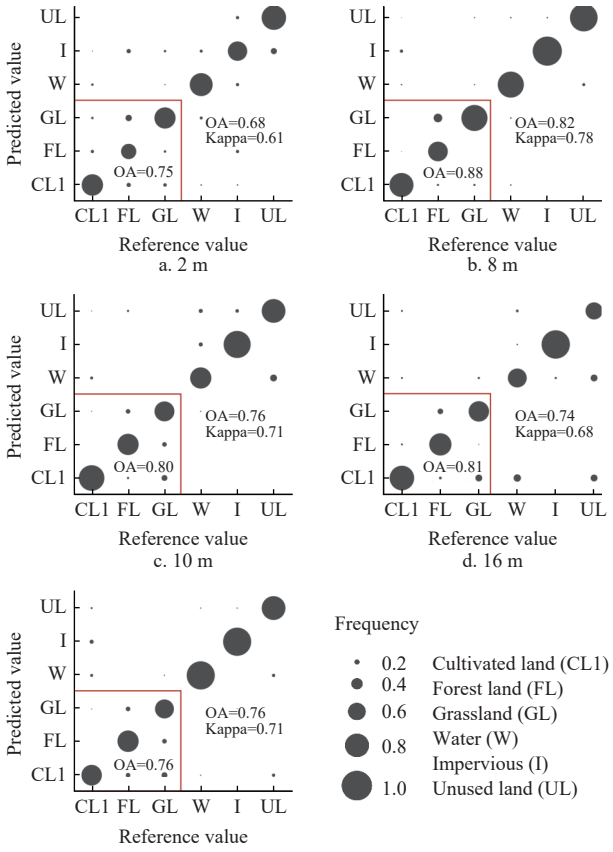
Note: True-color (Bands 3, 2, and 1) and false-color (Bands 4, 3, and 2) images are from Landsat.

Figure 5 Comparison between the land cover maps by using different machine learning algorithms

3.2 Classification of different resolution images

Although the RF classifier showed lower UAs and PAs, this study focused more on the performance of cropland. Comparisons of the results from the four machine learning classification models indicated that RF performed the best by effectively distinguishing cropland and avoiding overfitting problems. Therefore, RF was selected to classify remote sensing data with different spatial resolutions to determine the impact of resolution on the classification results. To avoid the influence of other factors, the same features (Table 1) and model parameters were used for remote sensing data with different spatial resolutions. The typical area, containing all attribute features, was further selected from the study area (Figure 1) to improve the calculation speed and efficiency.

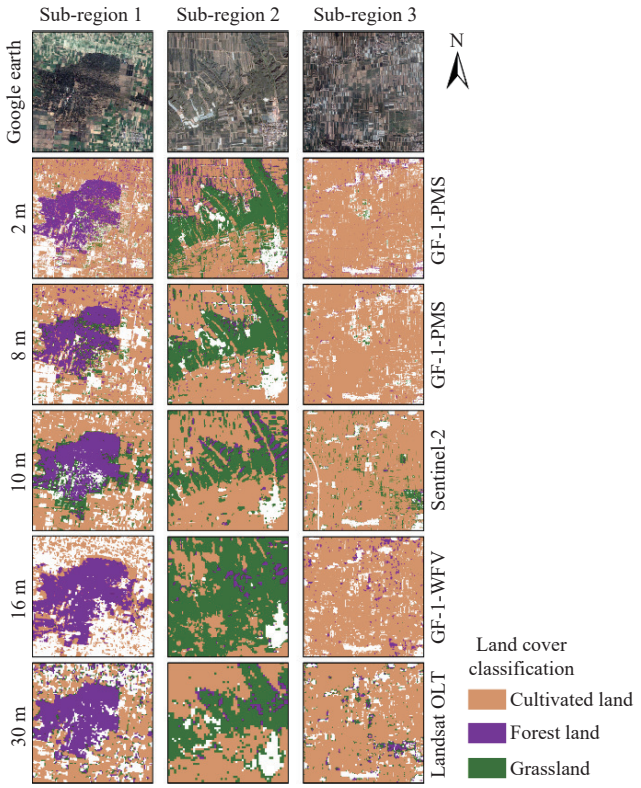
The overall accuracies of different-resolution images were as follows (in the order of highest to lowest): 82% (8 m), 76% (10 m), 76% (30 m), 77% (16 m), and 68% (2 m). The overall accuracies of cropland displayed a similar trend of variation; i.e., compared with the performances with other images, the performance with the 8 m resolution images (88%) reached the peak. Additionally, the Kappa coefficients of the five images were as follows (in the order of highest to lowest): 0.78 (8 m), 0.71 (10 m), 0.71 (30 m), 0.68 (16 m), and 0.61 (2 m) (Figure 6). The classification results for 8 m resolution images were the best, while the classification results for 2 m resolution images were the worst. The classification results for 10 m, 16 m, and 30 m resolutions were relatively close (Figure 6). The overall accuracies and Kappa coefficients of the 10 m and 30 m images were the same. The classification of each class in the 2 m image was unsatisfactory, with an accuracy of less than 70% for FL, GL, and CL2 and a classification accuracy of less than 80% for W and CL1 (Figure 6). Classification of the 8 m image for FL was poor, but other feature classes had relatively higher classification accuracies (the accuracies of both CL2 and UL were more than 90%). The 10 m image had the best classification accuracy for CL1, but the classification accuracy values for FL, GL, and W were relatively low. The classification accuracy of the 16 m image was better for CL2 (for which the accuracy was greater than 90%) than for GL, W, and UL (for which the accuracy was less than 70%). For the 30 m image, classification accuracies were better for W and CL2 than for the other four feature classes (Figure 6).



Note: Red boxes indicate the classification accuracy of cultivated land, forest land, and grassland. OA represents the overall accuracy. The value of each type was normalized based on the reference value to meet intercomparisons.

