

Removing Land cover spurious change by Geo-Eco zoning rule base

Ling Zhu^{1,*}, Yixuan La¹, Ruoming Shi¹, Shu Peng²

¹ Beijing University of civil engineering and architecture, No.1 Zhanlanguan Rd, Xicheng district, Beijing -zhuling@bucea.edu.cn

² National Geomatics center of China, No.28 Lianhuachi West road, Beijing-pengshu@ngcc.cn

KEY WORDS: Land cover product, Geo-eco zoning, Spurious change detection, Rule base.

ABSTRACT:

The speed of change in land cover is growing faster with the development of social science and technology. Remote sensing has become the most effective way to monitor change information. However, remote sensing images reflect only the instantaneous state of the Earth's surface. Spectral characteristics cannot correctly reflect the actual state, and this inability results in the limited classification accuracy of land cover products. In order to obtain high accuracy change detection results, it is necessary to identify and eliminate spurious changes.

At present, the spurious changes are generally identified by visual interpretation which not only labor and time consuming, but also easily lead to misjudgment due to the lack of identification experience of the interpreter. Therefore, it is urgent to establish a spurious change rule base to automatically identify spurious changes. In this study, the global geo-eco zoning can be used to build a rule base to identify and eliminate spurious changes.

The structure and content of the rule base are designed, the rules are represented and put into the rule library, the plugins are designed to remove spurious changes, and a rule base management system is established to identify the spurious changes using the rules in the rule base. 30m Land cover products of Laos were selected as the experimental area to verify the accuracy of the change patches after eliminating spurious changes. Results show that the accuracy of change detection is improved by using the rule base of geo-eco zoning to identify spurious changes.

1. INTRODUCTION

1.1 Spurious change identification status

Remote sensing change detection is a process to determine the surface change using multi-temporal remote sensing image data. The classification and change detection of land cover are mostly based on the analysis of spectral characteristics of remote sensing images. Because the remote sensing image only reflects the instantaneous characteristics of the surface, the phenomena of "same object different spectrums" and "different objects same spectrum", as well as the differences caused by seasonal phase changes, the spectral features cannot correctly reflect the actual changes and will produce spurious changes. Therefore, it is very necessary to identify and eliminate spurious changes for obtaining high precision change detection results.

At present, the identification of spurious changes is mostly manual identification, which wastes time, wastes resources and is easy to cause wrong judgment due to personnel problems. Although many land cover classification products are generated in the process of taking knowledge as auxiliary data to improve accuracy. For example, the GlobeLand30 product of the National Geomatics center of China uses auxiliary data in addition to land cover data, as well as special data such as lakes and wetlands, MODIS NDVI data, and global DEM data(Chen,2017). However, the relevant auxiliary data and available knowledge are relatively fragmented, and it is urgent to establish a knowledge base to systematically store auxiliary data. Eco-geographical zoning(Herbertson,1905) can be used to construct knowledge base to assist change detection because of its global nature, stability of land features within the partition, regularity of landform changes and large amount of information(Bailey,1983).

1.2 Geo-eco zoning

The "ecological zone" was first proposed by Canadian forest research scholar Orié Loucks in 1962(Bailey,1986). Ecological zoning is defined as a relatively large land or water unit that contains a unique combination of natural communities that represents the original distribution of different combinations of species and communities(Bailey,1989).

At present, the most famous ecological zonings in the world are as shown in Table1(Bailey,2004):

The most famous ecological zonings in the world
Holdridge model(Holdridge,1967)
World biogeographical biomes
world ecosystems
Continental ecoregions
Ecozones of the world
Terrestrial ecoregions of the world(Olson,2001)

Table 1. The most famous ecological zonings in the world

This paper uses the Terrestrial ecoregions of the world established by the World Wide Web Fund (WWF) as the basic framework. The world is divided into 8 biogeographic realms and 14 biomes. Based on these two basic layers, there are 867 ecological zones in the world. Each ecological zone has a unique 6-digit number. Geographical divisions and biomes are shown in Figure 1(a). The vector data of the global ecological partition SHP format is shown in Figure 1(b). The names and codes of the 8 biogeographical realms and the names and codes of the 14 biomes are shown in Tables 2 and 3.

