



Article

Using Eco-Geographical Zoning Data and Crowdsourcing to Improve the Detection of Spurious Land Cover Changes

Ling Zhu ^{1,2,*} , Dejun Gao ¹, Tao Jia ³ and Jingyi Zhang ¹

¹ School of Geomatics and Urban Spatial Information, Beijing University of Civil Engineering and Architecture, Beijing 100044, China; 2108160220011@stu.bucea.edu.cn (D.G.); 2108160216001@stu.bucea.edu.cn (J.Z.)

² Beijing Key Laboratory of Urban Spatial Information Engineering, Beijing 100038, China

³ School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430072, China; tao.jia@whu.edu.cn

* Correspondence: zhuling@bucea.edu.cn

Abstract: To address problems in remote sensing image change detection, this study proposes a method for identifying spurious changes based on an eco-geographical zoning knowledge base and crowdsourced data mining. After preliminary change detection using the super pixel cosegmentation method, eco-geographical zoning is introduced, and the rules of spurious change are collected based on the knowledge of expert interpreters, and from statistics on existing land cover products according to each eco-geographical zone. Uncertain changed patches with a high possibility of spurious change according to the eco-geographical zoning rule were published in the form of a map service on an online platform, and then crowd tagging information on spurious changed patches was collected. The Hyperlink-Induced Topic Search (HITS) algorithm was used to calculate the spurious change degree of changed patches. We selected the northern part of Laos as the experimental area and the Chinese GF-1 Wide Field View (WFV) images for change detection to verify the effectiveness of the method. The results show that the accuracy of change detection improves by 23% after removing the spurious changes. Spurious changes caused by clouds, river water turbidity, spectral differences in cultivated land before and after harvest, and changes in shrubs, grassland, and forest density, can be removed using an eco-geographical zoning knowledge base and crowdsourced data mining methods.

Keywords: eco-geographical zoning; knowledge base; crowdsourcing data mining; land cover; spurious change



Citation: Zhu, L.; Gao, D.; Jia, T.; Zhang, J. Using Eco-Geographical Zoning Data and Crowdsourcing to Improve the Detection of Spurious Land Cover Changes. *Remote Sens.* **2021**, *13*, 3244. <https://doi.org/10.3390/rs13163244>

Academic Editors: Shin Nagai, Timoaki Miura, Narumasa Tsutsumida and Brian Alan Johnson

Received: 19 June 2021

Accepted: 12 August 2021

Published: 16 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 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

Land cover constitutes a series of complex surface elements covered by natural structures and artificial buildings. It has specific temporal and spatial attributes, and its shape and state can change at a variety of spatial and temporal scales [1]. Land cover is a concept that has emerged with the development of remote sensing technology, and remote sensing is the only effective means of large-area land cover mapping [2]. Since the 1980s, the international scientific community has paid considerable attention to the problem of global land cover mapping and has developed various global, regional, and national 1 km, 300 m, 30 m, and 10 m resolution land cover products [3]. Land cover changes with time, mainly due to natural and human impacts. Natural forces, such as continental drift, glacier action, floods, and tsunamis, in addition to human forces, such as the transformation from forest to agriculture land, urban expansion, and the dynamic change in forest planting, have changed the types of land cover. Multitemporal remote sensing image change detection technology can be used to monitor the changes in the ecological environment and track urban development, which is of great significance for the study of the interaction between humans and the natural environment [4].

1.1. Land Cover Change Detection Methods

There are two kinds of land cover change detection methods used in the remote sensing community, namely the post-classification method and the direct comparison method [5,6].

Land cover products obtained from remote sensing image classification inevitably contain a large number of false classifications or uncertain pixels due to spectral confusion, image resolution limitations, and the complexity of land features [7]. The classical supervised and unsupervised classification methods are mature, but their accuracy is not high, ranging from about 60% to 70% [8]. For example, the accuracy of Globecover V2 products is about 67.5%, that of global land cover (GLC) 2000 is about 68.6%, and that of the Moderate-Resolution Imaging Spectroradiometer (MODIS) land cover product is 74.8% [9]. The overall accuracy of Globeland30 classification products is relatively high, reaching 83.51% [10], which can be attributed to manual correction and the assistance of expert knowledge. Artificial intelligence methods, such as deep learning, have developed rapidly in recent years, but are still in the research stage [11]. The post-classification change detection method compares land cover products obtained from the classification of remote sensing images at different times. However, due to the limited accuracy of the land cover products in each period, the accuracy of the change results obtained by the superposition of the two products is lower and pertains to the product (multiplication) of the accuracy of the two land cover products.

Another means of extracting land cover change is by directly extracting the changes using bitemporal image change detection. Remote sensing images of the same area obtained at different times can be used to identify and determine the types of surface changes and their spatial distribution. Most land cover change detection is based on an analysis of spectral, shape, texture, and other characteristic factors in images. The methods used for change information extraction include mathematical analysis, feature space transformation, feature classification, feature clustering, and neural networks [12]. However, remote sensing images only reflect the instantaneous state of the surface, resulting in many errors and uncertainties. The main challenge in change detection is how to preserve real changes while eliminating spurious changes [5].

1.2. Main Reasons for Spurious Changes

Due to the global complexity and diversity of land types, numerous phenomena from the same object have different spectra, and different objects have the same spectrum, inevitably resulting in spurious changes in change detection. The main reasons for spurious changes are as follows:

- (1) Bitemporal images are usually acquired under different atmospheric conditions, including different sun heights, sun angles, and off-nadir distances, which result in varying illumination levels in dual images. Sood et al. [13] focused on the problem of change caused by different shadows in high mountain areas and found that, in mountainous regions, satellite imagery change detection is more complex due to the presence of a rugged topography, slope variability, and topographic (shadow) effects. As a result, a topographically controlled subpixel-based change detection model was used to reduce the detection of spurious pixels. To remove the false change in vegetation resulting from building shadows caused by the shooting angle and sunlight in urban high-resolution aerial change detection, Zhou et al. [14] adopted a series of spatial analyses at different levels and reduced spurious changes. Spatial analysis was mainly conducted using self-adaptive morphology;
- (2) Large differences in dryness or humidity between the two phases—for example, abundant precipitation in one phase and scarce precipitation in another phase—result in a large difference in soil moisture, and a change in the position of the boundary lines of rivers and lakes. Similar examples occur for other land types, including grassland before and after wilting, and wetlands when water levels change;

- (3) Bitemporal images with various types of phenological phenomena may be acquired at different collection times that cover the same area. For example, the ground reflectivity values of farmland in various agricultural pheno-phases are different before and after harvesting. The landscape is different in summer and winter due to the changes that occur across the four seasons. Deciduous forests have varied spectral characteristics in different seasons. Grassland turns yellow and withers in autumn and winter, and turns green in spring. Lu et al. [15] used Landsat images from March and June in different years in the Shanxi Province of China for change detection to test a method intended to weaken spurious change. June was the peak period, whereas no crops were planted in March. The evident differences in the spectrum of cultivated land found by Lu et al. were due to the influence of phenology; hence, change detection has many false changes;
- (4) Forestland is affected by pests or diseases. *Dendrolimus punctatus*, *Bursaphelenchus xylophilus*, desert locust, pine leaf bee, and gypsy moth all lead to corresponding changes in spectral brightness and greenness [16];
- (5) Water turbidity also has an influence. Water bodies can be classified as clear, semi-turbid, or turbid [17]. A change in water turbidity between two phases causes change detection to identify a false change. The authors of [5] list the spurious changes caused by changes in water turbidity.

1.3. Review of Spurious Change Alleviation Methods

In a direct comparison of the gray levels of the corresponding pixels in bitemporal images, such as in the difference, ratio, and change vector methods, the result of change detection is easily affected by the above factors, resulting in spurious changes. Several scholars designed different methods that do not directly compare the gray values of the corresponding pixels of the two-phase image but convert them into another feature space for comparison. Lv et al. [18] used an Adaptive Histogram Trend (AHT) similarity approach to assess the changes. This method generates an adaptive region by comparing the spectral similarity between the central pixel and its surrounding pixels. A histogram for the pixels within the adaptive region is then built. Through comparison, the pairwise histogram trends generate an image that shows the magnitude of the changes. Spurious change noise can also be reduced by AHT. In another study, Chen et al. [5] used spectral gradient space to calculate the magnitude of land cover change to detect change/no-change areas. This method conducts change detection in spectral gradient space and determines the land cover change types by pattern matching them with a knowledge base of reference Spectral Gradient Difference (SGD) patterns. Spurious changes caused by interference factors that alter the spectral values without changing the shape characteristics of the spectrum can then be eliminated. The author pointed out that SGD is not suitable for analyzing the images acquired in different phenological seasons because the vegetation spectrum does not have a stable shape in different seasons.

In addition to bitemporal remote sensing images, several studies introduced auxiliary data, particularly Normalized Difference Vegetation Index (NDVI) time series data, to remove spurious changes caused by phenological factors. Liu et al. [19] integrated multiple shape parameters, including the phase angle cumulant, baseline cumulant, relative cumulation rate, and zero-crossing rate of the NDVI multitemporal curve and spectrum, to detect land cover changes. Spurious changes caused by phenological differences in Landsat images could then be eliminated. Lu et al. [15] proposed the Object-Based Spatial and Temporal Vegetation Index Unmixing Model (OB-STVIUM) to solve the issue of phenological differences in land cover change detection. Firstly, Landsat images are segmented by multiscale segmentation. Then, the OB-STVIUM is employed to disaggregate MODIS NDVI time series into Landsat image object time series through a spatial analysis and linear mixing theory. Lastly, NDVI-GD (the gradient difference) is used to calculate the shape and value differences of NDVI time series and determine change and no-change objects. Combined with annual MODIS NDVI 16-day composite grid data time series information,

land cover types that are easily affected by phenological factors and result in spurious changes, such as cultivated land, can be correctly identified in the change detection results. Hu et al. [6] found that using only spectral values or spectral change vector features from two independent time profiles is insufficient to meet the needs of accurate land change detection. The focus of their study was on the phenological differences between remote sensing images. Change vector (CV) analysis in posterior probability space was applied to conduct the land change detection of bitemporal images of western China. NDVI temporal variation analysis was used to increase the reliability of the change results. Supported by the Google Earth Engine (GEE), the classification and regression tree analysis method was used to supervise and classify images of the latter phase to ascertain the types of changes that had taken place.

Previous methods of spurious change detection were mainly based on remote sensing image data, through parameters extracted from multiple images or by considering neighboring pixels to reduce the spurious changes caused by direct comparisons, or by combining them with time series information to eliminate the spurious changes caused by vegetation phenology. Applications of spurious change recognition combined with nonremote sensing data are not common.

1.4. Concept of Eco-Geographical Zoning to Aid in Spurious Change Detection

Land cover is an element relating to resources and the environment and depends on the geographical environment of a specific region. The geographical environment comprises regional differentiation and spatial correlations. The distribution of resources and environmental elements can be analyzed through geoscience knowledge. Therefore, the development of geoscience knowledge, including geographical units, has become an important research area to ensure the accuracy of land cover classification and change detection in large areas. Remote sensing geoscience analysis focuses on information composition, including multisource remote sensing information composition analysis, and remote sensing information and nonremote sensing information composition analysis. Previously, scholars often used auxiliary data to improve the accuracy of products for land cover mapping [20]. However, the use of expert system knowledge and reference auxiliary data in land cover mapping is sporadic and unsystematic, and no special system exists to manage and accumulate all of the auxiliary data for reuse. Therefore, establishing an auxiliary database to store, manage, and transfer expert knowledge is necessary. Due to its global characteristics, rich mining information, and stable land types, the eco-geographical division, or the so-called eco region, can be used to build a knowledge base to improve the classification accuracy of remote sensing images.

The eco-geographical zoning refers to the partition of geographical divisions combined with the interference of human activities, using relevant principles and methods of ecology, to integrate and distinguish the differences and similarities between various ecological regions, and thus identify the ecological environment units [21].

Each eco-geographical zone has a similar biological community, and the land cover types are relatively stable over certain periods. However, in the case of changes, certain rules apply that allow for the change trend to be used as a reference. In addition, the related knowledge with regard to land cover and change detection can be mined according to the different geographical attributes of the eco-geographical zone.

In this study, the concept of eco-geographical zoning is introduced, and the spurious change rules of land cover are stored according to different zones. Possible spurious changes are identified by the eco-geographical zoning rule base. Supplementary eco-environmental factors can effectively reduce the errors in change detection, making the results more consistent with the eco-geographical zoning rules, thereby improving accuracy.

1.5. Concept of Crowdsourced Data Mining to Aid in Spurious Change Detection

Since the increase in the popularity of the Internet, geographic information has increased exponentially [22], resulting in a large amount of geographic information data.

In particular, due to the significant progress of location technology and the wide spread of Web 2.0 technology, citizens passively or actively collect and contribute an increasing amount of geographic data. This trend continues to expand because almost any information can now be geotagged [23].

Crowdsourced geographic data are generally defined as open geospatial data that are actively or passively obtained by numerous nonprofessionals [24]. Representative crowdsourced geographic data include OSM (Open Street Map) data from an online map collaboration planning platform, social media check-in data from various social networking sites (such as Weibo or Instagram), mobile photo signaling data, and floating car Global Positioning System (GPS) trajectory data. These kinds of crowdsourced geographic data contain rich information and knowledge, but often need to be processed before mining and use [25]. The development of crowdsourced geographic information and the public demand for geographic information complement one another. With the increase in the public's familiarity with, and demand for, data with spatial location information, crowdsourcing geographic data will play an increasingly important role in future social development. For example, social media check-in data are rich in geographic information and semantic information and have important research value for the study of land cover spatial distribution, such as the spatial structure of cities and the distribution of functional areas [26].

At present, research on the processing of crowdsourced geographic data mainly focuses on data quality analysis and evaluation, providing complete, accurate, and reliable data for subsequent applications. Crowdsourced geographic data are obtained by non-professional users. As the quality of these data is varied, the data quality has a high degree of uncertainty; thus, the integrity, accuracy, and reliability of the data should be fully considered when they are used [27,28].

Crowdsourcing technology also has significant potential in remote sensing land cover and land use research. This technology also benefits the collection of large quantities of reference data needed for accuracy validation. Clark and Aide [29] developed the Virtual Interpretation of Earth Web-Interface Tool, which is a collaborative browser-based tool for the crowdsourced interpretation of reference data from high-resolution imagery to produce and verify the accuracy of land use/land cover (LULC) maps. The principal component is the Google Earth plugin, which allows users to visually estimate the percentage cover of seven basic LULC classes. See et al. [30] and Schepaschenko et al. [31] used crowdsourced data from Geo-wiki for model training and validation. These studies demonstrated the immense potential of crowdsourced data as training samples and for validation in land cover integration. In addition to Geo-wiki, several other crowdsourced data sources are used in the land cover field, such as Degree Confluence Project (DCP) volunteer pictures, the Flickr photo-sharing website, and the Land Cover Validation Platform (LACO) Wiki open access online portal.