Figure 6 Scatterplots of classification accuracies for different-resolution images

The classification results of the five spatial-resolution images exhibited different spatial distributions (Figure S2). The distribution of the main classes could be displayed accurately, but there were significant differences in a single feature class due to the different spatial resolutions of the images (Figure S2). The CL1 areas showed similar spatial patterns for the different spatial resolutions. The 2 m and 16 m images predicted more FL and GL areas, which were distributed mainly in the northern region of the study area (Figure S2). There were many prediction errors with the 16 m image, especially for W (distributed mainly in the south of the study area), as the predicted classification was inconsistent with the actual ground observations. Compared with the other images, the 8 m image classified more CL2 areas, and these areas were distributed mainly in the middle of the study area. Compared with the 30 m image, the 2 m, 8 m, 10 m, and 16 m images predicted more UL areas, which were distributed mainly in the south of the study area. For cropland, the details of the spatial distribution were further compared based on the images with different spatial resolutions (Figure 7). The increased mixed pixels were observed with an improvement in resolution. More high-resolution images missed information on cropland; e.g., 16 m and 30 m images missed such information in subregion 1 (Figure 7). Additionally, high-resolution images lose more details of ground information, especially for cropland, and fine land ubiquitously exists in practical production. Although a 2 m resolution image could capture more ground details, these details misled classifiers and generated additional confusion (e.g., mulch on cropland was misclassified as impervious) (as shown in Figure 7).



Note: Non-cropland was masked in each image. GF-1-PMS, GF-1-WFV, and Landsat OLI represent the Gaofen-1 panchromatic/multispectral camera, Gaofen-1 wide field-of-view camera, and Landsat Operational Land Imager camera, respectively.

Figure 7 Land cover maps using the optimal model for the different-resolution images

Figure 8 shows the importance of classification features for different-resolution images using the RF model. Overall, the importance of classification features in different-resolution images has a similar variation pattern, which is in descending order of vegetation index, spectral reflectance, and texture. Among these indices, the DEM, NIR, and Texture mean indices are highly important. The importance of the DEM index under the 8 m image is higher than in other resolutions, the NIR index under 10 m images is more important than in other resolutions, and the importance of Texture in 30 m images is higher than in other resolutions.

3.3 Comparisons with existing LULC datasets

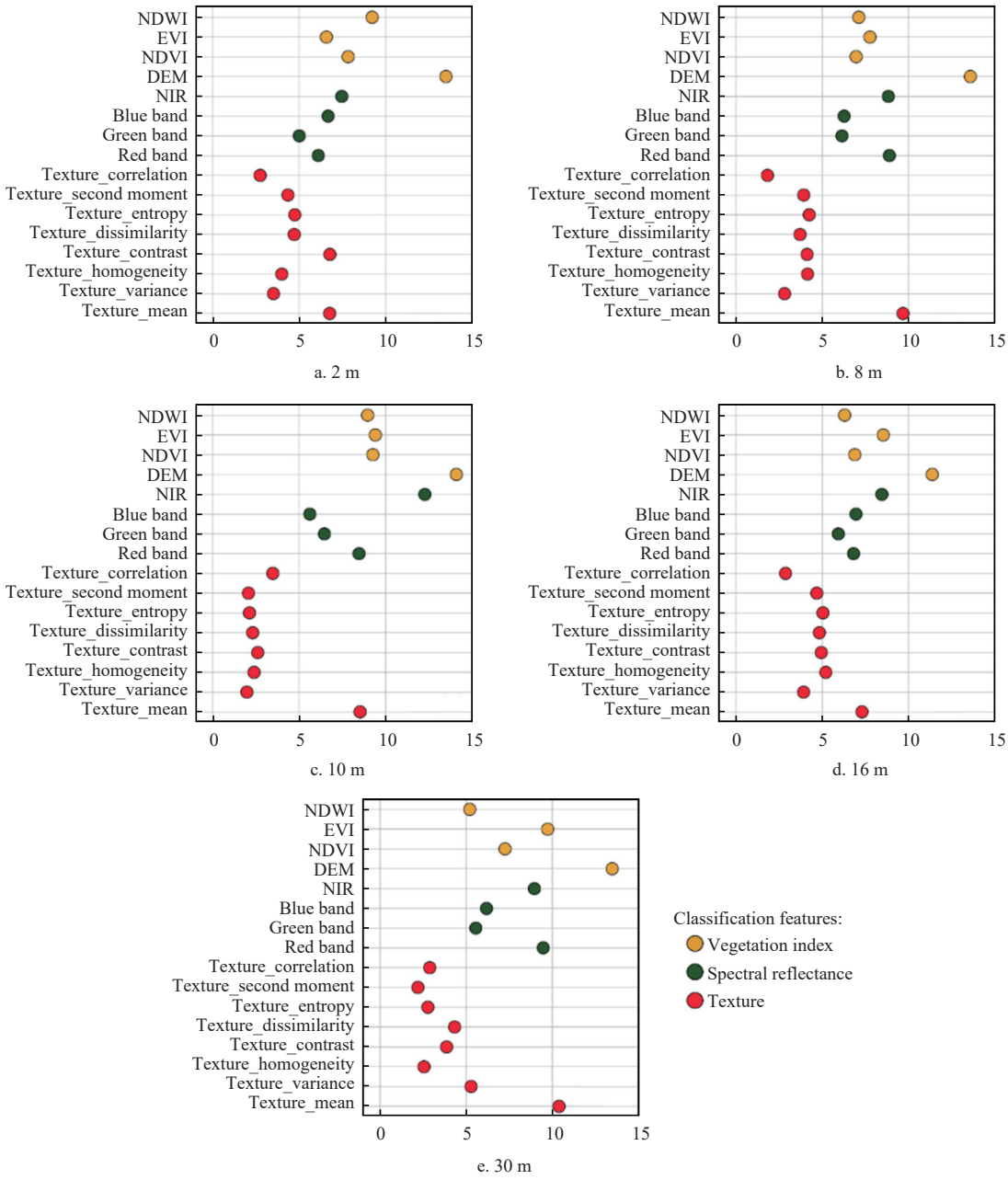
Although 8 m resolution images proved to be the best for land use classification, there are no land cover products with the same resolution comparable to the results of this study. Given that the different-resolution images may introduce additional uncertainty, in this study, popular global datasets with the same spatial resolution (LC10 m and LC30 m) were used to further verify the effectiveness of the classification method and the reliability of the results. For ease of comparison, the classes of the land cover dataset in the study area were reclassified to correspond with our six classification categories (Table S2). Then, the longitude and latitude information for the 702 measured validation points were used to extract classes of land cover and generate confusion matrices for LC10 m, LC30 m, RF10 m, and RF30 m (Figure 9). The results showed that the overall accuracies of the RF10 m, LC10 m, RF30 m, and LC30 m images were 93.59%, 49.43%, 94.59%, and 55.98%, respectively; and the Kappa coefficients were 0.92, 0.36, 0.93, and 0.45, respectively.