* Corresponding author

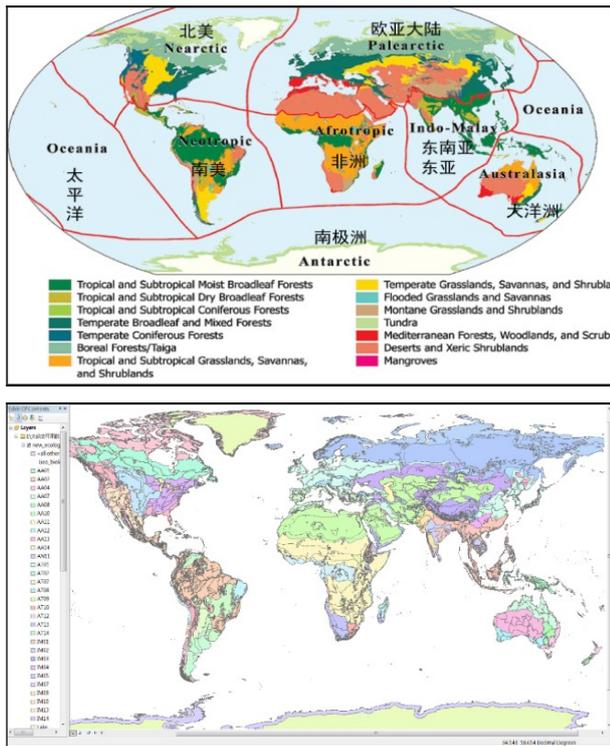


Fig1(a).8 Biogeographic realm and 14 Biomes

Fig1(b).Shape file of Eco-regions

Code	PA	NT	NA	AT
Name	Paleartic	Netotri c	Nearctic	Afrotropic
Code	IM	OC	AA	AN
Name	Indo- Malay	Oceania	Australasia	Antarctic

Table 2 Names and codes of 8 Biogeographic realm

Code	Name
01	Tropical and Subtropical Moist Broadleaf Forests
02	Tropical and Subtropical Dry Broadleaf Forests
03	Tropical and Subtropical Coniferous Forests
04	Temperate Broadleaf and Mixed Forests
05	Temperate Coniferous Forests
06	Boreal Forests/Taiga
07	Tropical and Subtropical Grasslands, Savannas, and Shrublands
08	Temperate Grasslands, Savannas, and Shrublands
09	Flooded Grasslands and Savannas
10	Montane Grasslands and Shrublands
11	Tundra
12	Mediterranean Forests, Woodlands, and Scrub
13	Deserts and Xeric Shrublands
14	Mangroves

Table 3 Names and codes of 14 Biomes

There are similar biomes in each ecological zone. The land cover type is relatively stable in a certain period of time, even if the surface changes, this change also has a certain regularity, and this trend can be used as a reference. At the same time, we

can mine knowledge related to land cover and change detection based on different geographic attributes. In summary, we choose ecological zone as the basic data to build a knowledge base, further sorting out and excavating this knowledge, and designing a reasonable knowledge base structure combined with expert experience knowledge and summarizing the spurious change rules of each partition to identify and eliminate spurious change.

2. METHOD

2.1 Geo-eco zoning rule base design

Spurious changes in geo-eco zoning rules require the consideration of a large amount of data, given the 867 ecological zones. Summing up the spurious change rule for each zone is a crucial but time-consuming task. Therefore, this study designed a unique object-oriented approach to establish the geo-eco zoning rule database. The advantage of object-oriented knowledge representation is that it is a structured knowledge representation technology and has inheritance and derivation characteristics. That is, subclasses can inherit all of the characteristics of their parents. Thus, this approach significantly reduces the workload and solves problems rapidly and efficiently.

The four-layer rule database establishes the derivation and inheritance relationship among the different layers from top to bottom, as depicted in Fig.2. On the basis of the 8 eco-geographical zones and 14 biological communities, a number of factors are considered attribute factors, and these include temperature, humidity, normalized difference vegetation index (NDVI), elevation, and slope.

The first layer of the rule base has two parallel branches, namely, left and right branches, which are designed for the ecological geography class and natural attributes, respectively. The two parallel branches comprise the top layer, and all other objects are a subclass of the two parallel branches.

The left branch of the second layer includes 8 eco-geographical zones and 14 biological communities. This branch is the subclass of the ecological geography class of the first layer. The right branch, which includes temperature, humidity, NDVI, elevation, and slope, is the subclass of natural attributes.