Crowdsourced data, especially Point of Interest (POI) data, can also be used to help determine the type of land cover. Human-made and non-man-made land cover can be distinguished. Xing et al. [32] proposed using POI data for the automatic classification of land cover. POI textual data were analyzed to calculate word and topic distributions using the latent Dirichlet allocation topic model. The Support Vector Machine (SVM) algorithm was applied with topic distributions of POIs to construct a land cover classification model. Fan et al. [33] proposed an optimized method for impervious surface mapping based on Sentinel-2 multispectral imagery and OSM POI data.

In this study, crowdsourcing data were used to identify spurious changes in land cover. In addition, the spurious change degree value was calculated using the Hyperlink-Induced Topic Search (HITS) algorithm based on the change degree value provided by volunteers in order to remove spurious changes and improve the accuracy of change detection.

This study has two main contributions. The first is that a novel solution to the problem of spurious changes in remote sensing change detection is proposed. The combination of an eco-geographical zoning database and online crowdsourced data mining is adopted.

The second contribution is the use of the crowdsourced data mining method to remove spurious changes in land cover from satellite remote sensing images. To the best of the authors' knowledge, this is the first time this has been undertaken. Crowdsourced data mining is an advanced big data method that can enrich land cover research and data production in the remote sensing community.

2. Method

2.1. Method Overview

The workflow of land cover spurious change detection using an eco-geographical zoning knowledge base and crowdsourced data mining is shown in Figure 1. The whole process can be divided into two parts: preliminary change detection and spurious change identification. Based on image preprocessing, the method of cosegmentation change detection is used to extract the changed patches of land cover, and the changed patches are classified to obtain the types of land cover before and after the change. In the spurious change identification step, first, the offline eco-geographical zoning knowledge base is used to identify the spurious changed patches. The identified spurious changed patches are then screened, and several are selected for online publishing. The spurious changed patches are further removed using the annotations of crowdsourcing volunteers, and the optimized changed patches are finally obtained.

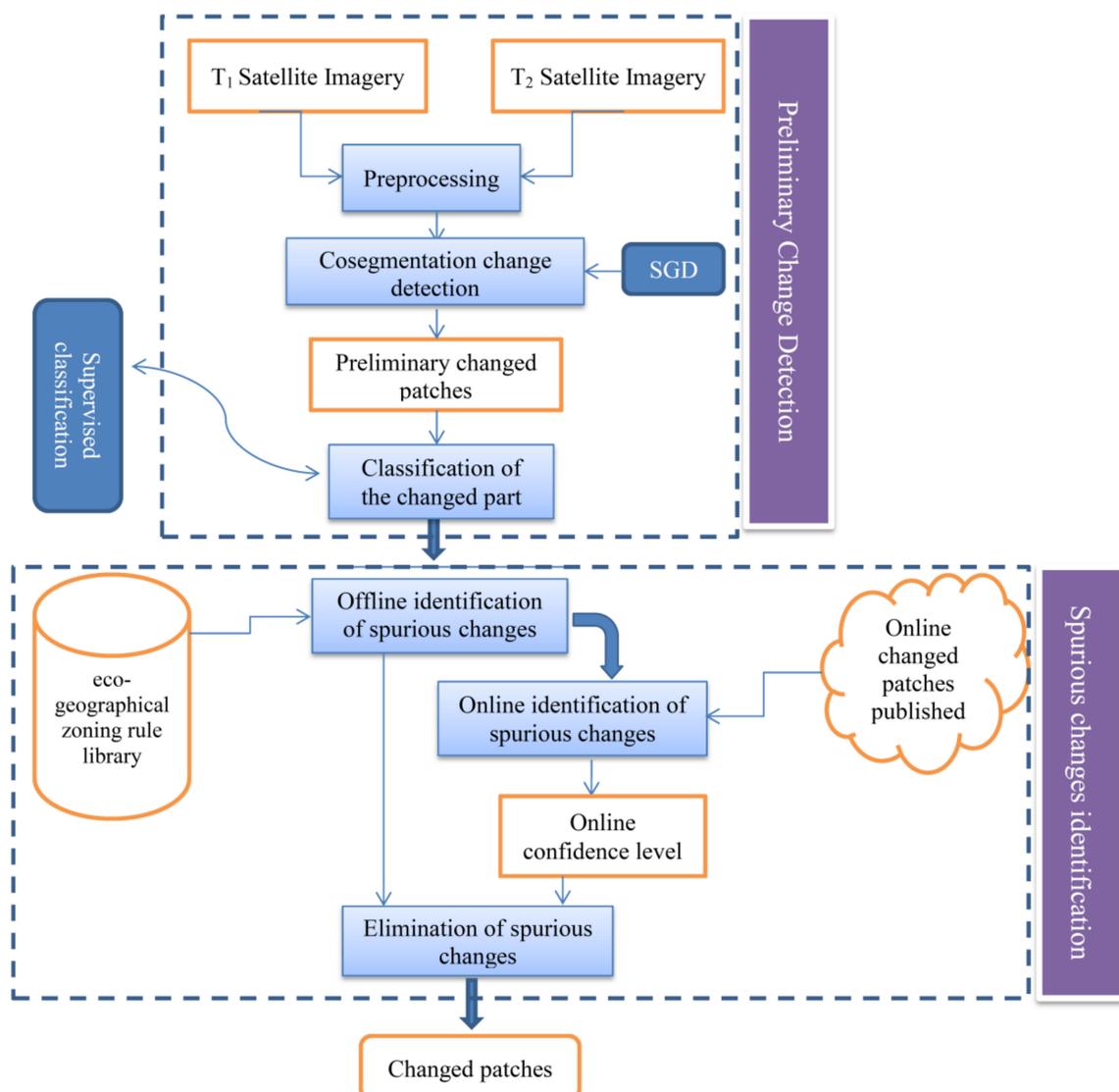


Figure 1. Workflow.

2.2. Image Preprocessing and Change Detection

2.2.1. Preprocessing

The GF-1 satellite is equipped with two 2 m resolution panchromatic/8 m resolution multispectral cameras and four 16 m resolution multispectral cameras. The combination of a high resolution and large width can be simultaneously realized on a single satellite. The imaging width of the 2 m resolution is greater than 60 km, and that of the 16 m resolution is greater than 800 km. Only 16 m resolution Wide Field View (WFV) images were used in this study. The main band parameters are shown in Table 1.

Table 1. GF-1 Satellite WFV camera parameters.

Multispectral Cameras	Band Number	Wave Range (μm)	Spatial Resolution (m)	Width (km)	Sway Ability	Revisit Period (Days)
	6	0.45–0.52		800		
	7	0.52–0.59	16	(four-camera combination)	$\pm 35^\circ$	2
	8	0.63–0.69				
	9	0.77–0.89				

The purpose of image preprocessing is to perform data registration and radiometric correction, weaken the influence of the external imaging environment, and simplify the problem of change detection. The specific steps are as follows:

- (1) Ortho correction: Ortho correction was carried out according to the Rational Polynomial Coefficients (RPC) file contained in the image download file and the corresponding Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) 30 m resolution DEM (Digital Elevation Model) of the area;
- (2) Radiometric calibration: The radiometric calibration coefficient of the GF-1 image was used to calibrate the image, and the image acquisition time, sensor type, aerosol mode, and other parameters for Fast Line-of-Sight Atmospheric Analysis of Spectral Hypercubes (FLAASH) atmospheric correction were inputted;
- (3) Geometric correction: The GF-1 WFV image has the characteristics of a wide coverage, large field of view, and complex geometric distortion. When using commercial software for image geometric correction, several local areas still appear to have deviations of a few pixels, resulting in an inability to achieve the accurate registration of two-phase images. For local geometric deformations, the conventional method is to use a large number of high-precision, evenly distributed control points to build a local correction model. Therefore, for GF-1 WFV images, the key to achieving a precise geometric correction is the acquisition of a large number of control points with high precision and a uniform distribution. In this study, the method of Shan et al. [34] was adopted. Based on a global mosaic image from the Landsat Thematic Mapper (TM), the hierarchical registration method was based on the Forstner operator and template matching. This method uses numerous high-precision, evenly distributed control points obtained by hierarchical matching to construct the Delaunay triangulation, which effectively solves the problem of geometric precision correction of GF-1 WFV images. The steps include GF-1 WFV image feature point extraction, hierarchical automatic image matching, control point homogenization, and triangulation correction. The registration error is less than 2 pixels.

2.2.2. Super Pixel Cosegmentation Change Detection

The cosegmentation algorithm in computer vision can segment the same or similar objects from multiview images of the same scene [35]. The algorithm can mine more image information because it uses the relationships between images. The central idea of cosegmentation is to optimize the energy function to obtain the optimal segmentation of objects of interest in the image group. The optimization of the energy function is realized

by finding the minimum cut of the network flow graph. The minimum cut/maximum flow algorithm is a method of energy function optimization. The energy function of cosegmentation change detection [36] consists of two parts, namely change features and image feature items, as shown in Equation (1):

$$E = \lambda E_1 + E_2 \quad (1)$$

where E_1 is the change feature item, E_2 is the image feature item, and λ is the weight of the change feature item used to balance the proportion of the image feature and the change feature in the formula.

Firstly, the image features and change features of each pixel are calculated. Then, the pixel image is mapped to the network flow graph. Finally, the minimum cut of the network flow graph is obtained to optimize the energy function.

The network flow diagram constructed in the minimum cut/maximum flow method of cosegmentation change detection takes each pixel as a node in the graph. Thus, the number of algorithm iterations is closely related to the total number of pixels in the graph, reducing the operation efficiency of the algorithm. Zhu et al. [36] introduced super pixel segmentation to make it suitable for the change detection of a large scene and enhance its practicability. Adjacent pixels with homogeneity were segmented into large pixels, that is, super pixels. Then, the super pixels were used as primitives for cosegmentation, thereby greatly improving the efficiency of the algorithm.

Zhu et al. [36] obtained the change features of each pixel by calculating the CV value, but the difference in the spectral values between the two phases was large due to the low quality of the GF-1 WFV image. The study area was located in a subtropical monsoon climate region, and cloudless images were difficult to obtain during the rainy season. Due to different climates, terrains, and solar height angles, cloud and mountain shadows can affect change detection. If the gray pixel level of two-phase images is directly compared to extract the changes, they will be affected by these factors and produce spurious changes. Therefore, in this study, for the change intensity image of the cosegmentation, we did not use the CV value. Instead, SGD, which was developed by Chen et al. [5], was used. SGD combines the adjacent spectral bands of multispectral remote sensing images to reflect the trend of the change in the reflectance of the adjacent spectrum by calculating the slope of the spectral band. Then, the slope series of all spectral bands was used to describe the shape of the whole spectrum curve. Through the comparison of the spectral curve shape of the same pixels in the two-phase images, change detection was carried out. The slope of the spectral curve is defined as follows:

$$g_{k,k+1} = \frac{\Delta R_{k,k+1}}{\Delta \lambda} = \frac{R_{k+1} - R_k}{\lambda_{k+1} - \lambda_k} \quad (2)$$

where k and $k + 1$ are two adjacent bands; R_{k+1} and R_k are corresponding gray levels; $\Delta R_{k,k+1}$ is the difference between the two gray levels; λ_{k+1} , λ_k are the central wavelengths of the k and $k + 1$ bands; $\Delta \lambda$ is the difference between the central wavelengths.

The spectral slopes of n bands are expressed by vectors, called spectral slope vectors, expressed as $G = (g_{1,2} \ g_{2,3} \ \cdots \ g_{n-1,n})$. $g_{n-1,n}$ is the spectral slope between bands $n - 1$ and n .

The difference in the spectral slope of a pixel in two phases is as follows:

$$|\Delta G| = \sum_{k=2}^n |g_{k-1,k}^2 - g_{k-1,k}^1| \quad (3)$$

where $g_{k-1,k}^2$ represents the spectral slope of $k - 1$ and k bands in the latter phase, whereas $g_{k-1,k}^1$ represents the spectral slope of $k - 1$ and k bands in the previous phase. The larger the value of $|\Delta G|$, the more likely it is to change. SGD can inhibit the spurious changes

caused by brightness differences, surface humidity differences, and turbidity differences in water bodies.

As shown in Table 1, the GF-1 WFV image has only four bands, namely blue, green, red, and near infrared, in which the spectral information is not rich. All four bands of the GF-1 image were used to calculate the difference in the spectral slope in this study. Cosegmentation is based on the graph cut theory and considers the image information, such as spectrum and texture, and mines the spatial neighborhood information between pixels. The energy function depends on the image and change features. Due to the large amount of computation in cosegmentation change detection, principal component transformation and gray value simplification were carried out by the algorithm, which reduce the importance of the spectral and radiometric resolution in the algorithm. The number of bands in the satellite image has little effect on the cosegmentation algorithm and is only reflected in the change intensity map [36].

2.2.3. Change Patch Classification

The initially extracted changed patches should be classified to obtain the from-to change information and provide type information for subsequent spurious change recognition. The land cover classification system follows the classification standard of the project “identification and extraction of typical resources and environmental elements and quantitative remote sensing technology” funded by the Chinese government. This study is one part of the project. The taxonomy is based on the Globeland30 land cover product and contains 10 first-level categories. The category name, code, and contents are shown in Table 2.

The algorithm for changed patch classification was developed by the Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences. The method extracts typical resource elements supported by geoscience knowledge. Its extraction accuracy is no lower than 85%. Different classification strategies were adopted for different land cover types. The classifier mainly uses the SVM method of supervised classification. In addition, NDVI time series data, a DEM, and derived slope data were combined. For more detail regarding the classification method, please refer to the website of the national earth system science data center (<http://www.geodata.cn/data/datadetails.html?dataguid=9657818&docid=825>) (accessed on 11 August 2021).

2.2.4. Two-Phase Overlay Analysis

The results are in the form of patches because the change detection algorithm uses super pixel cosegmentation. In a change patch, the classification results of the two phases of all pixels may not be uniform; that is, several types of land cover may exist in a change patch, which creates challenges for the subsequent use of the eco-geographical zoning rule base to judge the spurious changes. Therefore, after image classification, the image classification result and the change detection patches are overlapped and analyzed, and the changed patches are subdivided according to the types of the two phases of the land cover map, such that the types of the T_1 and T_2 phase images in each change patch are uniform.