The confusion matrix provided more detailed classification results (Figure 9). Both RF10 m and RF30 m had high classification accuracy, with the accuracy of most feature classes being greater than 90% and few misclassifications. However, the classification accuracies of both LC10 m and LC30 m were poor, with many of the classes misclassified (Figure 9). Figure 10 shows the conspicuous differences in spatial distributions. Specifically, the areas and distributions of CL2 and UL were different between the RF results and the LC results.

4 Discussion

4.1 Effect of machine learning models on classification results

In this study, four popular ML models were employed in classifying high-resolution images. This study's results demonstrated that the overall accuracy of the RF classifier was higher than those of the other three classifiers. However, the results should be treated with some caution due to the substantial differences in the accuracies that were calculated for different classes by the four ML models. Specifically, the RF classifier outperformed the other three algorithms on cropland, and the user and producer accuracies on cultivated land and forestland were over 80%. This result was further validated in a past study[66]. However, for the water class, the SVM and DT outperformed RF and ANN in terms of user accuracy. ANN was found to be the best classifier on unused land but significantly confounded cultivated land and forestland. The difference between different algorithms leads to the difference in the accuracy of specific land cover classification. The previous studies of model comparison reached a similar conclusion. For example, Ngo et al.[67] reported that with an OA of 94.81%, the SVM outperformed the RF classifier for cropland and forest in the Mekong Delta. Additionally, Prasad et al.[68] found that compared with RF and ANN, the logit boost model had higher robustness. It



Note: The details of each feature are listed in Table 2. The scores of the importance for each feature were calculated by the increase in the mean square error (MSE) percentage (% IncMSE).

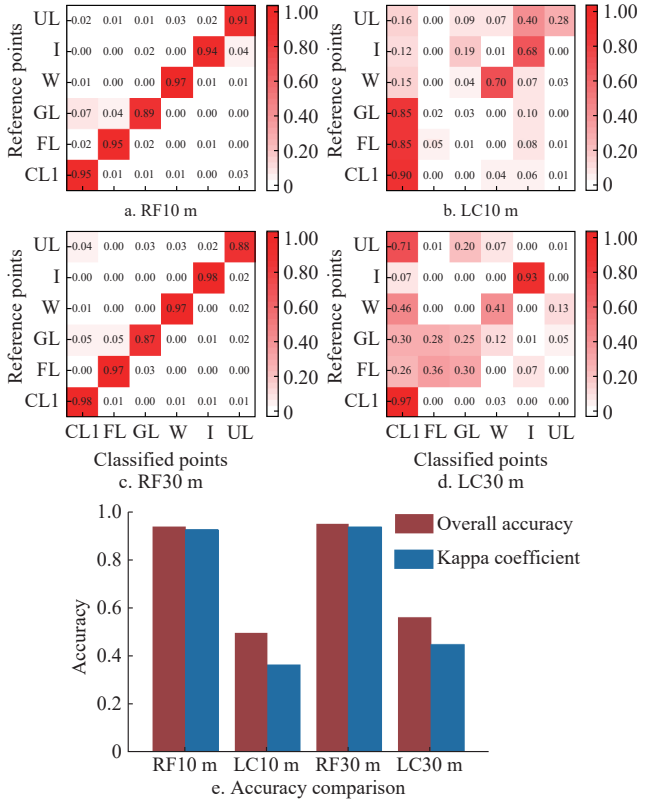
Figure 8 Importance of random forest-based classification features for different-resolution images

should be pointed out, however, that the four algorithms all show bad performance in grassland (Figure 4). According to a previous study, the grasslands, cultivated lands, and forests had a similar spectral response, which is likely the reason for the confusion[69]. Further, the classification performance will be dominated by different spectra with the same features and different features with the same spectra[70]. It is difficult to distinguish a similar land cover solely based on the vegetation spectrum. A classification system that incorporates phenological information would increase the accuracy of classifying vegetation.

Uncertainties among algorithms are often large and easy to overlook because the more popular classifier based on single evaluations is usually the preferred classifier for land mapping. Therefore, we highlight the importance of an appropriate classifier and workflow for developing land planning and spatial mapping. This is the first step to reducing uncertainty.

4.2 Impact of spatial resolution on classification results

This study used a random forest classification method to determine the effect of spatial resolution on classification accuracy. The use of the DEM in land cover classification can improve classification accuracy, especially when distinguishing grassland from forestland (Figure 8)[71]. Remote sensing images with different spatial resolutions will affect classification accuracy[72,73]. It found that increasing the spatial resolution of remote sensing images generally improved classification accuracy, with the best accuracy seen with 8 m resolution images. Higher-resolution images reduced classification accuracy. This result was consistent with the results of Roth et al.[74] More detailed features in high-resolution satellite images may affect the judgment of the classifier during classification, thereby resulting in confusion of results at different levels. Due to high landscape heterogeneity and spectral confusion among different land classes (especially in cropland, grassland, and



Note: RF10 m and RF30 m, random forest models with 10 m and 30 m resolutions, respectively; LC10 m and LC30 m, land classification datasets with 10 m and 30 m resolutions, respectively.

Figure 9 Land use confusion matrix for different datasets (upper four panels), and overall accuracies and Kappa coefficients of different classification models (lower panel) in Dali County, Shaanxi Province, China

forestland), it is difficult to map high-precision cropland[75]. An increase in the resolution of data sources does not seem to solve this problem. In the future, the heterogeneity of land cover will be combined with an appropriate pixel size to obtain better cropland mapping.