In the third layer, the left branch includes large geographical and ecological zones, namely, cross-inheritance of geographical areas and biological communities with 64 zones, which are labeled IM01–PA01. In accordance with the different natural conditions of each ecological zone, the right branch of the third layer is a subclass of elevation (E), slope (S), NDVI value (N), temperature (T), and humidity (W) depending on the values of these factors, as shown in Fig. 2. Given that the ecological zone is relatively large and

contains multiple regions, using atmospheric temperature to express the temperature is difficult. Therefore, the temperature is divided into tropical (T1), subtropical (T2), temperate (T3), and cold (T4) zones. The temperature zone is divided according to latitude as follows: 0°–23.5° is tropical, 23.5°–40° is subtropical, 40°–66.5° is temperate, and 66.5°–90° is cold. The average attribute value of the geo-eco zone is considered because each large geographic area corresponds to a large area of the Earth. The rule base of this layer stores its corresponding spurious change rules.

The fourth level is the lowest and composed of 867 geo-eco partitions that store specific attribute information, and the spurious change rule is determined from this layer only.

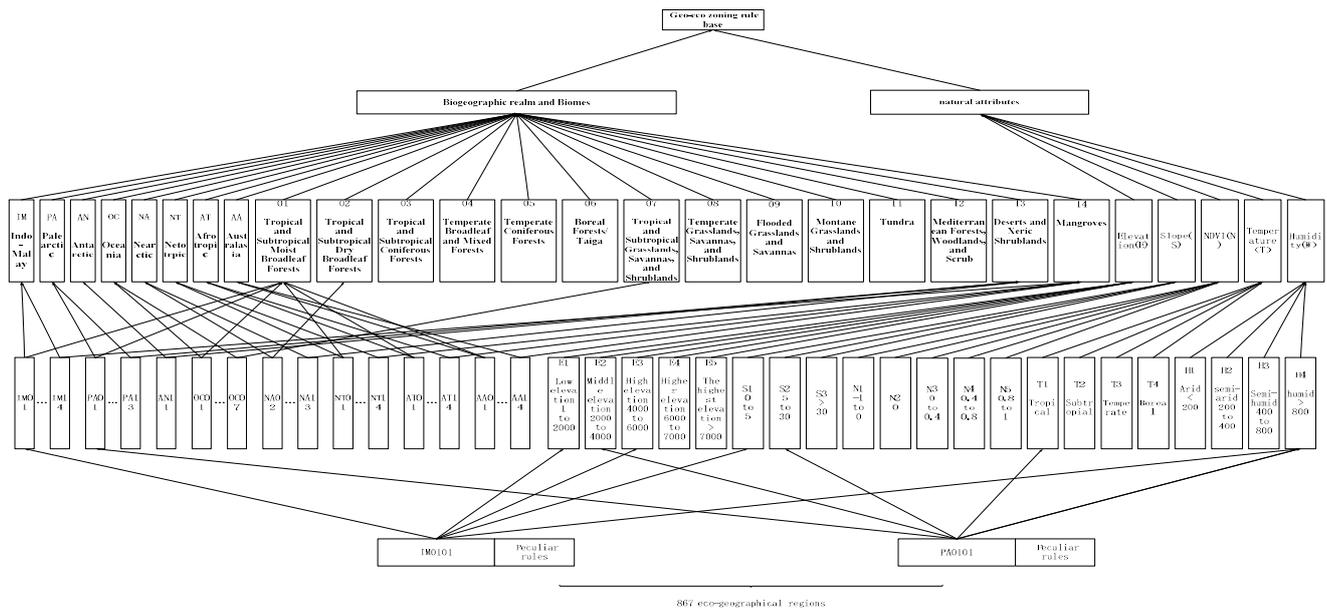


Fig. 2 Geo-eco zoning rule database

2.2 Establishment and expression of spurious change rules

The rules are summarized for the five attributes of elevation, slope, NDVI value, temperature and moisture, and some rules are given.

(1) Elevation: In general, there are few trees at altitudes above 4000 meters; cultivated land rarely exists above 4700; grassland rarely exists above 6000 meters; tundra is rarely present at altitudes above 7000 meters.

(2) Slope: The distribution of cultivated land generally does not exceed 30 degrees, the distribution of water body generally does not exceed 5 degrees, and the distribution of wetlands generally does not exceed 3 degrees. In developed countries, the slope of cultivated land is generally less than 25°. In developing countries, especially in countries with large populations, the slope of cultivated land can reach 30° or even higher.