The overlay analysis is shown in Figure 2. Two types of land cover exist in the T_1 phase of a change patch: forests (020) and bare land (080). Two types of land cover exist in T_2 , artificial land (070) and shrubs (030). Therefore, the types of land cover are not uniform. The patch is transformed into four patches through the overlay analysis of the two phases and, thus, land cover types with uniform T_1 and T_2 phases can be obtained.

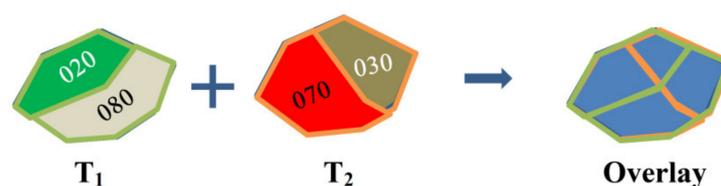


Figure 2. Overlay analysis of the changed patch of two phases.

Table 2. Land cover definition and code.

Code	Land Cover Class	Definition
010	Cultivated land	Refers to land used for cultivating crops. Paddy fields, irrigated upland, rainfed upland, vegetable land, cultivated pastures, and greenhouse land are included in this category.
020	Forests	Refers to land covered with trees, the top density of which occupies over 30%. Deciduous broadleaf forests, evergreen broadleaf forests, deciduous coniferous forests, evergreen coniferous forests, mixed forest and sparse woodlands, the top density of which covers 10–30%, are included in this category.
030	Shrublands	Refers to land covered with shrubs areas with a cover density over 30%. Mountain shrubs, deciduous and evergreen shrubs, and desert jungles in desert areas with a cover density over 10% are included in this category.
040	Grasslands	Refers to land covered by herbaceous vegetation. This includes all kinds of grasslands with more than 10% grass coverage, including shrub grasslands dominated by grazing and sparse forest grasslands with less than 10% forest coverage.
050	Wetlands	Refers to the junctions of land and water areas, which are constantly covered by biogas or hygrophyte plants and shallow water or wet soils. Inland marshes, lake marshes, river floodplain wetlands, forest/shrub wetlands, peat bogs, mangroves, salt marshes, etc., are included in this category.
060	Water bodies	Refers to liquid water-covered regions in land areas. Rivers, lakes, reservoirs, pit ponds, etc., are included in this category.
070	Artificial surfaces	Refers to surfaces formed by human-made activities. All kinds of habitation in urban and rural areas, industrial and mining areas, transportation facilities, etc., are included in this category, whereas interior contiguous green land and water bodies are not included.
080	Bare land	Refers to naturally covered land with a vegetation cover density lower than 10%. Deserts, sand, gravel ground, bare rocks, saline and alkaline land, etc., are included in this category.
090	Permanent snow and ice	Refers to land covered by permanent snow, glaciers, and icecaps.
100	Tundra	Refers to land covered by lichen, moss, hardy perennial herbs and shrubs in cold and high mountain areas. Shrub tundra, grass tundra, wet tundra, alpine tundra, barren tundra, etc., are included in this category.

2.3. Eco-Geographical Zoning Rule Base

2.3.1. Eco-Geographical Zoning

To assist in the updating of land cover based on remote sensing change detection, this study used the terrestrial eco-regions of the world [37] established by the World Wild

Fund for nature conservation as the basic framework of the global eco-geographical zoning knowledge base. The eco-regions divide the world into eight biogeographic realms and 14 biomes. Based on these two basic layers, 867 eco-regions are identified. Each division has a unique six-digit code. The naming rules are as follows: the first two digits are the geographical division where the region is located, the middle is the type of biological community present in the area, and the last two digits are distinguished according to the natural attributes; the sequence codes are 01, 02, etc. For example, PA0101 is a broad-leaved and mixed forest and is located in Guizhou Plateau, among tropical and subtropical moist broad-leaved forests found in Eurasia. The global ecological division is shown in Figure 3. Different colors indicate different zones.

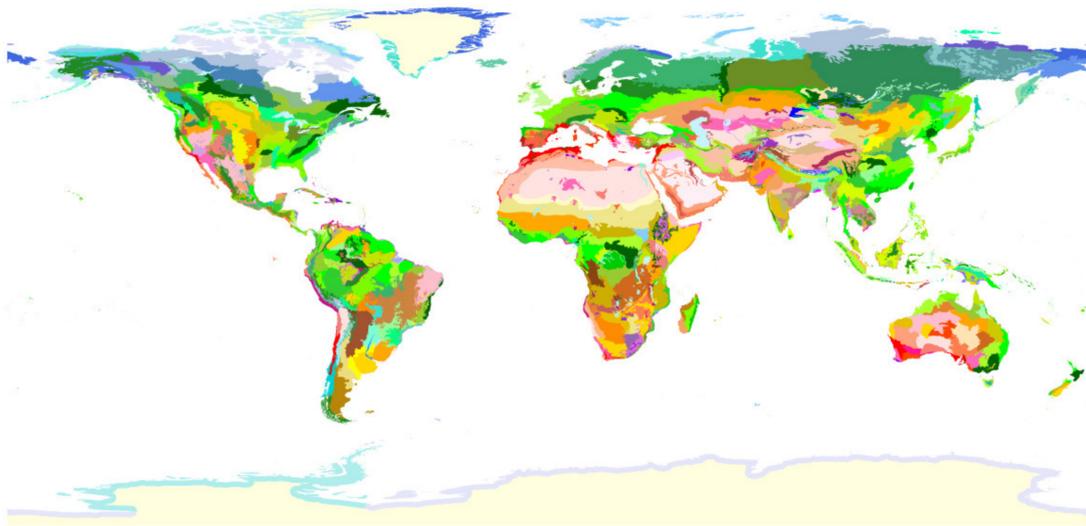


Figure 3. Map of 867 global ecological zones (different colors indicate different zones).

2.3.2. Eco-Geographical Zoning Rule Base

Framework of Eco-Geographical Zoning Rule Base

Rules are long-term information or knowledge extracted from the experience of experts. The quality of rules can directly affect the performance of an expert system. In this study, the local geoscience knowledge was sorted and mined, and the spurious change rules of each zone were summarized and used to identify and eliminate spurious changes. Among the knowledge in the area, the knowledge related to land cover includes the following:

- (1) Each eco-geographical zone is divided according to certain geographical attributes, and the information of different land types can be collected based on these different geographical attributes (such as elevation, slope, precipitation, and NDVI);
- (2) Different land cover types in each eco-geographical zone follow certain rules, and these rules can be extracted from the geographical attribute information collected on various categories.

As the world is divided into 867 different eco-geographical zones, data collection and storage rules involve a large amount of work. Zhu et al. [3] designed an object-oriented means to establish the rule base. Considering eight geographical divisions and 14 biological communities, and the natural attributes such as slope, elevation, temperature, moisture, and NDVI, the knowledge base is divided into four layers, as shown in Figure 4. The inheritance and derivation relationship between each layer, from top to bottom, is established. The third layer contains 64 large ecological divisions which come from combining of eight geographical divisions and 14 biological communities. The names of the ecological divisions are IM01, PA01, etc. The right-hand side of the third layer shows the subclasses of the natural attributes. General spurious change rules, collected

mainly from expert interpreters, are stored in this layer. The fourth layer consists of the 867 eco-geographical zones.

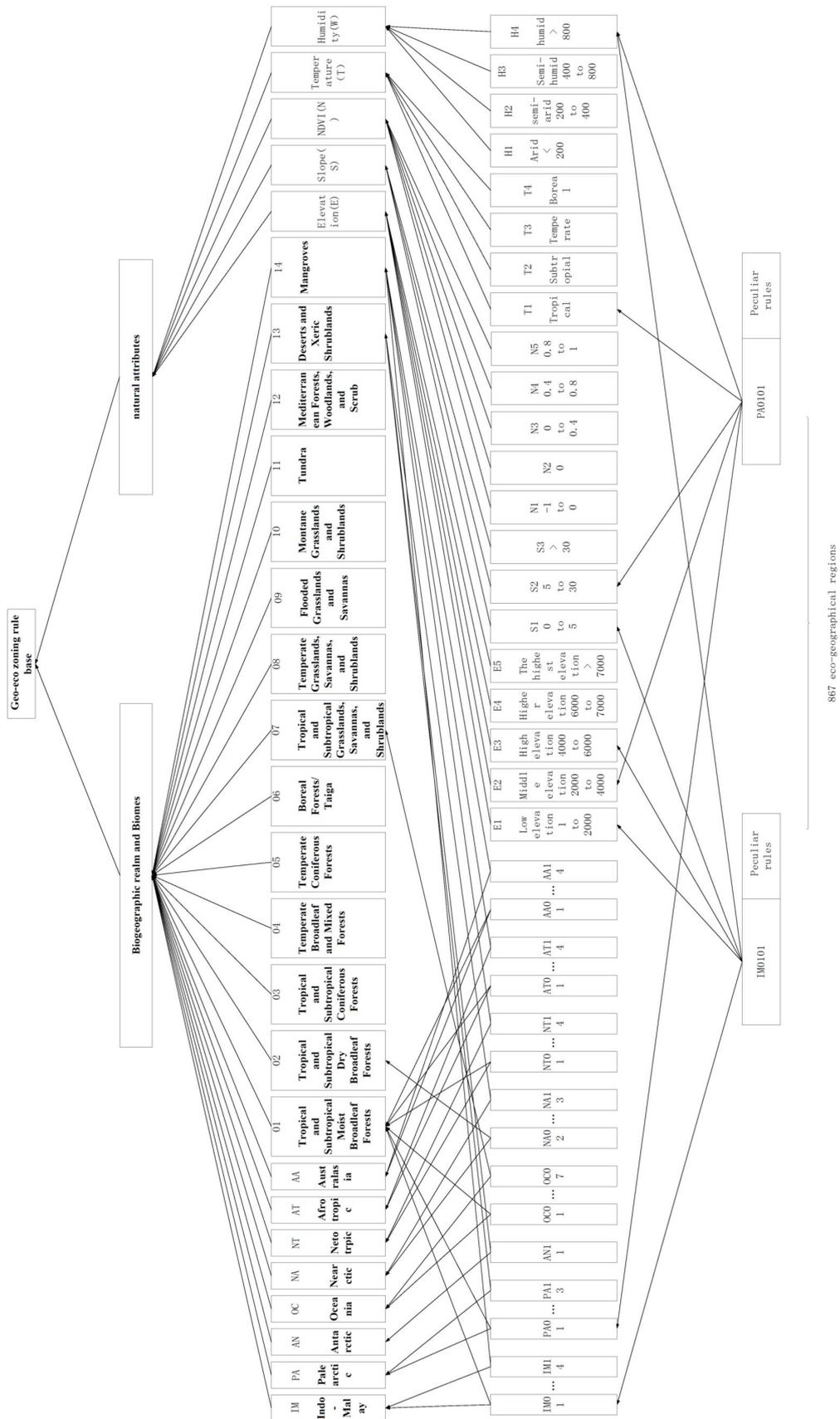


Figure 4. Framework of the global eco-geographical zone rule base (from Zhu et al. [3]).

Prior Knowledge Collection

To effectively identify spurious changes in land cover, it is particularly important to define our prior knowledge. Two kinds of methods can be used to collect spurious changes from prior knowledge: one involves the use of expert knowledge and the other involves mining rules from the existing land cover products.

Using different attribute conditions, prior knowledge of the global land cover distribution is collected from remote sensing expert interpreters and expressed as knowledge rules that can be used to judge spurious changes.

As the spurious change rules are stored in a four-tier structure, the rules of the upper level are inherited by the next level to reduce the repetition of rules. General rules apply to a wide range of zones. In addition, mining the unique rules for each eco-geographical zone is necessary. Although the accuracy of existing land cover products is limited, the statistical regularity of land cover at different times can express the change laws of land cover. In this study, we used the method of mining rules from existing land cover products. We compared the same pixel locations of the existing land cover products in the same region at different times and counted the types of conversion relationships of pixels to obtain the land cover class conversion rules of each eco-geographical zone. These statistics were carried out sequentially according to the eco-geographical zone to reflect the internal laws of different divisions. The specific method was as follows: first, the types of land cover in the zone were extracted, and if the number of types was n , a matrix of size $n \times n$ was established. The two matched phases of land cover products were cut to each area according to the eco-geographical zone. The land cover types of the two periods were sequentially counted, and the number accumulated at the corresponding position in the matrix. Finally, the conversion matrix of the land cover type of each eco-geographical region was obtained. The land cover class conversion matrix is shown in Table 3, which assumes that the eco-geographical zone has six types of land cover, namely cropland, forests, grass, shrubs, wetlands, and water.

Table 3. Conversion matrix of the land cover type of each eco-geographical region.

Phase 2 Phase 1	Cropland	Forests	Grass	Shrubs	Wetlands	Water	Sum
Cropland	N11	N12	N13	N14	N15	N16	$\sum N1$
Forests	N21	N22	N23	N24	N25	N26	$\sum N2$
Grass	N31	N32	N33	N34	N35	N36	$\sum N3$
Shrubs	N41	N42	N43	N44	N45	N46	$\sum N4$
Wetlands	N51	N52	N53	N54	N55	N56	$\sum N5$
Water	N61	N62	N63	N64	N65	N66	$\sum N6$
Sum	$\sum N1$	$\sum N2$	$\sum N3$	$\sum N4$	$\sum N5$	$\sum N6$	

For example, in Table 3, N11 is the number of pixels from cropland to cropland in the two-phase land cover products. $\sum N1$ is the sum of N11 to N16 and represents the number of cropland pixels in the phase 1 land cover product.

Then, the number statistics table was transformed into the probability statistics table, as shown in Table 4, where P11 is the probability of cropland remaining as cropland and P12 is the probability of cropland becoming a forest, etc. The sum of the transition probability is 1 for each row and column in the matrix.

In the transformation probability matrix of land cover types in the eco-geographical zone, transformation relationships with low probabilities, such as values close to 0, are generally considered to be nonexistent, that is, spurious changes.

The expression of spurious change rules is carried out in the form of a premise and a conclusion, which are mapped to the table in the database and stored in the common object-relational database.

Table 4. Conversion matrix of the land cover type of each eco-geographical region.

Land Cover	Cropland	Forests	Grass	Shrubs	Wetlands	Water	Sum
Cropland	P11	P12	P13	P14	P15	P16	1
Forests	P21	P22	P23	P24	P25	P26	1
Grass	P31	P32	P33	P34	P35	P36	1
Shrubs	P41	P42	P43	P44	P45	P46	1
Wetlands	P51	P52	P53	P54	P55	P56	1
Water	P61	P62	P63	P64	P65	P66	1
Sum	1	1	1	1	1	1	1

Spurious change rules are established layer by layer according to the framework structure of the knowledge base of the eco-geographical zone. The spurious change rules are represented by a six-digit code; the first three digits represent the category code before the change, and the last three digits represent the category code after the change. The fourth layer of the model contains the spurious change rules of each small eco-geographical zone, including the rules inherited from the upper layer and the rules from statistics on existing land cover products according to each eco-geographical zone.