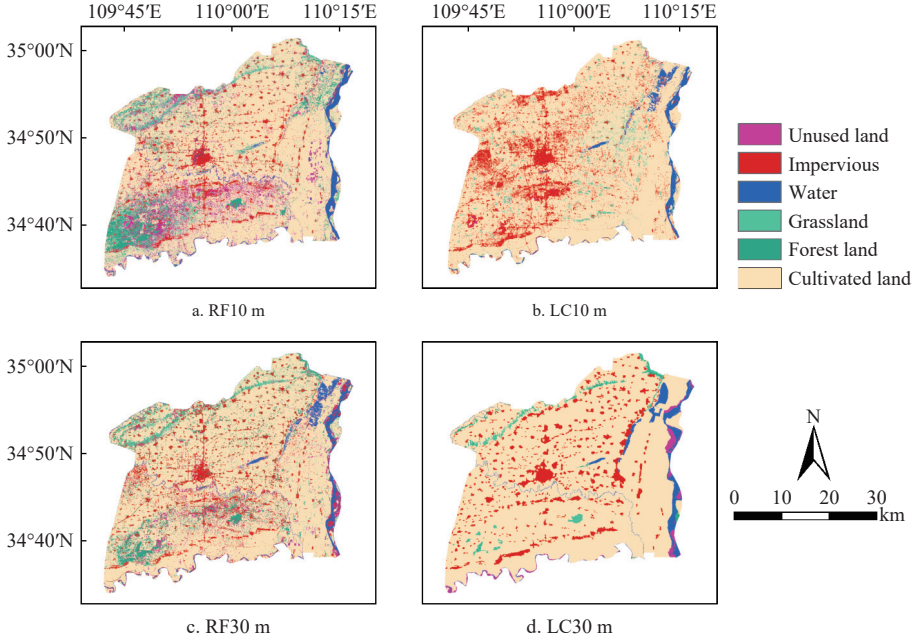
Another factor affecting classification results comes from the differences in satellite sensors. Specifically, different satellites have different sensors for receiving and transmitting information, and the calculation methods of these satellites differ. This study controlled the number and kinds of features, but it is impossible to quantitatively estimate the error from satellite differences.

4.3 Comparison with existing LULC datasets

This paper used global land cover data (LC30 m and LC10 m) as comparative data to test the RF classification model. The classification accuracy in this study was much greater than that of the global land cover dataset at the same spatial resolution. The difference may be related to the land cover class, the object orientation, and the models. Firstly, global land cover types are more complex than local ones, resulting in a greater variety of training data globally. With the collection of global data, there will be more heterogeneity within a category; for example, different spectral responses for different types of croplands will be observed. In addition, it should be noted that global data contain fewer local representative samples than local studies, thus increasing their uncertainty. It is therefore preferable for local studies to use local training and validation data.

Moreover, the results indicate that the local-based model contains higher accuracy in a local study than the global model. Different training tasks are required for global and local models, depending on the target object. And global and local models can all achieve high accuracy in their target domains. Due to differences in outliers or noise generated by different objects, it is imperative for users when working with specific local tasks to choose appropriate-scale land cover data. More noise, outliers, and insufficient training samples will interfere with test results.

However, readers should note that LC30 m is a dataset based on 2010 remote sensing images, and the random forest model was not used as the classification method. In contrast, LC10 m used random forest classification[34], but samples were from all over the



Note: RF10 m and RF30 m, random forest models with 10 m and 30 m resolutions, respectively; LC10 m and LC30 m, land classification datasets with 10 m and 30 m resolutions, respectively.

Figure 10 Spatial distribution of land use and land cover classes obtained from different datasets in Dali County, Shaanxi Province, China

world. Hence, the classification results of a local area cannot provide more accurate and clearer information.

4.4 Limitations and perspectives

This study used validation data to build classification models and to verify the accuracy of the results. The validation data were from Google Earth high-resolution images. The ground-truth data used in this study were available mainly for farmland. However, ground-truth data for other features were less available, and we thus had to rely on Google Earth images. This study was relatively weak regarding ground-truth points. In future studies, more ground-truth data will be added to improve the accuracy of verification.

The selection of features in this study was based on previous studies. The number of attribute features has a complex impact on classification accuracy, including reducing the correlation between any two classification trees in a random forest and increasing the error rate of the random forest[76]. In future research, the band combination method will be used to optimize the attribute features to reduce the errors from the original features.

It was noted that salt-and-pepper noise appeared in the different results for each image. However, this is an inevitable problem when using grid-based LULC classification. To reduce this noise in the future, the object-oriented methods will be used to segment the image into objects and classify the features of the objects to gradually reduce such problems.

With the future increase in remote sensing mapping, additional attention should be given to the problem of matching data with algorithms and data with classifiers. Human activities and meteorological conditions may affect the correlation between terrain and land cover category in large-scale land cover mapping[77]. The introduction of crop phenology can improve the precision of classification[78] when attempting to more accurately classify a single category, such as cultivated land. LULC classification of multitemporal images of appropriate resolution combined with machine and deep learning algorithms can provide a reference and basis for local decision-making regarding food security policies.

5 Conclusions

This study used multiple machine learning models to classify remote-sensing images with different spatial resolutions. The impact of classification algorithms and spatial resolutions on land use classification results was tested to explore the potential uncertainty of land mapping. The results show that the RF model outperformed other machine learning classifiers in agriculture land mapping and could better distinguish cropland, forest, and grassland. SVM and ANN showed a higher accuracy on the water and unused land classes for the local region, respectively. Additionally, the accuracy and reliability of land use classification results were affected by spatial resolution. A potential trend was observed between the classification accuracy and the spatial resolution, which initially increased together, followed by a decrease in classification accuracy as the spatial resolution continued to increase. Images with a suitable resolution level provided the optimal classification results. Finally, classifications from the random forest model were more reliable than the classifications from the existing large-scale land use datasets for the local regions. The results of this study highlight that the classifier is one of the sources of uncertainty for land mapping. Classification errors caused by the insufficient or excessive spatial resolution of remote sensing images cannot be ignored in land use monitoring. Selecting images by using the appropriate classification model and spatial resolution is very essential for obtaining reliable land use classification results.

Acknowledgements

This study was supported by the Natural Science Research General Program of Shanxi Province Basic Research Project (Grant No. 202 203 021 221 231).