(3) NDVI: When $-1 < NDVI < 0$, there is no possibility of arable land, woodland, grassland, shrub, wetland, tundra, bare land and artificial cover on the surface; when $NDVI = 0$, there is no arable land or forest land on the surface. Grassland, shrubs, wetlands, water bodies, tundra, man-made cover and snow and ice; when $NDVI > 0$, it means vegetation coverage. At this time, as the density of vegetation increases, the NDVI value increases, and there is no bare land and ice on the surface.

(4) Temperature: Due to the large ecological division area, the temperature range is too small, and the division of each ecological division is difficult. Therefore, the temperature is expressed as temperature zone, which is divided into tropical, subtropical, temperate and cold zones. According to the prior knowledge under different temperature zones, it can be concluded that there is no possibility of ice, snow and tundra in the tropics. In the frigid zone, only the tundra is the tundra, and there are no tall woody plants such as woodland.

(5) Moisture: The moisture attribute is divided into wet area and arid area according to the different standard annual precipitation. Ice and snow and tundra are generally not available in arid regions, and wetlands in arid and semi-arid regions are deteriorating as global warming and human causes occur.

A spurious change rule is expressed in the form of production, that is, the forms of premise and conclusion, and mapped to the

table in the object-relational database. A simple spurious change expression is designed. Spurious change rule representation uses a six-digit code, namely, "XXXXXX", where the first three digits represent the land cover class of the early phase image and the last three digits represent the class code of the new image. The first class code of GlobeLand30 land cover product was used in this study. The specific spurious change rules must be determined based on the expert knowledge of remote sensing interpreters and geographical specialists. Knowledge of the test area at different levels was collected and stored in the database. Several examples are presented in Table4.

Attribute	Prior knowledge	Spurious change rule
Elevation	When elevation >4000 m, there is no cultivated land and forest on the surface of the Earth.	Cultivated land to other land types Other land types to cultivated land Forest to other land types Other land types to forest
	When elevation >6000 m, there is no grassland on the surface of the Earth.	Grassland to other land types Other land types to grassland
	When elevation >7000 m, tundra cannot exist on the surface of the Earth.	Tundra to other land types Other land types to tundra
Slope	When slope >5°, there is no water and wetlands on the surface of the Earth	Water to other land types Other land types to water Wetlands to other land types Other land types to wetlands
	When slope >30°, there is no cultivated	Cultivated land to other land types

	land on the surface of the Earth	Other land types to cultivated land
NDVI	When $-1 < NDVI < 0$, there is no cultivated land, forest, grassland, shrub, wetland, tundra, bare land and artificial cover on the surface of the earth	cultivated land, forest, grassland, shrub, wetland, tundra, bare land and artificial cover to other land types and other land types to them
	When $NDVI = 0$, there is no cultivated land, forest, grassland, shrub, wetland, water, tundra, bare land, artificial cover and ice and snow on the surface of the earth	cultivated land, forest, grassland, shrub, wetland, water, tundra, bare land, artificial cover and ice and snow to other land types and other land types to them
	When $NDVI > 0$, there is no bare land, on the surface of the earth	Bare land to other land types Other land types to bare land
temperature	In the tropics (latitude range 0–23.5), there is no possibility of snow and tundra on the surface of the Earth.	Ice and snow to other land types Other land types to ice and snow Tundra to other land types Other land types to tundra
	In the frigid zone (latitude range 66.5–90), there is only tundra on the surface of the Earth, and no forest can be found.	Forest to other land types Other land types to forest
Moisture	When precipitation range (0–400mm), there is no Tundra, ice and snow on the surface of the earth	Ice and snow to other land types Other land types to ice and snow Tundra to other land types Other land types to tundra

Table 4 Example of experience knowledge and spurious change rules of different properties

3. RESULTS

3.1 Data

The experimental data selected in this study were GlobeLand30 2000 and 2010 land cover products. The experimental area is located in Laos. Laos has a land area of 236,800 square kilometers. Laos is a landlocked country located in the northern part of the Indo-China Peninsula. 80% of Laos is mountainous and plateau, and is mostly covered by forests. It is known as the “Indochina ridge”. The terrain is high in the north and low in the south. Laos has a tropical and subtropical monsoon climate. The rainy season is from May to October, and the dry season is from November to April. The annual average temperature is

about 26 °C. There is plenty of rainfall throughout Laos. The Mekong River, which originated in China, is the largest river in Laos.