2.4. Online Crowdsourced Spurious Change Tagging

2.4.1. HITS Crowdsourced Data Mining

HITS is an algorithm for spurious change detection based on crowdsourced data. This algorithm was first used to sort documents according to the link information in a group of documents [38]. It has been adopted by most search engine websites and proven to be a classic, effective method. The basic idea of the algorithm is that a high-level hub often points to many other document nodes, whereas a high-quality authority is pointed to by many document nodes. The hub node and authority node are mutually reinforcing; that is, a high-level hub node often points to many high-level authority nodes, whereas a high-level authority node is often pointed to by many high-level hub nodes.

The HITS algorithm generally uses hub and authority values to represent the hub and authority of a node, respectively. In each iteration, the algorithm has two basic steps: The first step is to update the authority value of each node to the sum of the hub values of the nodes connected to it. The second step is to update the hub value of each node to the sum of the authority values of the nodes connected to it. Suppose a directed graph $G(V, E)$, V is a set of nodes and E is a set of directed edges. If m and n belong to V and the directed edges (m, n) belong to E , then a direct connection relationship exists between nodes m and n . The out-degree value of m is the total number of nodes pointing to other nodes, and the in-degree value of n is the total number of nodes pointing to n nodes. If the authority value of each node p is denoted as $a(p)$ and the hub value is denoted as $h(p)$, then each iteration has the following calculation:

- (1) Calculate the authority value of the node, that is:

$$|\Delta G| = \sum_{k=2}^n |g_{k-1,k}^2 - g_{k-1,k}^1| \quad (4)$$

- (2) Normalize $a(m)$ by:

$$\alpha(m) = a(m) / \sqrt{\sum_{q \in n} [h(q)]^2} \quad (5)$$

- (3) The hub value of the node is calculated:

$$h(n) = \sum_{(m,n) \in E} a(m) \quad (6)$$

(4) Normalize $h(n)$ by:

$$h(n) = h(n) / \sqrt{\sum_{q \in m} [a(q)]^2} \quad (7)$$

In general, to obtain stable hub or authority values, Formulas (4) and (6) must be iteratively calculated until the algorithm converges, which is usually achieved by setting a small threshold. For example, if the threshold value is set to 10^{-5} , then the algorithm is convergent if the difference between the hub value and the authority value of all nodes of two consecutive times is less than the threshold value.

The HITS algorithm was applied to the detection of spurious change based on crowdsourced data in this study. Firstly, a bipartite graph for the evaluation of the connection relationship between multiple users and the changed patches must be established, as shown in Figure 5. As one type of node, the users of the crowdsourcing network have no association with one another. Moreover, as another type of node, the changed patches have no association with one another. However, a spurious change evaluation relationship exists between these two types of nodes; that is, a change patch with a high spurious change degree is more likely to be given a higher spurious change degree value by experienced users, and an experienced user's spurious change degree value is more reliable.

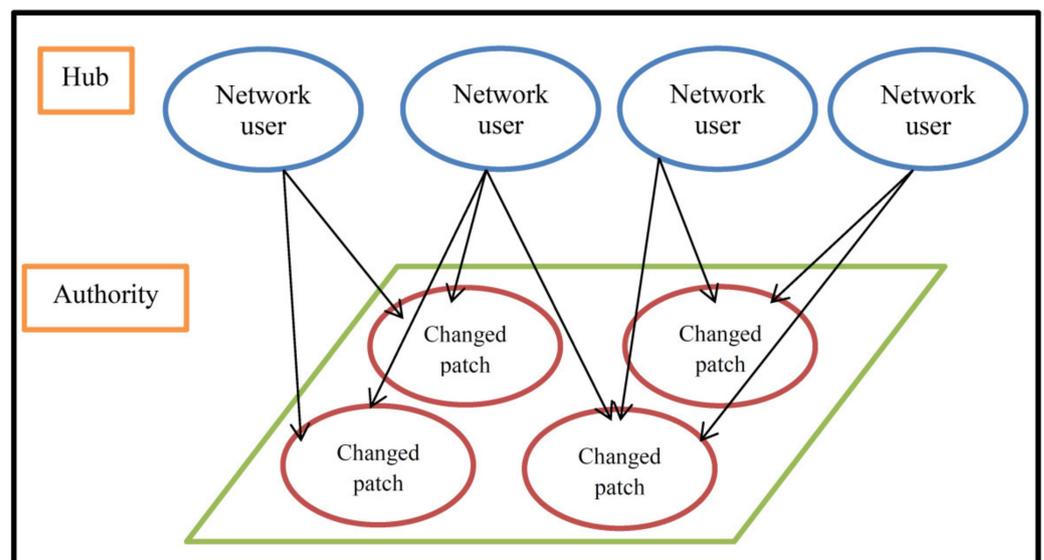


Figure 5. Bipartite graph of multiple users and changed patches.

Based on the bipartite network between the network user and the change patch, this study introduces the idea of the HITS algorithm to calculate the hub value of the network user and the authority value of the change patch to judge the spurious change degree of the patch, and thus identify the most likely spurious change patch. Network users not only evaluate the change patch, but also provide a quantitative score for the spurious change degree according to their own experience and knowledge. In the bipartite network, the score value of the spurious change degree of the change patch is the weight value of the directed edge. The higher the weight, the greater the probability of false changes in the change patch. The lower the weight, the smaller the probability of false changes in the change patch.

To describe this algorithm, a weighted network matrix W , whose elements W_{mn} represent the spurious change degree score of hub node m to authority node n , was defined. According to the basic principle of the HITS algorithm, the nodes meet the following definition in a limited number of iterations:

- (1) $a(m)$ is node m 's authority value:

$$\alpha(m) = \sum_{(m,n) \in E} h(n) \times W_{mn} \quad (8)$$

- (2) After the standardization of (m) , the following results are obtained:

$$\alpha(m) = a(m) / \sqrt{\sum_{(m,q) \in E, q \in n} [h(q) \times W_{mq}]^2} \quad (9)$$

- (3) $h(n)$ is the hub value of node n :

$$h(n) = \sum_{(m,n) \in E} a(m) \times W_{mn} \quad (10)$$

- (4) After the standardization of $h(n)$, the following results are obtained:

$$h(n) = h(n) / \sqrt{\sum_{(m,q) \in E, q \in m} [a(q) \times W_{mq}]^2} \quad (11)$$

The convergence condition of the algorithm is the same as the original HITS algorithm, which is that the change in the hub and authority values of all nodes in two consecutive iterations is less than the threshold value. One of the important reasons for using the weighted HITS algorithm is that it more easily finds the difference between the head hub node and the authority node compared with the original HITS algorithm, which makes finding changed patches with high spurious change degree values and experienced network users easier. To improve the operation efficiency of the algorithm and make it converge as quickly as possible, the following improvements to the normalization of the algorithm were made: In each iteration of the algorithm, the hub value and authority value were normalized according to Formulas (9) and (11), such that only the nodes connected with themselves needed to be normalized. Instead of summing and normalizing the entire hub node or authority node, the purpose of using the connection relationship of the small network to replace the connection information of the whole large network was achieved to improve network computing efficiency.

2.4.2. Design of Online Crowdsourcing Spurious Change Detection Platform

Hadoop is a distributed system infrastructure developed by the Apache Foundation. Users can develop distributed programs without understanding the underlying details of distributed computing. The core technologies of the Hadoop framework include the Hadoop Distributed File System (HDFS) and MapReduce. HDFS provides distributed storage for massive data, whereas MapReduce provides distributed computing for massive data. The implementation of the HITS algorithm in the Hadoop platform mainly uses MapReduce core technology to complete the calculation of the authority values of patches and the hub values of user nodes. A complete HITS iteration requires two map/reduce operations. The first calculates and normalizes the authority values of related nodes, and the second calculates and normalizes the hub values of related nodes. In detecting spurious changes in the patch, the evaluation score plays a key role in the confidence in the patch. The HITS algorithm, based on evaluation score weighting, inputs the weight as a multiplication factor when updating the hub and authority values in each iteration. A specific diagram illustrating this is shown in Figure 6.

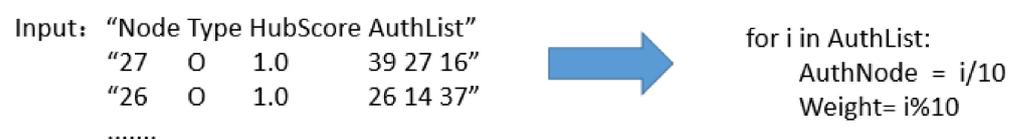


Figure 6. Iterative code of weighted HITS algorithm based on the Hadoop platform.

The functional module diagram of the platform is shown in Figure 7, which includes two main modules: the collection of evaluation data from crowdsourcing network users and the detection of spurious changes in the change patch. Network users can directly input their spurious change degree value and a corresponding text description of the change patch in the platform according to their own knowledge of the land cover change detection region. This information is collected and stored by the crowdsourcing network user acquisition module. The weighted HITS algorithm calculates a large amount of network user evaluation data in parallel and updates the authority score of the change patch and the central scores of users.

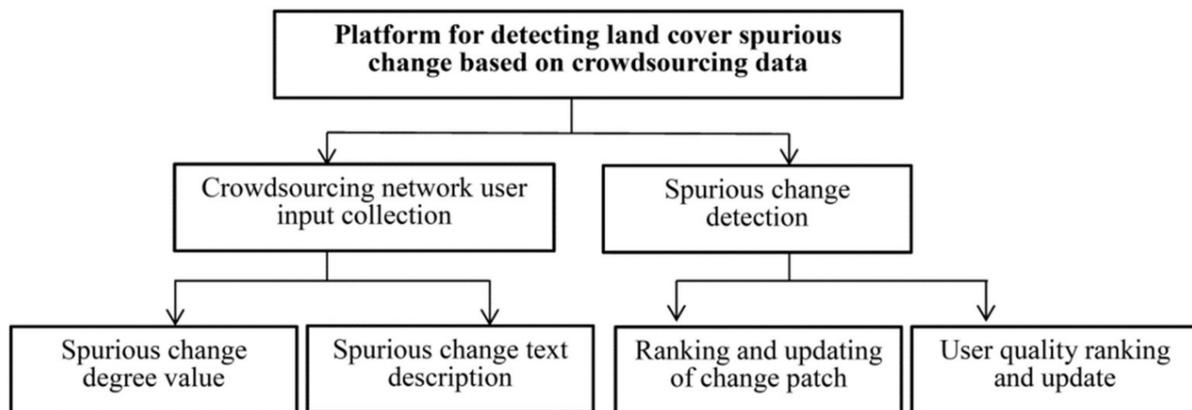


Figure 7. The function module diagram of the platform for detecting spurious land cover changes based on crowdsourced data.

(1) Design and implementation of user-input collection module for crowdsourcing network:

The crowdsourcing network user input collection module is used to collect users' evaluation data on the spurious change degree of a change patch, which mainly includes users' scores of the spurious change degree value of a change patch and a text description of the land cover class in the two time phases of the patch.

Figure 8 shows that after registering and logging in to the platform, users can move the map to locate their own locations, or locations they are familiar with or interested in. The map interface is divided into two parts: the left and right sides show a base map of the remote sensing image before and after the change, and the change patch layer is overlaid on the left and right base maps. In addition to the conventional base map zooming in and out, to help users better judge the change patch according to their own experience and knowledge, the platform adopts the layered release method for the base map remote sensing image, and uses tile map technology, which not only improves loading speed, but also clarity, and helps enhance users' interactive experience. Moreover, the platform adds a hidden button to help user better judge the changes in the geographical environment around the change patches.

(2) Design and implementation of the spurious change detection module:

The spurious change detection module mainly imports the user evaluation data stored in the database into the big data parallel processing framework based on SPARK (full name: Apache Spark). This is a general data analysis engine that specifically deals with a large amount of streaming data. Spark is a general parallel framework, similar to Hadoop MapReduce, which is available on an open-source basis from the UC Berkeley AMP lab. For details, please refer to the website <http://spark.apache.org/> (accessed on 11 August 2021). The technology uses the weighted spurious change detection HITS algorithm implemented in SPARK to obtain the degree scores of all changed patches and the scores of network users.

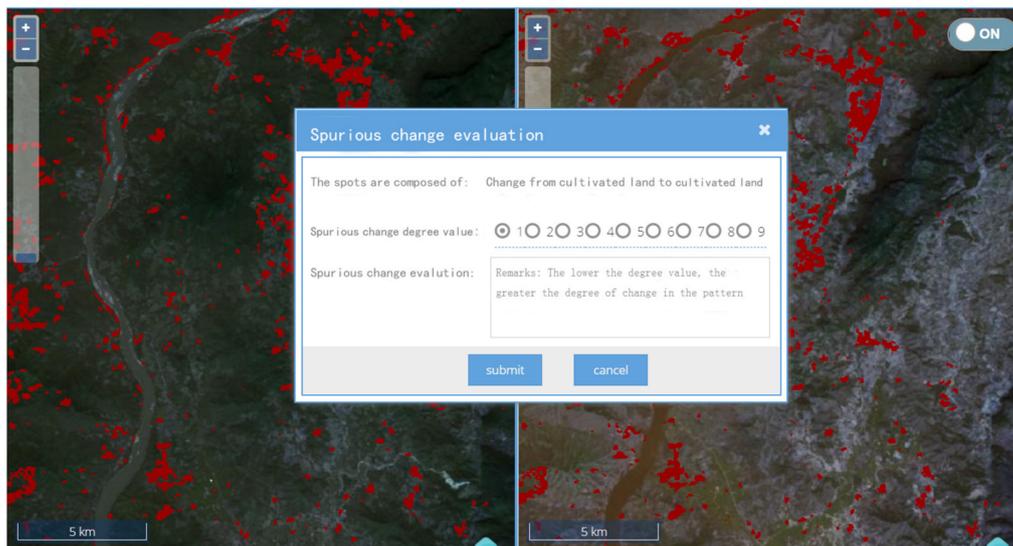


Figure 8. Acquisition interface of the online platform for the spurious change degree values of the changed patches.

In this module, the specific steps taken by the system administrator to detect spurious changes in the changed patches are as follows:

Firstly, the administrator uses the professional data transfer tool Sqoop (the full name of which is Apache Sqoop, an open-source tool. This tool can import data from a data warehouse or relational database into Hadoop HDFS or HBase for MapReduce to analyze. For more details, please refer to the website <http://sqoop.apache.org/> (accessed on 11 August 2021)) to import the change patch evaluation data stored in the PostgreSQL database into the big data distributed file storage system, HDFS, and completes the preprocessing operations, such as data deduplication and the deletion of null values.