[References]

[1] Sun D, Zhang W X, Lin Y B, Liu Z F, Shen W J, Zhou L X, et al. Soil erosion and water retention varies with plantation type and age. Forest Ecology and Management, 2018; 422: 1–10.

[2] Lin D L, Xia J Y, Wan S Q. Climate warming and biomass accumulation of terrestrial plants: A meta-analysis. New Phytologist, 2010; 188(1): 187–198.

[3] Liang J Y, Xia J Y, Liu L L, Wang S Q. Global patterns of the responses of leaf-level photosynthesis and respiration in terrestrial plants to experimental warming. Journal of Plant Ecology, 2013; 6(6): 437–447.

[4] Geist H, Lambin E F (Eds.). Land-use and land-cover change: Local processes to global impacts. Springer, 2006; 205p.

[5] Bussi G, Whitehead P G, Gutierrz-Canovas C, Ledesma J L J, Ormerod S J, Counture R M. Modelling the effects of climate and land-use change on the hydrochemistry and ecology of the River Wye (Wales). Science of The Total Environment, 2018; 627: 733–743.

[6] Luo Z L, Shao Q X, Zuo Q T, Cui Y K. Impact of land use and urbanization on river water quality and ecology in a dam dominated basin. Journal of Hydrology, 2020; 584: 124655.

[7] Sun Z H, Di L P, Fang H, Burgess A. Deep learning classification for crop types in North Dakota. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2020; 13: 2200–2213.

[8] Zhang C, Di L P, Lin L, Guo L Y. Machine-learned prediction of annual crop planting in the U.S. Corn Belt based on historical crop planting maps. Computers and Electronics in Agriculture, 2019; 166: 104989.

[9] Fritz S, See L, Rembold F. Comparison of global and regional land cover maps with statistical information for the agricultural domain in Africa. International Journal of Remote Sensing, 2010; 31(9): 2237–2256.

[10] Pu R L, Bell S. Mapping seagrass coverage and spatial patterns with high spatial resolution IKONOS imagery. International Journal of Applied Earth Observation and Geoinformation, 2017; 54: 145–158.

[11] Hansen M C, Potapov P V, Moore R, Hancher M, Turubanova S A, Tyukavina A, et al. High-resolution global maps of 21st-century forest cover change. Science, 2013; 342(6160): 850–853.

[12] Wang C, Tandeo P, Mouche A, Stopa J E, Gressani V, Longepe N, et al. Classification of the global Sentinel-1 SAR vignettes for ocean surface process studies. Remote Sensing of Environment, 2019; 234: 111457.

[13] do Nascimento Bendini H, Fonseca L M G, Schwieder M, Korting T S, Rufin P, Del Arco Sanches I, et al. Detailed agricultural land classification in the Brazilian cerrado based on phenological information from dense satellite image time series. International Journal of Applied Earth Observation And Geoinformation, 2019; 82: 101872.

[14] Preidl S, Lange M, Doktor D. Introducing APiC for regionalised land cover mapping on the national scale using Sentinel-2A imagery. Remote Sensing of Environment, 2020; 240: 111673.

[15] Lu D, Weng Q. A survey of image classification methods and techniques for improving classification performance. International Journal of Remote Sensing, 2007; 28(5): 823–870.

[16] Lyu Z Y, He H Q, Benediktsson J A, Huang H. A generalized image scene decomposition-based system for supervised classification of very high resolution remote sensing imagery. Remote Sensing, 2016; 8(10): 814. doi: A generalized image scene decomposition-based system for supervised classification of very high resolution remote sensing imagery.

[17] Wang J, Xiao X M, Liu L, Wu X C, Qin Y W, Steiner J L, et al. Mapping sugarcane plantation dynamics in Guangxi, China, by time series Sentinel-1, Sentinel-2 and Landsat images. Remote Sensing of Environment, 2020; 247: 111951.

[18] Zhong Y F, Su Y, Wu S Q, Zheng Z D, Zhao J, Ma A L, et al. Open-source data-driven urban land-use mapping integrating point-line-polygon semantic objects: A case study of Chinese cities. Remote Sensing of Environment, 2020; 247: 111838.