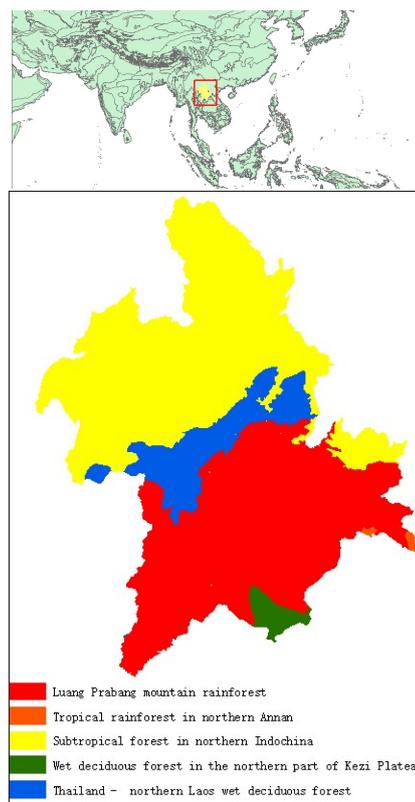


Fig. 3 (a) Location of the experimental area in the global geo-eco zoning and (b) geo-eco zones in the experimental area

The experimental area includes five ecological zones, namely, Luang Prabang mountain rainforest (IM0121), Tropical rainforest in northern Annam (IM0136), Subtropical forest in northern Indochina (IM0137), Wet deciduous forest in the northern part of Kezi Plateau (IM0138), and Thailand - northern Laos wet deciduous forest (IM0139), as shown in Fig. 3(b). The land cover products in 2000 and 2010 are shown in Figs. 4(a) and 4(b). The number of pixels and proportion of each category in 2000 and 2010 are shown in Table 5 and 6

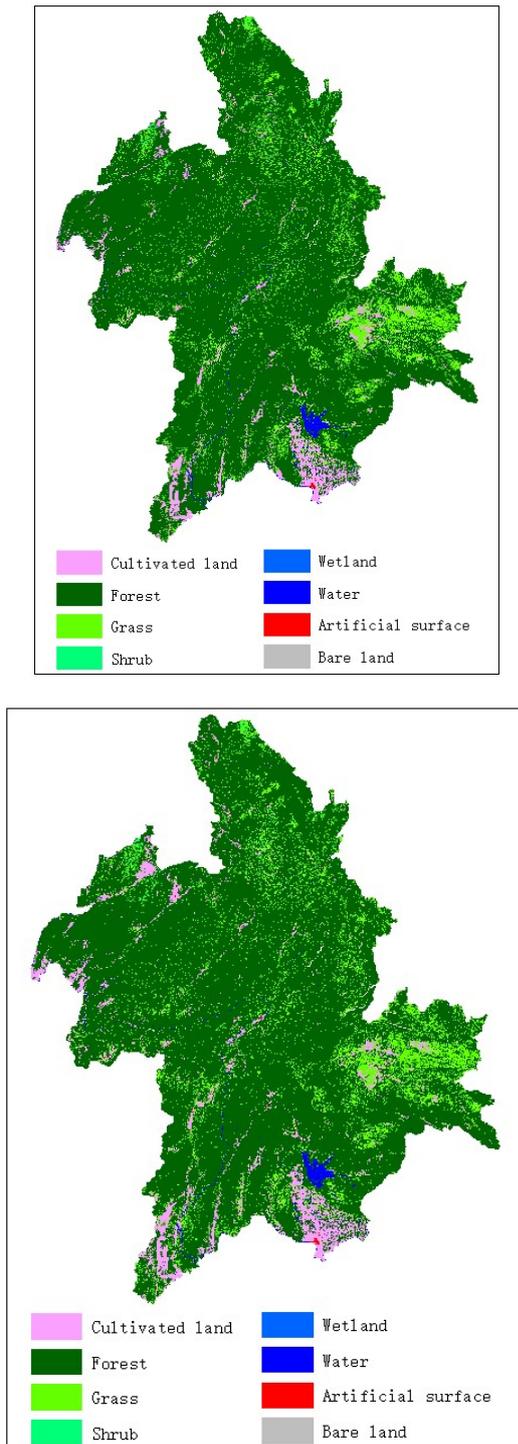


Fig. 4 GlobeLand 30 land cover maps in (a) 2000 and (b) 2010

2000		
Class	Number of pixels	Proportion
Cultivated land	6345596	4.8931%
Forest	107982423	83.2651%
Grass	13824404	10.6600%
Shrub	362188	0.2793%
Wetland	9181	0.0071%
Water	1101713	0.8495%
Artificial surface	57121	0.0440%
Bare land	2542	0.0020%