Secondly, the administrator applies for the SPARK computing resources on the cluster, inputs the corresponding parameters, and calls the weighted HITS algorithm to calculate the spurious change degree of the imported data.

Thirdly, the administrator uses the Sqoop tool to again import the spurious change degree value and user centrality value of the change patch into the PostgreSQL database, and the platform updates the spurious change degree score and user quality score of the change patch according to the updated data.

Finally, the administrator analyzes the spurious change degree score and user quality score of the updated change map and completes the land cover change detection.

2.4.3. Screening Changed Patches for Online Crowdsourced Publishing

The online crowdsourcing method is based on the detection and removal of spurious changes from the eco-geographical zoning knowledge base. After online publishing, volunteers manually determine the degree of spurious changes of each image patch. To improve the degree of automation of change detection, due to the limitation of the network data volume, publishing all of the changed patches in a large change detection area is impossible. The current approach involves the publication of the suspicious changed patches screened, which are most likely to be spurious changed patches or patches with the greatest amount of uncertainty.

In general, vegetation change detection can easily produce uncertainty. Due to the seasonal changes in vegetation, single-scene remote sensing images with biased information and satellite images have difficulty obtaining vegetation extraction images of key phases. Obtaining images in the same phenological phase is also difficult. For example, cultivated land shows different spectral characteristics in different months. Due to the influence of mountain shadows, the spectral characteristics of vegetation change vary with different solar elevation angles.

Moreover, publishing these data online involves rules that refer to the type of land cover with a low degree of confidence in each partition transformation matrix from each eco-geographical zone. In addition, to overcome the influence of phenology on the spectrum of cultivated land, the changed patches of cultivated land in any phase are selected and published.

3. Data

3.1. Study Area

The selected study area is in the north part of Laos, which is an inland country located in the north of the Indochina Peninsula; 80% of the land is mountainous or plateaus and is mostly covered by forest. The terrain is high in the north and low in the south. Laos is bordered by the western Yunnan Plateau in Yunnan Province, China. In the east, Laos and Vietnam border a plateau composed of the Changshan Mountains. In the west lies the Mekong River Valley, basins, and small plains along the Mekong River and its tributaries. Laos has a tropical and subtropical monsoon climate, with a rainy season from May to October, and a dry season from November to the following April. The annual average temperature is about 26 °C. Laos has abundant rainfall. The reason for the selection of this study area is that the land is covered with different kinds of vegetation in fine particles, and change detection is difficult. Thus, spurious changes can be easily produced. The commonly used land cover products often achieve a low accuracy in this area. To verify the effectiveness of the method, an area with complex land cover types and distribution in northern Laos was selected for the test. Figure 9 shows the location of the selected experimental area in Laos.

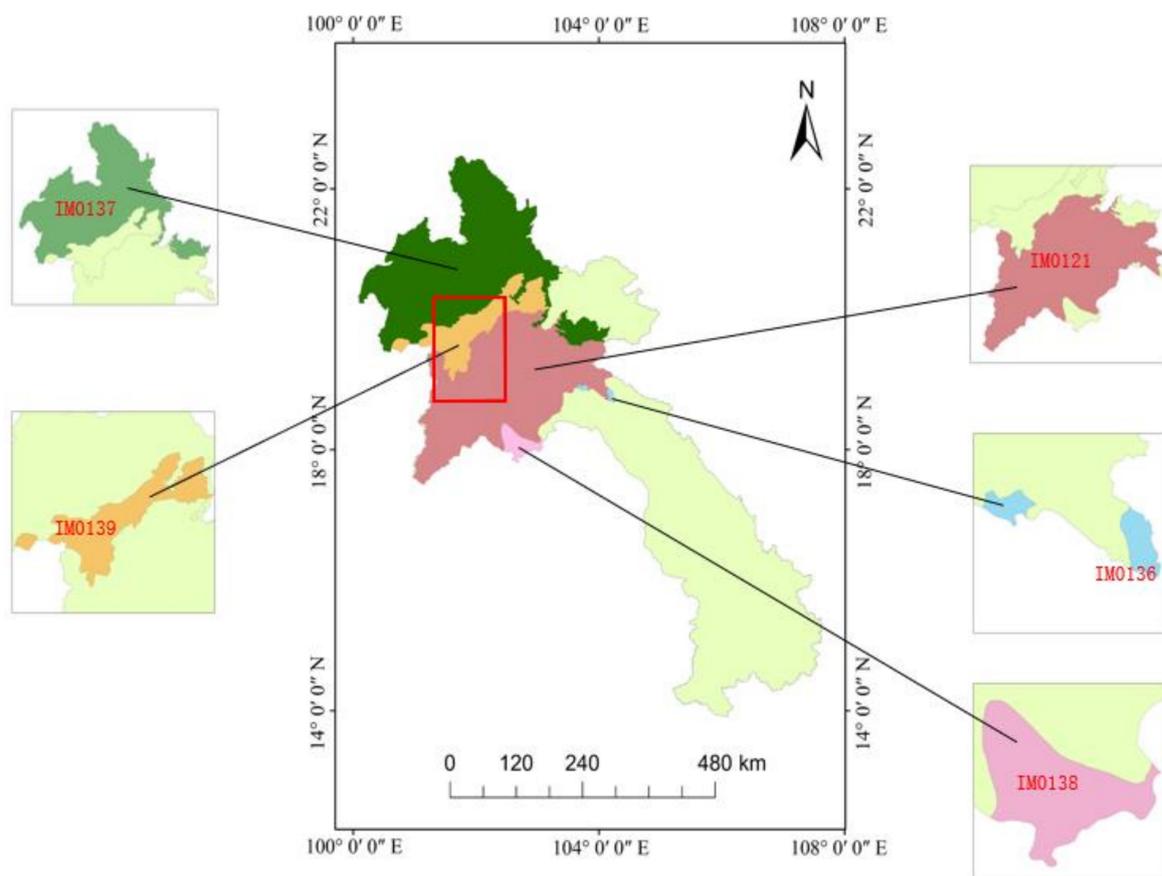


Figure 9. Eco-geographical zoning of the experimental area, where the red box is the study area.

The study area consists of three eco-geographical zones, namely the Luang Prabang montane rain forests (IM0121), the northern Indochina subtropical forests (IM0137), and

the northern Thailand–Laos moist deciduous forests (IM0139). In the Luang Prabang montane rain forests (IM0121), the biome comprises subtropical moist broadleaf forests. This zone has had limited biological exploration. Although more than 70% of this zone’s original montane vegetation has been converted into shrubs or degraded forests, the remaining area presents several of the best opportunities for large mammal conservation in Indochina. In the northern Indochina subtropical forests (IM0137), which is globally famous for its biodiversity, the biome comprises tropical and subtropical moist broadleaf forests. The northern Thailand–Laos moist deciduous forests (IM0139) contain large blocks of teak-dominated forests. The biome in this zone consists of tropical and subtropical moist broadleaf forests.

3.2. 2015 and 2020 GLC_FCS30 Products

The product selected for the land cover transformation probability statistics in 2015 and 2020 was GLC_FCS30. GLC_FCS30 was produced by the Aerospace Information Research Institute, Chinese Academy of Sciences [39], by combining time series from Landsat imagery and high-quality training data from the Global Spatial Temporal Spectra Library on the GEE computing platform. GLC_FCS30-2015 was validated using three different validation systems (containing different land cover details) using 44,043 validation samples. The validation results indicate that the GLC_FCS30-2015 achieved an overall accuracy of 82.5% and a kappa coefficient of 0.784 for the level-0 class (nine basic land cover types). GLC_FCS30 level-0 land cover types have a one-to-one correspondence with the classification system in this study. The GLC_FCS30 2015 and 2020 products of the northern part of Laos are shown in Figure 10. According to the related statistics (Table 5), forests and shrubs accounted for more than 80% of the land cover types in the experimental area in 2015 and 2020, whereas forests, shrubs, and cropland were the main land cover types. The proportions of all of these land cover types changed, considering the classification errors and the confusion caused when distinguishing between forests, shrubs, and grass, and the overall distribution and proportion changed slightly.

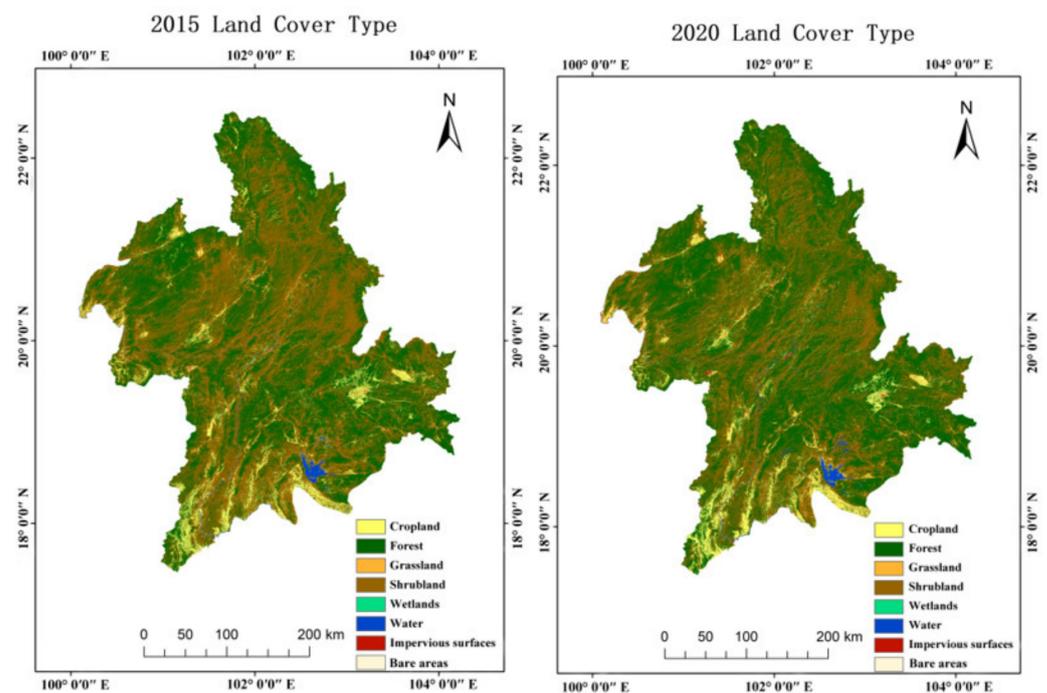


Figure 10. GLC_FCS30 2015 and 2020 products.

Table 5. Statistical table of GLC_FCS30 products in two phases (2015 and 2020).

Land Cover	2015		2020	
	Pixel Count	Proportion	Pixel Count	Proportion
Cropland	1,369,323	5.70%	1,468,986	6.12%
Forests	10,515,595	43.79%	12,025,140	50.12%
Grass	15,723	0.07%	10,602	0.04%
Shrubs	11,927,722	49.67%	10,182,737	42.44%
Wetlands	11,560	0.05%	3129	0.01%
Water	119,395	0.49%	163,235	0.68%
Impervious surfaces	55,088	0.23%	140,651	0.59%

The transformation probability matrix of the land cover types in each eco-geographical zone was calculated, and the results are shown in Table 6.

Table 6. Transformation probability matrix of land cover types in each eco-geographical zone based on GLC_FCS30 products (2015–2020).

IM0121								
Land Cover	Cropland	Forest	Grass	Shrub	Wetland	Water	Artificial	
Cropland	0.655662	0.138724	0.000083	0.172982	0.000358	0.011288	0.020905	
Forest	0.025709	0.797330	0.000019	0.174539	0.000116	0.001765	0.000522	
Grass	0.518326	0.292829	0.001112	0.037833	0.006076	0.128777	0.015047	
Shrub	0.061172	0.273834	0.000016	0.662740	0.000148	0.000636	0.001454	
Wetland	0.528367	0.172955	0.000632	0.153999	0.057118	0.066415	0.020514	
Water	0.068448	0.006729	0.000153	0.003885	0.001945	0.914736	0.004104	
Artificial	0.297608	0.013116	0.000000	0.014344	0.000054	0.007900	0.666979	
IM0137								
Land Cover	Cropland	Forest	Grass	Shrub	Wetland	Water	Artificial	Bare Land
Cropland	0.501647	0.178256	0.005283	0.261301	0.000155	0.005802	0.047551	0.000006
Forest	0.008681	0.829920	0.001236	0.159641	0.000025	0.000226	0.000271	0.000000
Grass	0.160823	0.580679	0.059788	0.121092	0.000208	0.034025	0.043385	0.000000
Shrub	0.037345	0.281683	0.001239	0.678366	0.000019	0.000279	0.001068	0.000000
Wetland	0.445588	0.208367	0.001345	0.137880	0.003228	0.015268	0.188324	0.000000
Water	0.112686	0.069924	0.032684	0.036265	0.004447	0.729387	0.014606	0.000000
Artificial	0.190790	0.005327	0.005062	0.007981	0.000038	0.001232	0.789494	0.000076
IM0139								
Land Cover	Cropland	Forest	Grass	Shrub	Wetland	Water	Artificial	
Cropland	0.489749	0.203831	0.002754	0.211230	0.000668	0.026297	0.065472	
Forest	0.013997	0.783616	0.000702	0.199963	0.000053	0.000365	0.001304	
Grass	0.252498	0.222519	0.022801	0.054469	0.005630	0.281070	0.161013	
Shrub	0.038376	0.240848	0.000739	0.717184	0.000066	0.000657	0.002129	
Wetland	0.323329	0.269231	0.000899	0.188893	0.061467	0.087527	0.068656	
Water	0.050101	0.018540	0.005219	0.005432	0.001748	0.904035	0.014924	
Artificial	0.189507	0.021239	0.000378	0.012762	0.000113	0.012064	0.763937	

In zone IM0121, the statistics show that approximately 65% of cultivated land, 79% of forests, 66% of shrubland, 91% of water bodies, and 66% of artificial surfaces did not change from 2015 to 2020. Approximately 52% and 29% of grasslands changed into cultivated land and forests, respectively. Approximately 52%, 17%, and 15% of wetlands changed into cultivated land, forests, and shrubland, respectively.

In zone IM0137, the statistics show that approximately 50% of cultivated land, 83% of forests, 68% of shrubland, 73% of water bodies, and 79% of artificial surfaces did not change from 2015 to 2020. Approximately 16% and 58% of grasslands changed into cultivated land and forests, respectively. Approximately 45%, 21%, and 14% of wetlands changed into cultivated land, forests, and shrubland, respectively.