[19] Hu T Y, Yang J, Li X C, Gong P. Mapping urban land use by using Landsat images and open social data. Remote Sensing, 2016; 8(2): 151.
[20] Sharma R C, Tateishi R, Hara K, Iizuka K. Production of the Japan 30-m land cover map of 2013-2015 using a random forests-based feature optimization approach. Remote Sensing, 2016; 8(5): 429.
[21] Gómez C, White J C, Wulder M A. Optical remotely sensed time series data for land cover classification: A review. ISPRS Journal of Photogrammetry and Remote Sensing, 2016; 116: 55–72.
[22] Shi Y, Li L C, Wu B Y, Zhang Y J, Wang B, Niu W H, et al. Predicting rice productivity for ground data-sparse regions: A transferable framework and its application to North Korea. Science of The Total Environment, 2024; 946: 174227.
[23] Cortes C, Vapnik V. Support-vector networks. Machine Learning, 1995; 20(3): 273–297.
[24] Shi Y, Jin N, Ma X L, Wu B Y, He Q S, Yue C, et al. Attribution of climate and human activities to vegetation change in China using machine learning techniques. Agricultural and Forest Meteorology, 2020; 294: 108146.
[25] Pal M, Mather P M. Support vector machines for classification in remote sensing. International Journal of Remote Sensing, 2006; 26(5): 1007–1011.
[26] Schneider A. Monitoring land cover change in urban and peri-urban areas using dense time stacks of Landsat satellite data and a data mining approach. Remote Sensing of Environment, 2012; 124: 689–704.
[27] Hackman K O, Gong P, Wang, J. New land-cover maps of Ghana for 2015 using Landsat 8 and three popular classifiers for biodiversity assessment. International Journal of Remote Sensing, 2017; 38(14): 4008–4021.
[28] Maxwell A E, Warner T A, Fang F. Implementation of machine-learning classification in remote sensing: an applied review. International Journal of Remote Sensing, 2018; 39(9): 2784–2817.
[29] Biradar C M, Thenkabail P S, Noojipady P, Li Y J, Dheeravath V, Turral H, et al. A global map of rainfed cropland areas (GMRCA) at the end of last millenium using remote sensing. International Journal of Applied Earth Observation and Geoinformation, 2009; 11(2): 114–129.
[30] Thenkabail P S, Birandar C M, Noojipady P. Global irrigated area map (GIAM), derived from remote sensing, for the end of the last millennium. International Journal of Remote Sensing, 2009; 30(14): 3679–3733.
[31] Ramankutty N, Evan A T, Monfreda C, Foley J A. Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000. Global Biogeochemical Cycles, 2008; 22(1): GB1003.
[32] Vancutsem C, Marinho E, Kayitakire F, See L, Fritz S. Harmonizing and combining existing land Cover/Land use datasets for cropland area monitoring at the African continental scale. Remote Sensing, 2013; 5(1): 19–41.
[33] Chen J, Chen J, Liao A P, Cao X, Chen L J, Chen X H, et al. Global land cover mapping at 30m resolution: A POK-based operational approach. ISPRS Journal of Photogrammetry and Remote Sensing, 2015; 103: 7–27.
[34] Pérez-Hoyos A, Rembold F, Kerdiles H, Gallego J. Comparison of global land cover datasets for cropland monitoring. Remote Sensing, 2017; 9(11): 1118.
[35] Feng X M, Fu B J, Piao S L, Wang S, Ciais P, Zeng Z Z, et al. Revegetation in China's Loess Plateau is approaching sustainable water resource limits. Nature Climate Change, 2016; 6(11): 1019–1022.
[36] Gong P, Li X C, Zhang W. 40-Year (1978-2017) human settlement changes in China reflected by impervious surfaces from satellite remote sensing. Science Bulletin, 2019; 64(11): 756–763.
[37] U. S. Geological Survey. Landsat-Earth observation satellites (ver. 1.2, April 2020): U. S. Geological Survey Fact Sheet 2015-3081, 2016; 4p.
[38] Adler-Golden S, Matthew M W, Bernstein L S, Levine R Y, Berk A, Richtsmeier S C, et al. Atmospheric correction for shortwave spectral imagery based on MODTRAN4. In: Proceedings of SPIE - International Symposium on Optical Science, Engineering and Instrumentation, 1999; 3753: 366315. doi: 10.1117/12.366315.
[39] European Space Agency. Sentinel-2 User Handbook, 2015. Available: https:// sentinel.esa.int/web/sentinel/user-guides. Accessed on [2024-02-24].
[40] Clevers J G P W, Kooistra L. Using hyperspectral remote sensing data for retrieving canopy chlorophyll and nitrogen content. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2012; 5(2): 574–583.
[41] Yang N, Li J H, Mo W B, Luo W J, Wu D, Gao W C, et al. Water depth retrieval models of East Dongting Lake, China, using GF-1 multi-spectral remote sensing images. Global Ecology and Conservation, 2020; 22: e01004.
[42] Laben C A, Brower B V. Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening. US, 2000; Patent notice No. US06011875A.
[43] Daughtry C S T, Walthall C L, Kim M S, de Colstoun E B, McMurtrey J E Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. Remote Sensing of Environment, 2000; 74: 229–239.
[44] Nicolaï B M, Beullens K, Bobelyn E, Peirs A, Saeys W, Theron K I, et al. Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. Postharvest Biology and Technology, 2007; 46: 99–118.
[45] Yue L W, Shen H F, Zhang L P, Zheng X W, Zhang F, Yuan Q Q. High-quality seamless DEM generation blending SRTM-1, ASTER GDEM v2 and ICESat/GLAS observations. ISPRS Journal of Photogrammetry and Remote Sensing, 2017; 123: 20–34.
[46] 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 Sensing of Environment, 2002; 83: 195–213.
[47] Gao B. NDWI - A normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sensing of Environment, 1996; 58: 257–266.
[48] Ghimire B, Rogan J, Miller J. Contextual land-cover classification: incorporating spatial dependence in land-cover classification models using random forests and the Getis statistic. Remote Sensing Letters, 2010; 1(1): 45–54.
[49] Pal M, Mather P M. An assessment of the effectiveness of decision tree methods for land cover classification. Remote Sensing of Environment, 2003; 86(4): 554–565.
[50] Eisavi V, Homayouni S, Yazdi A M, Alimohammadi A. Land cover mapping based on random forest classification of multitemporal spectral and thermal images. Environmental Monitoring and Assessment, 2015; 187(5): 291.
[51] Tian S H, Zhang X F, Tian J, Sun Q. Random forest classification of wetland landcovers from multi-sensor data in the arid region of Xinjiang, China. Remote Sensing, 2016; 8(11): 954.
[52] Liaw A, Wiener M. Classification and regression by randomForest. R News, 2002; 2: 18–22.
[53] Geiß C, Aravena Pelizari, P, Blickensdörfer L, Taubenbock H. Virtual Support Vector Machines with self-learning strategy for classification of multispectral remote sensing imagery. ISPRS Journal of Photogrammetry and Remote Sensing, 2019; 151: 42–58.
[54] Zendehboudi A, Baseer M A, Saidur R. Application of support vector machine models for forecasting solar and wind energy resources: A review. Journal of Cleaner Production, 2018; 199: 272–285.
[55] Dou J, Yunus A P, Fbui D T, Merghadi A, et al. Assessment of advanced random forest and decision tree algorithms for modeling rainfall-induced landslide susceptibility in the Izu-Oshima Volcanic Island, Japan. Science of The Total Environment, 2019; 662: 332–346.
[56] Bui D T, Pradhan B, Lofman O, Revhaug I. Landslide susceptibility assessment in Vietnam using support vector machines, decision tree, and Naive Bayes Models. Mathematical Problems in Engineering, 2012; 2012: 974638.
[57] Ara Rahman S, Chakrabarty D. Sediment transport modelling in an alluvial river with artificial neural network. Journal of Hydrology, 2020; 588(5): 125056.
[58] Dong Y H, Fu Z T, Peng Y Q, Zheng Y J, Yan H J, Li X X, et al. Precision fertilization method of field crops based on the Wavelet-BP neural network in China. Journal of Cleaner Production, 2020; 246: 118735.
[59] Hemeida A M, Hassan S A, Mohamed A A, Alkhalaf S, Mahmoud M M, Senjyu T, et al. Nature-inspired algorithms for feed-forward neural network classifiers: a survey of one decade of research. Ain Shams Engineering Journal, 2020; 11(3): 659–675.
[60] Wang L, Wang P X, Liang S L, Qi X, Li L, Xu L X. Monitoring maize growth conditions by training a BP neural network with remotely sensed vegetation temperature condition index and leaf area index. Computers and Electronics in Agriculture, 2019; 160: 82–90.
[61] Han X, Wei Z, Zhang B Z, Li Y N, Du T S, Chen H. Crop evapotranspiration prediction by considering dynamic change of crop coefficient and the precipitation effect in back-propagation neural network model. Journal of Hydrology, 2021; 596: 126104.
[62] Zhang J Q, Tian H R, Wang P X, Tansey K, Zhang S Y, Li H M. Improving wheat yield estimates using data augmentation models and remotely sensed biophysical indices within deep neural networks in the