In zone IM0139, the statistics show that approximately 49% of cultivated land, 78% of forests, 72% of shrubland, 90% of water bodies, and 76% of artificial surfaces did not change from 2015 to 2020. Approximately 25%, 22%, and 28% of grasslands changed into cultivated land, forests, and water bodies, respectively. Approximately 32%, 27%, and 19% of wetlands changed into cultivated land, forests, and shrubland, respectively.

In the three eco-geographical zones of the study area, consistent characteristics of land cover conversion were observed, and the proportion of grassland and wetland conversion into other types in the study area was larger. The main reason for this is that, in the GLC_FCS30 land cover products, the proportion of pixels of grasslands and wetlands is very small, as shown in Table 5, accounting for 0.07% and 0.05%, respectively. Moreover, GLC_FCS30 products based on time series Landsat images are limited by the resolution of Landsat series images, and forests, shrubs, and grass are easily confused in the spectrum, resulting in several commission and omission errors. Therefore, these types are liable to undergo large transformations into other land cover types. In Table 6, the change types with transformation probabilities of less than 0.0001 in the statistical results of each eco-geographical zone are marked with a gray background, and the likelihood of their occurrence is considered impossible.

3.3. GF-1 WFV Image

Obtaining images with less cloud coverage in Laos is difficult. The GF-1 images were acquired in February 2016 and February 2020, respectively, and the size of the image was $7532 \times 10,859$ pixels after preprocessing, as shown in Figure 11.

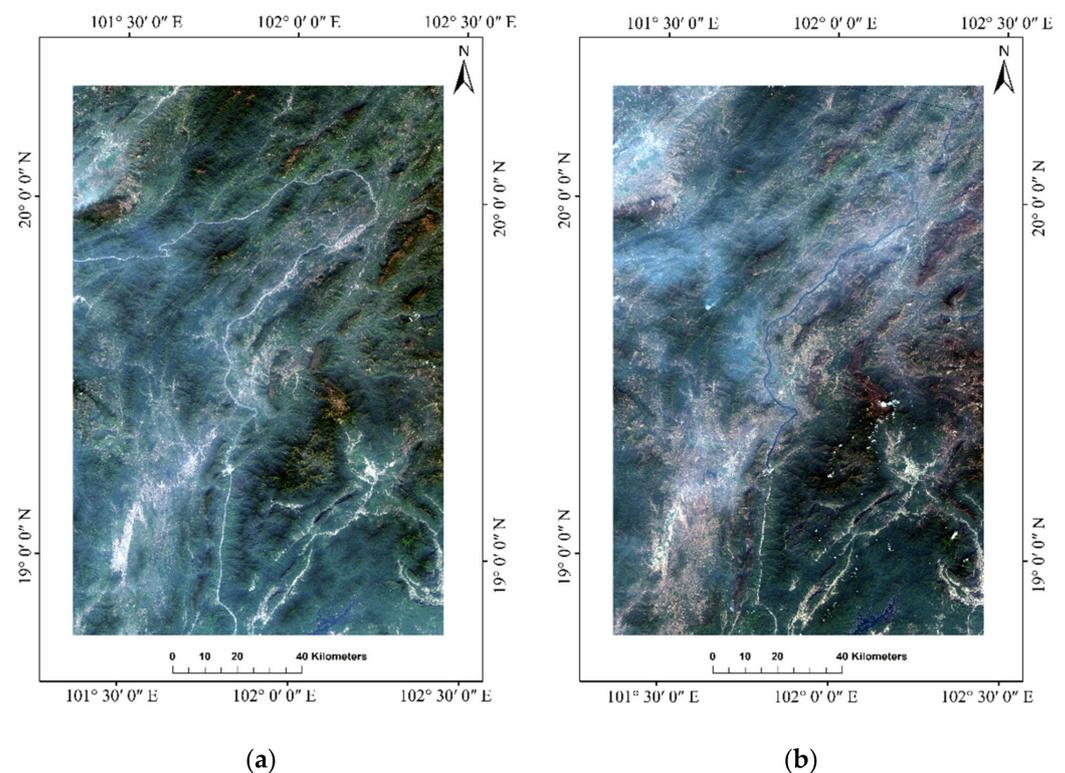


Figure 11. Chinese GF-1 WFV images of the study area: (a) 2016; (b) 2020.

The types of land cover in this area are cultivated land, forests, shrubs, grasslands, water bodies, artificial surfaces, and bare land. From a visual inspection, the main change types are from forests to grasslands, shrubs to farmland, and grasslands to farmland.

4. Results

4.1. Change Detection

As described in Section 2.2.2, the super pixel cosegmentation method was used to detect changes in the study area. As shown in Figure 12, the change detection result map had 61,502 changed patches. The changed pixels account for 20.8% of the total pixels. The large proportion of changes is due to a large number of spurious changes.

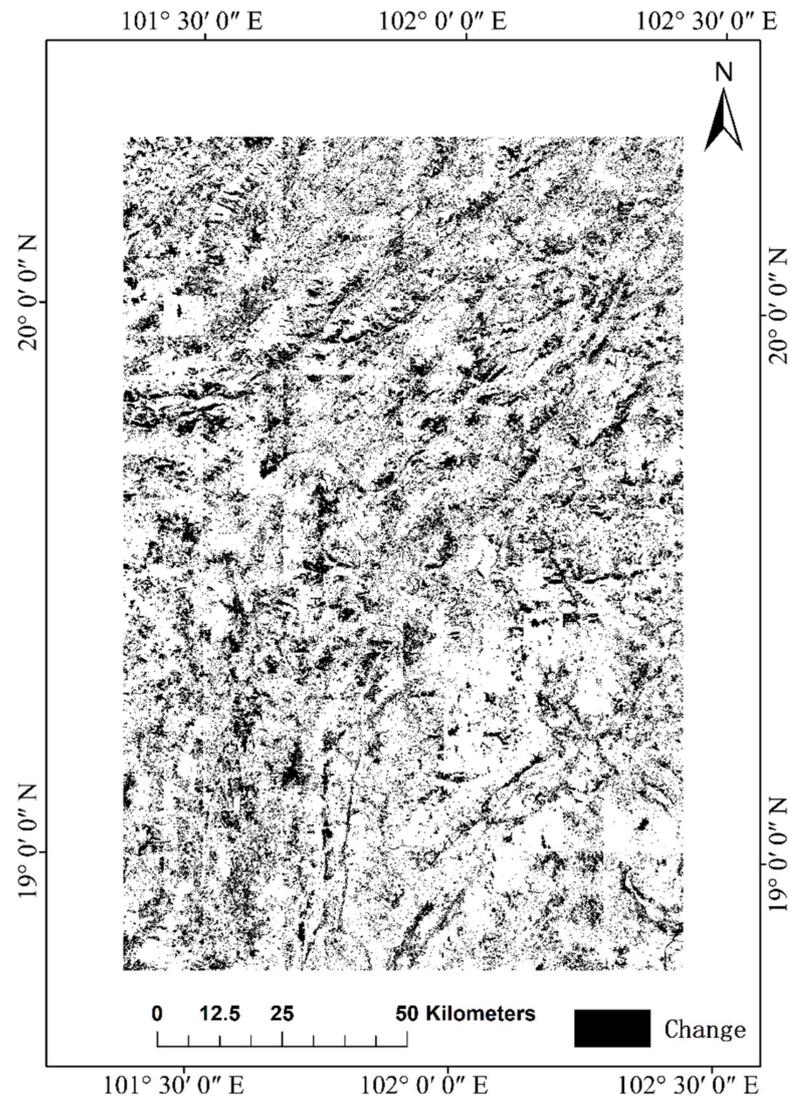


Figure 12. Change detection images of the study area.

4.2. Result of Classification

According to the method in Section 2.2.3, the images of 2016 and 2020 were classified, the method in Section 2.2.4 was used to overlay and analyze the obtained changed patches, and the number of uniform changed patches was 182,334. In 2016 and 2020, the seven types of land cover were croplands, forests, grasslands, shrubs, bare lands, water bodies, and impervious surfaces, as shown in Figure 13. There is no wetland in the study area.

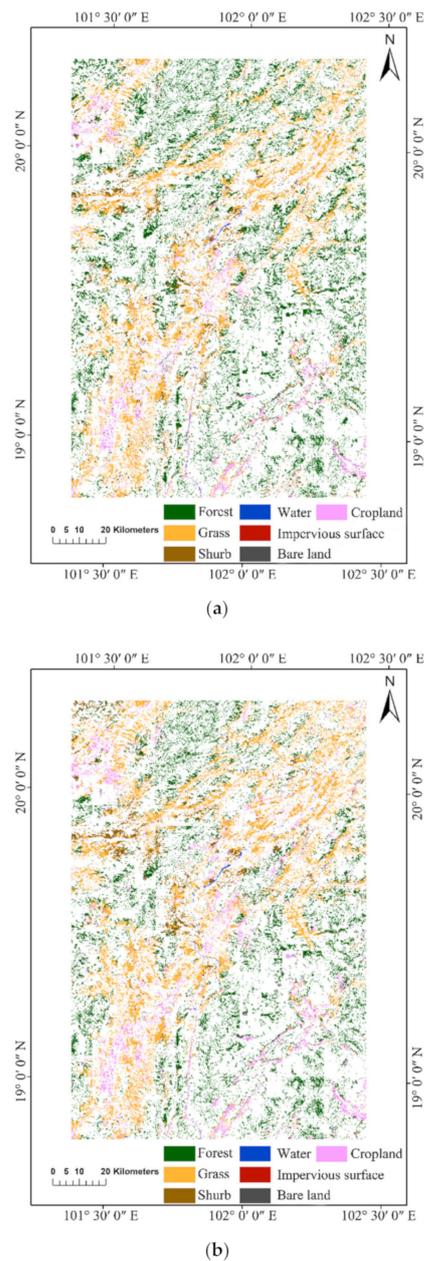


Figure 13. Changed patches classification result images of the study area: (a) 2016; (b) 2020.

4.3. Eco-Geographical Zoning Rule Base Spurious Change Detection Results

According to the method in Section 2.3.2, the offline spurious change detection using the eco-geographical zoning knowledge base was carried out for the classified changed patches. Amongst the 182,334 changed patches, 88,698 possible spurious changed patches were identified, including changed patches with the same type of conversion, prior knowledge collected from remote sensing expert interpreters, and change types with a transformation probability less than 0.0001, as shown in in Table 6 for each eco-geographical zone.

The rules used to identify spurious changes collected from the expert interpreters in these study areas are listed in Table 7. Each rule with a confidence level was accessed by the expert interpreter. These rules were spurious change rules stored in the third level of the eco-geographical zoning knowledge base, which are rules for large IM01 ecological divisions because all of the eco-geographical zones in this study were located in the IM01 eco-region. The higher the confidence level, the more likely it is that a change is a spurious change.

Table 7. Rules with confidence levels over 0.7 for large IM01 ecological divisions.

Ecological Divisions	Rule (A → B)	Confidence Level
IM01	Water–forest	0.8
	Artificial surface–forest	0.9
	Bare land–forest	0.7
	Cropland– forest	0.8

The spurious change detection results by eco-geographical zoning rule base are shown in Table 8. Most of the spurious changes identified are changed patches with the same type of land cover before and after transformation.

Table 8. Statistics of potential spurious changes in land cover.

Eco-Geographical Zone	Source of Rules	Types of Land Cover Transformation	Number of Potential Spurious Changed Patches
IM01	From an interpreter	Water–forest	11
		Artificial surface–forest	4
		Bare land–forest	252
		Cropland–forest	2056
	Changed patches with the same type of land cover before and after transformation	Cropland–cropland	12,511
		Forest–forest	33,609
		Shrub–shrub	23,464
		Grass–grass	8682
		Water–water	707
		Artificial surface–artificial surface	764
IM0121	Statistics from land cover products—low-probability events	Bare land–bare land	211
		Cropland–grass	960
		Forest–grass	3634
		Shrub–grass	1679
		Artificial surface–grass	9
IM0137	Statistics from land cover products—low-probability events	Cropland–bare land	3
		Forest–bare land	81
		Grass–bare land	31
		Shrub–bare land	29
Sum		Water–bare land	1
			88,698

The statistics in Tables 5 and 6 show that the proportion of forests and shrubs was the largest. In each eco-geographical zone, the self-conversion rate of forests and shrubs was also very high, which means that transforming into other land cover types is difficult for forests and shrubs. Based on this, the changed patches obtained by image change detection were affected by various causes of spurious change. Although changes in the images were detected, the possibility of actual changes was small. Hence, forest-to-forest and shrub-to-shrub conversions were considered to be spurious changes. Due to the high self-conversion rate of water bodies, artificial surfaces, and bare land, and the small proportion of these types, the changed patches detected as water body to water body, artificial surface to artificial surface, and bare land to bare land were considered to be spurious changes.

The other two kinds of land cover self-transformation changed patches (i.e., cropland–cropland, grassland–grassland) may be spurious or real changes, which results in significant uncertainty. If they are directly and arbitrarily considered to be spurious changes, they may contain incorrect judgments, which reduces the accuracy of change detection. Therefore, due to their uncertainty, these two kinds of changed patches were the next to be released via network crowdsourcing. Nine kinds of statistics from land cover products with small probability (less than 0.0001) events were also regarded as being uncertain and were released using network crowdsourcing. The total number of changed patches published online was 27,620 (shown as the gray background in Table 8) and accounted for approximately 15% of the total detected changed patches.

The spurious changed patches detected by the eco-geographical zoning knowledge base are shown in Figure 14. A total of 121,256 changed patches (account for 8% of the total pixels) were retained after the spurious changed patches were deleted.

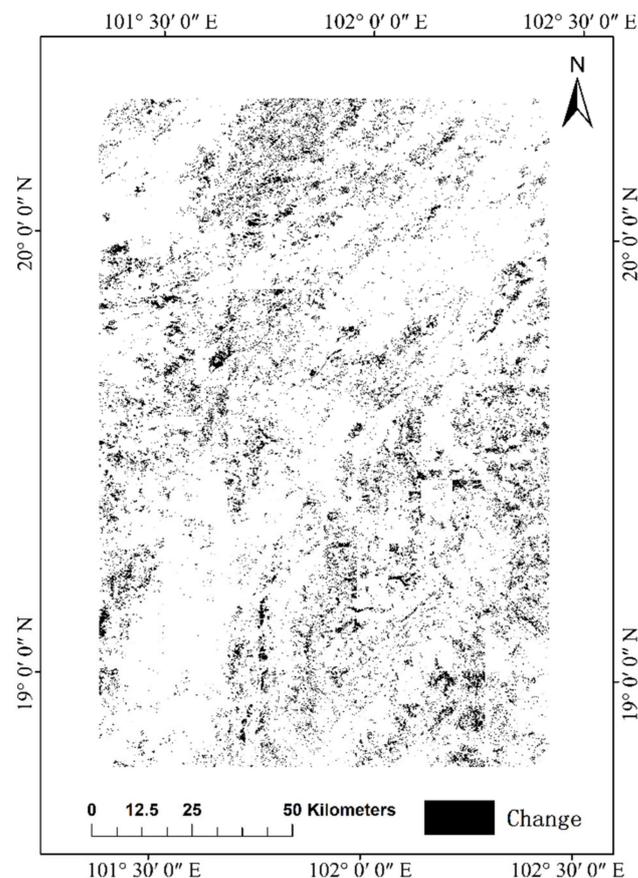


Figure 14. Spurious changes detected by the eco-geographical zoning knowledge base.