Guanzhong Plain, PR China. Computers and Electronics in Agriculture, 2022; 192: 106616.

[63] Dong J W, Xiao X M, Menarguez M A, Zhang G L, Qin Y W, Thau D, et al. Mapping paddy rice planting area in northeastern Asia with Landsat 8 images, phenology-based algorithm and Google Earth Engine. Remote Sensing of Environment, 2016; 185: 142–154.

[64] Foody G M. Status of land cover classification accuracy assessment. Remote Sensing of Environment, 2002; 80(1): 185–201.

[65] Congalton R G. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment, 1991; 37(1): 35–46.

[66] Tan J B, Zuo J Q, Xie X Y, Ding M Q, Xu Z K, Zhou F B. MLAs land cover mapping performance across varying geomorphology with Landsat OLI-8 and minimum human intervention. Ecological Informatics, 2021; 61: 101227.

[67] Ngo K D, Lechner A M, Vu T T. Land cover mapping of the Mekong Delta to support natural resource management with multi-temporal Sentinel-1A synthetic aperture radar imagery. Remote Sensing Applications: Society and Environment, 2020; 17: 100272.

[68] Prasad P, Loveson V J, Chandra P, Kotha M. Evaluation and comparison of the earth observing sensors in land cover/land use studies using machine learning algorithms. Ecological Informatics, 2022; 68: 101522.

[69] Hermosilla T, Wulder M A, White J C, Coops N C. Land cover classification in an era of big and open data: optimizing localized implementation and training data selection to improve mapping outcomes. Remote Sensing of Environment, 2022; 268: 112780.

[70] Zhao Y F, Zhu W W, Wei P P, Fang P, Zhang X W, Yan N N, et al. Classification of Zambian grasslands using random forest feature importance selection during the optimal phenological period. Ecological Indicators, 2022; 135: 108529.

[71] Fang H L, Zhang Y H, Wei S S, Li W J, Ye Y C, Sun T, et al. Validation of global moderate resolution leaf area index (LAI) products over croplands in northeastern China. Remote Sensing of Environment, 2019; 233: 111377.

[72] Liu T, Abd-Elrahman A, Morton J, Wihelm V L. Comparing fully convolutional networks, random forest, support vector machine, and patch-based deep convolutional neural networks for object-based wetland mapping using images from small unmanned aircraft system. GIScience and Remote Sensing, 2018; 55(2): 243–264.

[73] Liu W S, Yang J, Li P X, Han Y, Zhao J Q, Shi H T. A novel object-based supervised classification method with active learning and random forest for PolSAR imagery. Remote Sensing, 2018; 10(7): 1092.

[74] Roth K L, Roberts D A, Dennison P E, Peterson S H, Alonzo. The impact of spatial resolution on the classification of plant species and functional types within imaging spectrometer data. Remote Sensing of Environment, 2015; 171: 45–57.

[75] Lesiv M, Fritz S, McCallum I, Tsendbazar N, Herold M, Pekel J-F, et al. Evaluation of ESA CCI prototype land cover map at 20m. IIASA Working Paper, Laxenurg: IIASA, 2017; WP-17-021. Available: http://pure.iiasa.ac.at/14979/. Accessed on [2024-02-24].

[76] Wang C S, Shu Q Q, Wang X Y, Guo B, Liu P, Li Q Q. A random forest classifier based on pixel comparison features for urban LiDAR data. ISPRS Journal of Photogrammetry and Remote Sensing, 2019; 148: 75–86.

[77] Xie Z L, Chen Y L, Lu D S, Li G Y, hen E X. Classification of land cover, forest, and tree species classes with ZiYuan-3 multispectral and stereo data. Remote Sensing, 2019; 11(2): 164.

[78] Xia J Y, Wan S Q. Independent effects of warming and nitrogen addition on plant phenology in the Inner Mongolian steppe. Annals of Botany, 2013; 111(6): 1207–1217.

Appendix

Table S1 Parameters of GF-1 satellite

Payloads	Bands	Spectral range/μm	Spatial resolution/m	Swath width/km	Repetition cycle/d
WFV	Band 1	0.45–0.52	16	800	4
	Band 2	0.52–0.59	16		
	Band 3	0.63–0.69	16		
	Band 4	0.77–0.89	16		
PMS	Pan 1	0.45-0.90	2	60	41
	Band 1	0.45–0.52	8		
	Band 2	0.52–0.59	8		
	Band 3	0.63–0.69	8		
	Band 4	0.77–0.89	8		

Note: WFV and PMS represent wide field of view and panchromatic/multi-spectral sensors, respectively.