Figure 15a–d shows patches that were identified as forests transformed into forests, shrubs transformed into shrubs, cropland transformed into cropland, and grass transformed into grass.

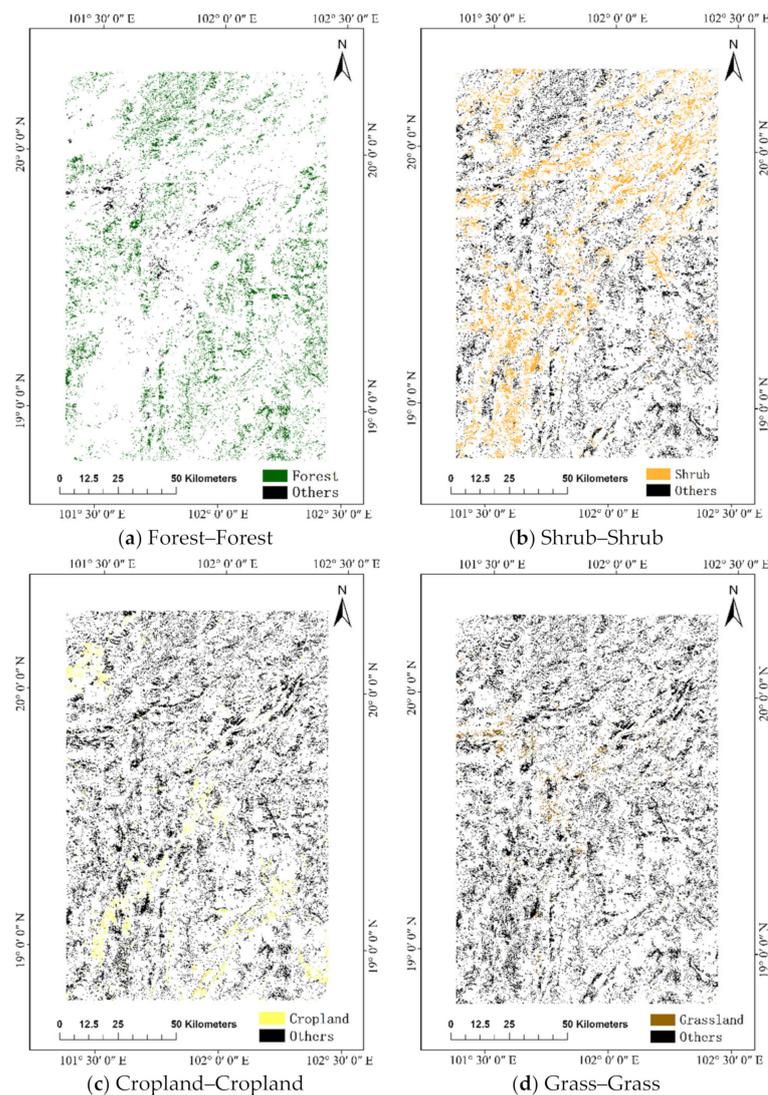


Figure 15. Spurious changes in (a) forests transformed into forests; (b) grass transformed into grass; (c) cropland transformed into cropland; (d) shrubs transformed into shrubs.

4.4. Crowdsourcing Online Spurious Change Tagging and Data Mining

The web site for crowdsourcing spurious change tagging of this study is <http://ggssc.wvu.edu.cn/tomcatproxy/proxy/MapProject/login.html> (accessed on 11 August 2021). After a period of time, we received annotation information on all 27,620 patches from volunteers. The higher the authority value, the higher the spurious change degree of the change patch, and the higher the hub value, the greater the likelihood of the user identifying a spurious change. Table 9 is an example of the results, including the change patch serial number and the authority value. Due to the normalization, the range of the authority value in the result is between 1 and 2.236. Through inspection and analysis, patches with an authority value greater than 1 were defined as spurious changes, and a total of 8608 patches were removed from the results.

4.5. Accuracy Assessment

Confusion matrix analysis was used to verify the accuracy of change detection after removing spurious changes. In the change detection result map, several reference samples of changed and unchanged patches were selected, and a high-resolution image of the Google Earth map was used to visually judge the correctness of the results.

Table 9. An example of online crowdsourcing data computing results.

Change Patch Serial Number	Authority Value
13317	1
14779	1
13643	1
3456	1
9166	1.4142
14713	1
7041	1
14882	1
667	1
7020	1.4108
8077	1
1325	1
3517	1.4142
3877	1
7988	1
204	1
956	1
1	1
13310	1
3706	1.4
3436	1.022
12540	1

Based on the cosegmentation change detection results, the changed and unchanged patches were converted into point files using the conversion tools in ArcGIS, and then changed and unchanged samples were randomly selected. The distribution of the verification samples is shown in Figure 16.

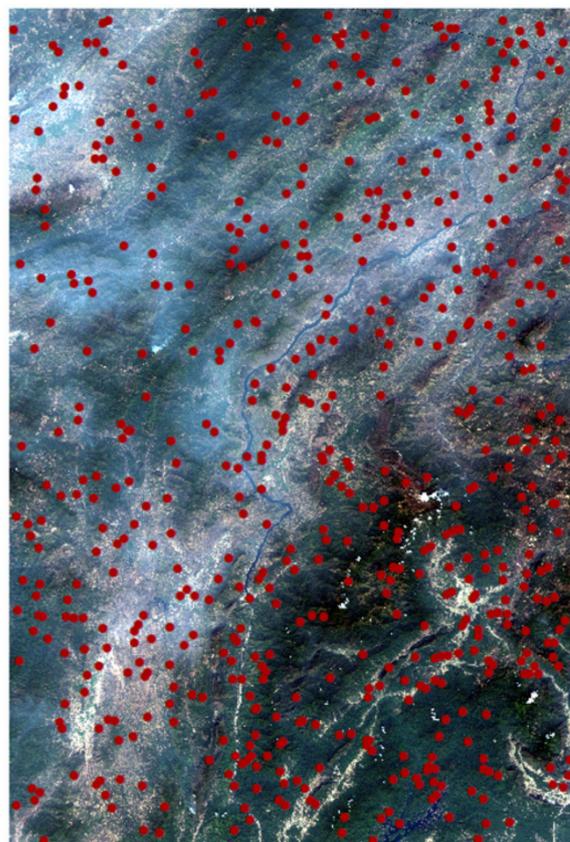
**Figure 16.** Verification samples.

Table 10 shows the confusion matrix of the original changes, whereas Table 11 shows the confusion matrix after removing the spurious change image patches in the experimental area. The results in Table 10 show that, amongst the 586 samples, 278 were unchanged and 308 were changed samples. The overall accuracy was only 66.72%, indicating that the phenomenon of excessive change detection occurred. Table 11 shows that after eliminating the spurious changes, the number of changed samples was reduced to 140. The overall accuracy was improved by 23.89%, reaching 90.61%.

Table 10. Super pixel cosegmentation change detection confusion matrix.

$T_2 \setminus T_1$	Changed	Unchanged	Sum
Changed	113	195	308
Unchanged	0	278	278
Sum	113	473	586
Overall Accuracy	66.72%		
KIA	0.3548		

Table 11. Confusion matrix after deleting spurious changes.

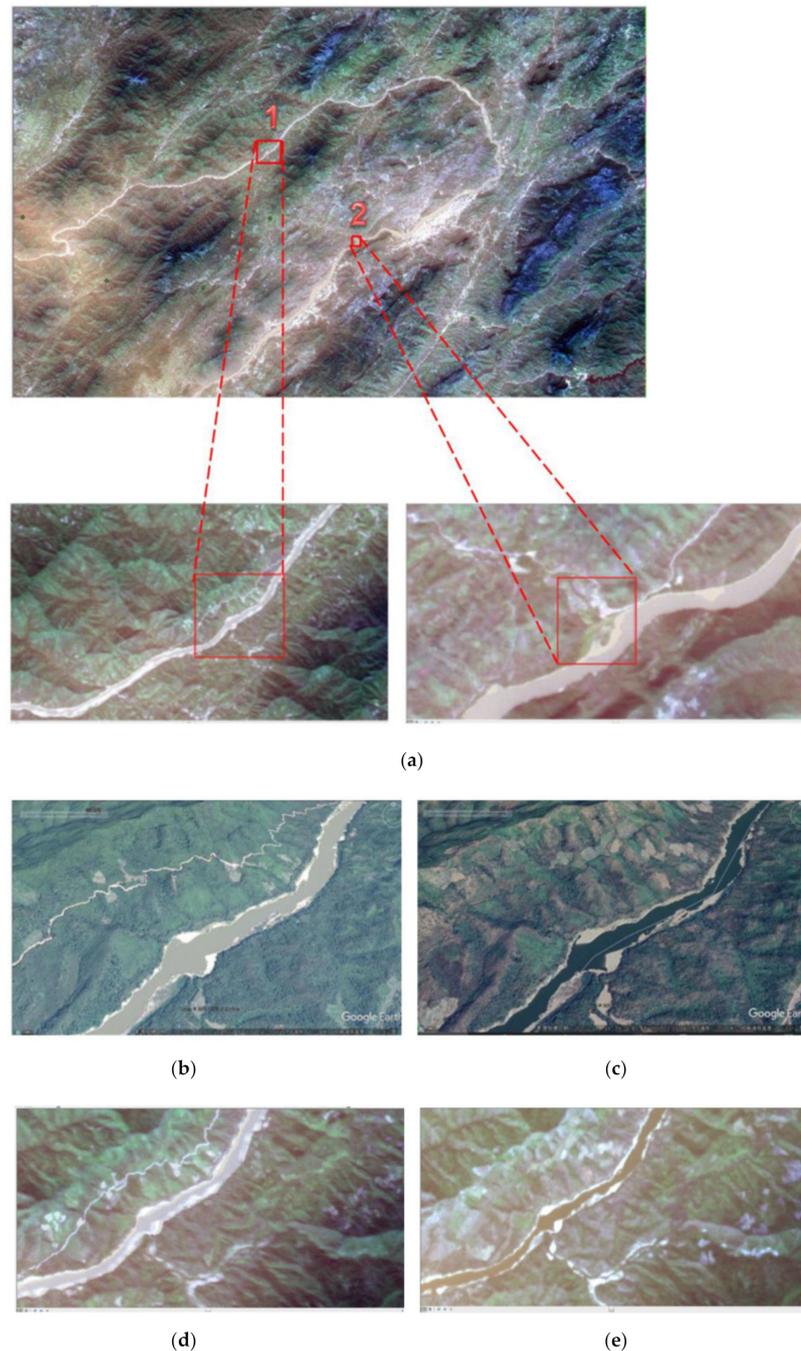
$T_2 \setminus T_1$	Changed	Unchanged	Sum
Changed	99	41	140
Unchanged	14	432	446
Sum	113	473	586
Overall Accuracy	90.61%		
KIA	0.7236		

5. Discussions

An important issue that must be considered in remote sensing image change detection is how change should be defined. This issue is very important for the design and verification of algorithms and for the accuracy of change detection. In a study by Benedek and Szirányi [40], the following changes were considered: new built-up regions, building operations, planted forests or individual trees, freshly ploughed lands, and the ground-works of future buildings. “Change” was defined as the change between land cover types because the purpose of change detection in this study is the subsequent updating of land cover products. A change in an image did not directly represent a change in the land cover type. For example, changes in river turbidity, spectral differences in cultivated land before and after harvest, and changes in shrub, grassland or forest density did cause spectral differences, but the type of land cover did not change. The purpose of this study is to remove these kinds of spurious changes using an eco-geographical zoning knowledge base and online crowdsourcing annotation method. Two examples were provided. Figure 17 shows two sites in the study area: Figure 17b,c presents Google Earth images of Area 1 in February 2016 and February 2020. Figure 17d,e presents GF-1 images of Area 1 in February 2016 and February 2020. Figure 17f shows the changed patches in Area 1 before the removal of spurious changes. Figure 17g shows the changed patches after the removal of spurious changes. A comparison of the two images reveals that, due to phenological reasons, several cultivated land plots on both sides of the river were green in the image from 2016, but became the color of bare land in the image from 2020. These spurious changes were identified and removed. Another example showed a similar phenomenon. Figure 17h,i presents Google Earth images of Area 2 in February 2016 and February 2020; Figure 17j,k presents GF-1 images of Area 2 in February 2016 and February 2020. Figure 17l shows the changed patches in Area 2 before the removal of spurious changes. Figure 17m shows the changed patches after the removal of spurious changes. A comparison of the two images reveals that, due to phenological reasons, several cultivated land plots on the bank of the river were relatively bare in the image from 2016 but became green in the image from 2020. Moreover, the forest land was relatively dense in 2016, but sparse in 2020. A

large hue difference in the GF-1 images was observed, and these spurious changes were therefore recognized.

From Figure 11, it can be seen that the GF-1 WFV images of the study area contain the impact of clouds, including thin and thick clouds. This will inevitably result in spurious changes in change detection if directly comparing the two-phase images. In the cosegmentation change detection method, SGD was used to obtain the change intensity image. SGD calculates the spectral differences in the adjacent spectral bands of multispectral remote sensing images, and then compares the two-phase images to obtain the changed pixels. As a result, spurious changes caused by thin clouds can effectively be avoided. In addition, some spurious changes caused by the existence of the remaining clouds, especially thick clouds, can be removed by the annotations of volunteers through crowdsourced data mining.



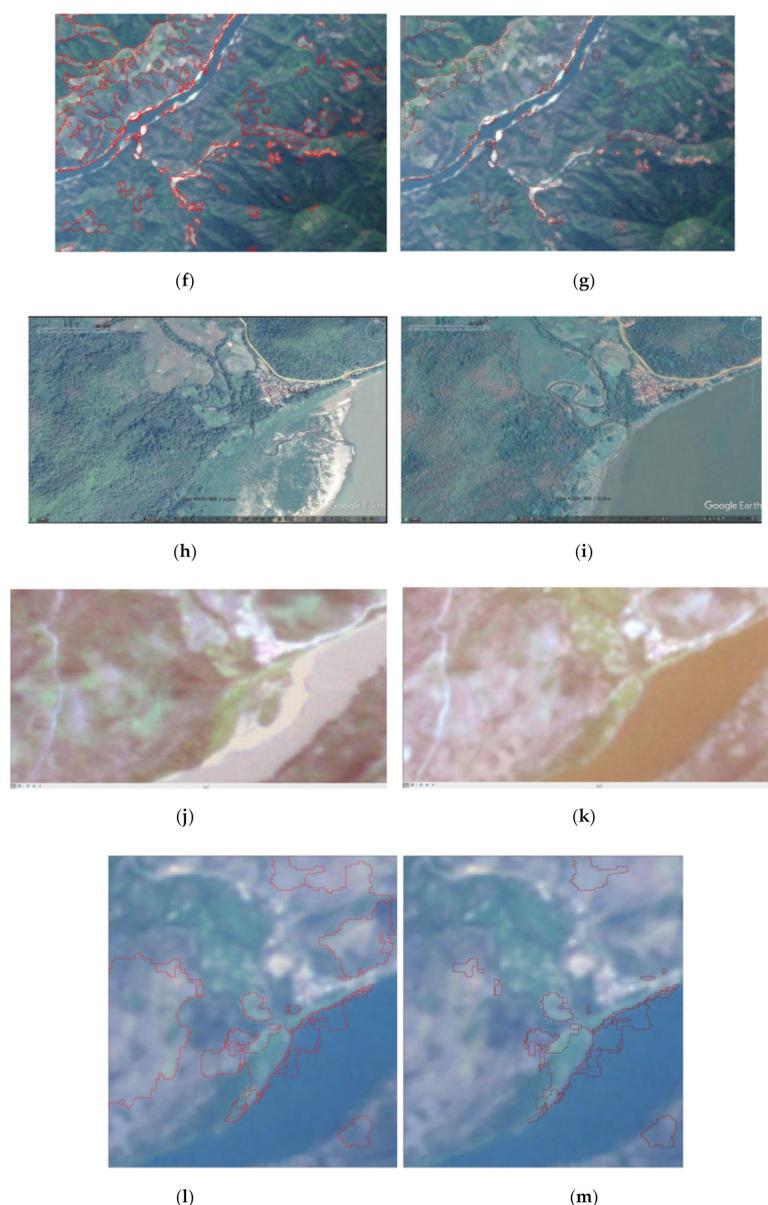


Figure 17. (a) Location of two small areas, (b) Area 1 Google Earth image of February 2016, (c) Area 1 Google Earth image of February 2020, (d) Area 1 GF-1 image of February 2016, (e) Area 1 GF-1 image of February 2020, (f) Area 1 changed patches before spurious change removal, (g) Area 1 changed patches after spurious change removal, (h) Area 2 Google Earth image of February 2016, (i) Area 2 Google Earth image of February 2020, (j) Area 2 GF-1 image of February 2016, (k) Area 2 GF-1 image of February 2020, (l) Area 2 changed patches before spurious change removal, (m) Area 2 changed patches after spurious change removal.

6. Conclusions

This study proposes a method for identifying spurious changes using the knowledge base rules of an eco-geographical zoning knowledge base and crowdsourced data mining. The experimental results show this method is feasible and effective for improving the accuracy of change detection and has certain practical relevance for updating land cover.

The identification of spurious changes using an eco-geographical zoning knowledge base depends on the rules collected. Two methods are used in this study, namely knowledge from expert interpreters, which is subjective, and knowledge from land cover product statistics. At present, the expression of the rules is based on the degree of confidence. In

the future, artificial intelligence can also be used to mine rules from data to improve the completeness of the search and the accuracy of the method.

The platform based on crowdsourced data can use the experience and knowledge voluntarily shared by network users to evaluate the spurious change degrees of changed patches. Based on the weighted HITS algorithm for spurious change detection, combined with the WebGIS framework, PostgreSQL database, SPARK data parallel processing, and other technologies, a platform for spurious change detection based on crowdsourced geographic data can be realized.

Change detection in Laos is challenging because the study area has complex land cover characteristics, frequent changes in land cover, and fine patches. The examples provided in this study show that this method still has several shortcomings. The registration errors of the two images are large due to the local geometric deformation of GF-1 satellite images. Approximately two pixel registration errors in many positions were observed, which can lead to many spurious changes. Several of these can be eliminated using the proposed method, but many such errors remain. Analysis-ready data (ARD) from GF-1 images need to be produced to improve the accuracy of change detection.

The rules for the removal of spurious changes in the knowledge base of eco-geographical zoning are based on the correct classification of changed patches. However, the accuracy of automatic classification methods is limited. Classification errors definitely affect the identification of spurious changes, as several real changes are removed, and false changes cannot be correctly identified. Technical routes also need to be improved to limit the effect of classification errors on the methods.

The method of crowdsourced data mining relies on the participation of a sufficient number of volunteers. The number of volunteers in emerging countries, such as Laos, is limited. As a result, crowdsourcing only plays a secondary role in the overall technical approach (61078 vs. 8608 spurious changed patches were removed by the eco-geographical zoning knowledge base and crowdsourced data mining, respectively).

Author Contributions: Conceptualization, funding acquisition, methodology, and writing—original draft, L.Z.; data curation, methodology and software, D.G.; crowdsourced data mining method design T.J.; data processing J.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by national key research and development program of China (grand number 2016YFB0501404), and Beijing key laboratory of urban spatial information engineering (grand number 20210217).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Zhu, L.; Jia, T.; Shi, R. *Global Surface Covering Product Update and Integration*; Science Press: Beijing, China, 2020. (In Chinese)
2. Chen, J.; Chen, J.; Liao, A. *Remote Sensing Mapping of Global Land Cover*; Science Press: Beijing, China, 2016. (In Chinese)
3. Zhu, L.; Sun, Y.; Shi, R.M.; La, Y.X.; Peng, S. Exploiting Cosegmentation and Geo-Eco Zoning for Land Cover Product Updating. *Photogramm. Eng. Remote Sens.* **2019**, *85*, 597–611. [[CrossRef](#)]
4. Zhang, L.; Wu, C. Current situation and Prospect of multi temporal remote sensing image change detection. *J. Surv. Mapp.* **2017**, *46*, 1447–1459.
5. Chen, J.; Lu, M.; Chen, X.; Chen, J.; Chen, L. A spectral gradient difference based approach for land cover change detection. *ISPRS J. Photogramm. Remote Sens.* **2013**, *85*, 1–12. [[CrossRef](#)]
6. Hu, Y.; Dong, Y. An automatic approach for land-change detection and land updates based on integrated NDVI timing analysis and the CVAPS method with GEE support. *ISPRS J. Photogramm. Remote Sens.* **2018**, *146*, 347–359. [[CrossRef](#)]
7. Li, W.; Zhang, C.; Willig, M.R.; Dey, D.K.; Wang, G.; You, L. Bayesian Markov Chain Random Field Cosimulation for Improving Land Cover Classification Accuracy. *Math. Geosci.* **2015**, *47*, 123–148. [[CrossRef](#)]
8. Giri, C.P. *Remote Sensing of Land Use and Land Cover: Principles and Applications*; CRC Press: Boca Raton, FL, USA, 2012.

9. Mora, B.; Tsendbazar, N.E.; Herold, M.; Arino, O. *Global Land Cover Mapping: Current Status and Future Trends*; Springer Press: Dordrecht, The Netherlands, 2014.
10. Chen, J.; Chen, J.; Liao, A.; Cao, X.; Chen, L.; Chen, X.; He, C.; Han, G.; Peng, S.; Lu, M.; et al. Global land cover mapping at 30 m resolution: A POK-based operational approach. *ISPRS J. Photogramm. Remote Sens.* **2015**, *103*, 7–27. [[CrossRef](#)]
11. Liu, T.; Chen, X.; Dong, Q.; Cao, X.; Chen, J. Application of deep learning in product classification accuracy optimization of globeland30-2010. *Remote Sens. Technol. Appl.* **2019**, *34*, 3–11. (In Chinese)
12. You, Y.; Cao, J.; Zhou, W. A Survey of Change Detection Methods Based on Remote Sensing Images for Multi-Source and Multi-Objective Scenarios. *Remote Sens.* **2020**, *12*, 2460. [[CrossRef](#)]
13. Sood, V.; Gusain, H.S.; Gupta, S.; Singh, S. Topographically derived subpixel-based change detection for monitoring changes over rugged terrain Himalayas using AWiFS data. *J. Mt. Sci.* **2021**, *18*, 126–140. [[CrossRef](#)]
14. Zhou, J.; Yu, B.; Qin, J. Multi-Level Spatial Analysis for Change Detection of Urban Vegetation at Individual Tree Scale. *Remote Sens.* **2014**, *6*, 9086–9103. [[CrossRef](#)]
15. Lu, M.; Chen, J.; Tang, H.; Rao, Y.; Yang, P.; Wu, W. Land cover change detection by integrating object-based data blending model of Landsat and MODIS. *Remote Sens. Environ.* **2016**, *184*, 374–386. [[CrossRef](#)]
16. Wang, L.; Luo, Y.; Zhang, X.; Ma, Q.; Kari, H.; Jia, L. Research progress on Application of remote sensing technology in forest pest monitoring. *World For. Res.* **2008**, *21*, 37–43.
17. Yan, D. Extraction of Inland Surface Water with Different Turbidity Based on Landsat Series Remote Sensing Images. Master's Thesis, Northwestern University, Xi'an, China, 2004. (In Chinese).
18. Lv, Z.Y.; Liu, T.F.; Zhang, P.; Benediktsson, J.A.; Lei, T.; Zhang, X. Novel Adaptive Histogram Trend Similarity Approach for Land Cover Change Detection by Using Bitemporal Very-High-Resolution Remote Sensing Images. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 9554–9574. [[CrossRef](#)]
19. Liu, B.; Chen, J.; Chen, J.; Zhang, W. Land Cover Change Detection Using Multiple Shape Parameters of Spectral and NDVI Curves. *Remote Sens.* **2018**, *10*, 1251. [[CrossRef](#)]
20. European Space Agency. *CCI-LC Product User Guide*; UCL-Geomatics: Louvain-la-Neuve, Belgium, 2017.
21. Herbertson, A.J. The Major Natural Regions: An Essay in Systematic Geography. *Geogr. J.* **1905**, *25*, 300–310. [[CrossRef](#)]
22. Carver, S.; Evans, A.; Kingston, R.; Turton, I. Public participation, GIS, and cyberdemocracy: Evaluating on-line spatial decision support systems. *Environ. Plan. B Plan. Des.* **2001**, *28*, 907–921. [[CrossRef](#)]
23. Hudson-Smith, A.; Batty, M.; Crooks, A.; Milton, R. Mapping for the Masses: Accessing web 2.0 through crowdsourcing. *Soc. Sci. Comput. Rev.* **2009**, *27*, 524–538. [[CrossRef](#)]
24. Li, D.; Qian, X. On data management of spontaneous geographic information. *J. Wuhan Univ.* **2010**, *35*, 379–383. (In Chinese)
25. Goodchild, M.F.; Glennon, J.A. Crowdsourcing geographic information for disaster response: A research frontier. *Int. J. Digit. Earth* **2010**, *3*, 231–241. [[CrossRef](#)]
26. Jia, T.; Yu, X.; Shi, W.; Liu, X.; Li, X.; Xu, Y. Detecting the regional delineation from a network of social media user interactions with spatial constraint: A case study of Shenzhen, China. *Phys. A Stat. Mech. Appl.* **2019**, *531*, 121719. [[CrossRef](#)]
27. Haklay, M. *Citizen Science and Volunteered Geographic Information: Overview and Typology of Participation*; Springer: Boston, MA, USA, 2013. [[CrossRef](#)]
28. Li, D. Looking forward to geospatial informatics in the era of big data. *J. Surv. Mapp.* **2016**, *45*, 379–384.
29. Clark, M.L.; Mitchell, A.T. Virtual Interpretation of Earth Web-Interface Tool (VIEW-IT) for Collecting Land-Use/Land-Cover Reference Data. *Remote Sens.* **2011**, *3*, 601–620. [[CrossRef](#)]
30. See, L.; Dmitry, S.; Myroslava, L.; Ian, M.; Steffen, F.; Alexis, C.; Christoph, P.; Christian, S.; Zhao, Y.Y.; Victor, M. Building a hybrid land cover map with crowdsourcing and geographically weighted regression—ScienceDirect. *ISPRS J. Photogramm. Remote Sens.* **2015**, *103*, 48–56. [[CrossRef](#)]
31. Schepaschenko, D.; See, L.; Lesiv, M.; McCallum, I.; Fritz, S.; Salk, C.; Moltchanova, E.; Perger, C.; Shchepashchenko, M.; Shvidenko, A.; et al. Development of a global hybrid forest mask through the synergy of remote sensing, crowdsourcing and FAO statistics. *Remote Sens. Environ.* **2015**, *162*, 208–220. [[CrossRef](#)]
32. Xing, H.; Meng, Y.; Hou, D.; Xu, H.; Liu, J. A Land-Cover Classification Method Using Point of Interest. *Geomat. Inf. Sci. Wuhan Univ.* **2019**, *44*, 758–764.
33. Fan, W.; Wu, C.; Jin, W. Improving Impervious Surface Estimation by Using Remote Sensed Imagery Combined With Open Street Map Points-of-Interest (POI) Data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2019**, *12*, 4265–4274. [[CrossRef](#)]
34. Shan, X.J.; Tang, P.; Hu, C.M.; Tang, L.; Zheng, K. Automatic geometric precise correction technology and system based on hierarchical image matching for HJ-1A/B CCD images. *J. Remote Sens.* **2014**, *18*, 254–266.
35. Rother, C.; Minka, T.; Blake, A.; Kolmogorov, V. Cosegmentation of Image Pairs by Histogram Matching—Incorporating a Global Constraint into MRFs. In Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), New York, NY, USA, 17–22 June 2006.
36. Zhu, L.; Zhang, J.; Sun, Y. Remote Sensing Image Change Detection Using Superpixel Cosegmentation. *Information* **2021**, *12*, 94. [[CrossRef](#)]
37. Olson, D.M.; Dinerstein, E.; Wikramanayake, E.D.; Burgess, N.D.; Powell, G.V.; Underwood, E.C.; D'amico, J.A.; Itoua, I.; Strand, H.E.; Morrison, J.C.; et al. Terrestrial Ecoregions of the World: A New Map of Life on Earth. *BioScience* **2001**, *51*, 933–938. [[CrossRef](#)]

-
38. Jon, M.K. Authoritative sources in a hyperlinked environment. *J. ACM* **1999**, *46*, 604–632.
 39. 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](#)]
 40. Benedek, C.; Szirányi, T. Change Detection in Optical Aerial Images by a Multi-Layer Conditional Mixed Markov Model. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 3416–3430. [[CrossRef](#)]