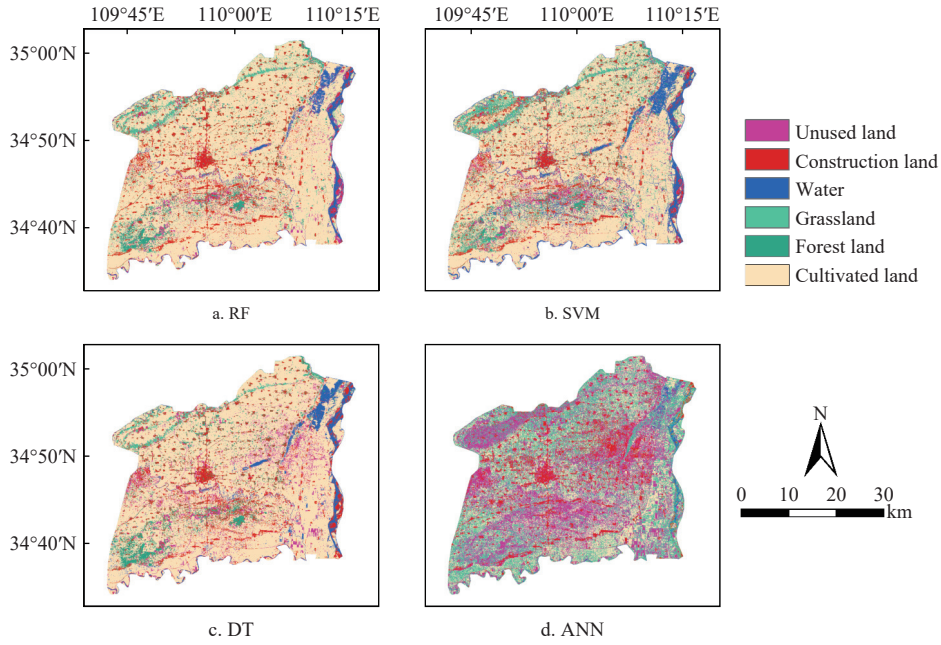
Table S2 Reclassification of global land cover datasets

Datasets	Before classification	After classification
LC10m/LC30m	Cropland	Cultivated land
	Forest	Forest land
	Grassland	Grassland
	Shrub land	Grassland
	Wetland	Unused land
	Water	Water
	Impervious area	Construction land
	Snow and Ice	Unused land

Table S3 Confusion matrix of different machine learning models

Algorithm		CL	FL	GL	W	I	UL	Total
RF	CL	50	1	0	1	1	1	54
	FL	0	24	3	0	0	0	27
	GL	5	5	20	0	1	2	33
	W	1	0	0	26	0	2	29
	I	0	0	0	0	35	3	38
	UL	4	0	3	3	2	18	30
	Total	60	30	26	30	39	26	211
SVM	CL	45	4	0	2	0	3	54
	FL	0	23	4	0	0	0	27
	GL	5	5	18	0	2	3	33
	W	0	0	0	28	1	0	29
	I	1	0	0	3	33	1	38
	UL	4	1	3	4	2	16	30
	Total	55	33	25	37	38	23	211
DT	CL	48	1	0	3	1	1	54
	FL	1	23	3	0	0	0	27
	GL	8	4	18	0	2	1	33
	W	1	0	0	27	1	0	29
	I	2	1	0	0	33	2	38
	UL	3	0	3	3	1	20	30
	Total	63	29	24	33	38	24	211
ANN	CL	42	5	1	2	1	3	54
	FL	4	23	0	0	0	0	27
	GL	8	8	10	1	2	4	33
	W	4	1	0	23	0	1	29
	I	1	0	7	0	25	5	38
	UL	1	1	3	2	1	22	30
	Total	60	38	21	28	29	35	211

Note: The column labels and row labels represent the observed and predicted values, respectively.



Note: RF, random forest; SVM, support vector machine; DT, decision tree; ANN, artificial neural network.

Figure S1 Spatial distribution of land use and land cover from Landsat OLI using different machine learning classification models in Dali County, Shaanxi Province, China

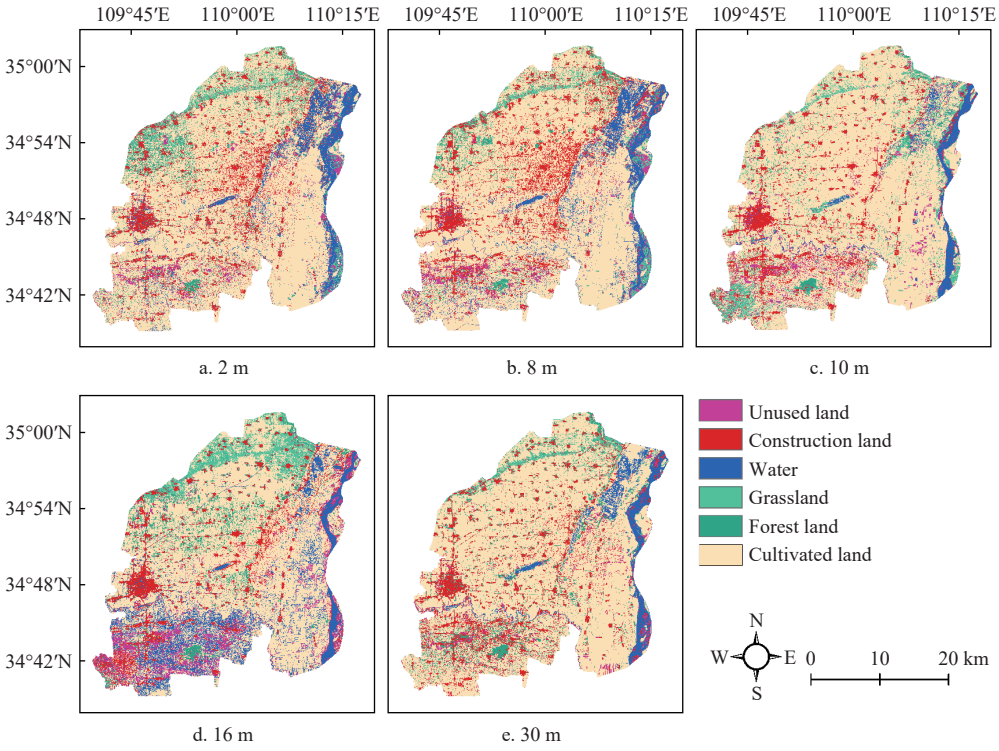


Figure S2 Spatial distribution of land use and land cover obtained from different-resolution images in a portion of Dali County, Shaanxi Province, China