Table 5 Statistics of surface coverage types of products in 2000

2010		
Class	Number of pixels	Proportion
Cultivated land	7998768	6.1681%
Forest	105948101	81.6997%
Grass	14224263	10.9687%
Shrub	413244	0.3187%
Wetland	4815	0.0037%
Water	1018223	0.7852%
Artificial surface	71800	0.0554%
Bare land	679	0.0005%

Table 6 Statistics of surface coverage types of products in 2010

3.2 Results of spurious change detection

After subtracting the two products (this step was completed in ArcGis), 6390 changed patches were obtained, as shown in Fig. 5(a). A total of 3786 spurious change patches were marked by the geo-eco zoning rule base and are shown in Fig. 5(b).

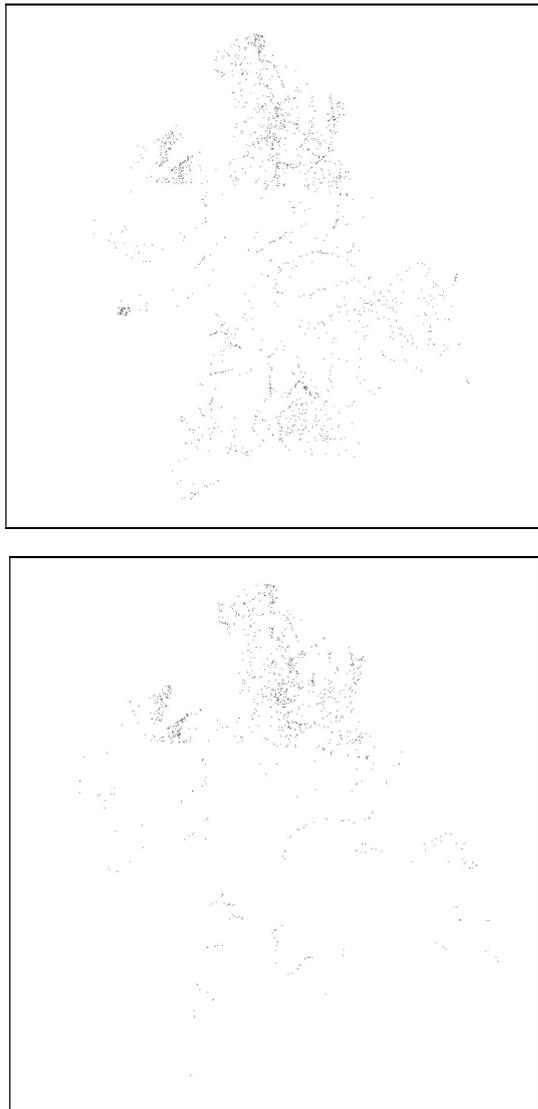


Fig. 5 (a) Changed patches and (b) spurious change patches

Among the spurious change patches detected by the geo-eco zoning rule base in the experimental area, the main land cover types of the 2000 map were forest, grassland, and shrub. The main land cover types in the 2010 map were grassland, shrub, and water. For example, grass to water, it is a kind of spurious change, usually due to spectral differences, and our geo-eco zoning rule base can identify it.

3.3 Accuracy analysis

The accuracy of spurious change detection by using the geo-eco zoning rule base was analyzed. The validation method used Google Earth images to visually interoperate the correctness of the results. A total of 370 spurious change patches out of 3786 were sampled, which accounted for about 10% of the total spurious changes. As shown in Table 7, the correct rate of false change detection in the experimental area was 95.13%.

Number of samples	Number of correct detection	Accuracy
370	352	95.13%

Table 7 Correctness rate of spurious change detection

Interpretation of the Google Earth images indicated that the transformation from grass to water was almost non-existent. However, due to the large difference in hue between the two original Landsat ETM images which were used for classification, misjudgment could easily arise if only spectral is adopted, but this spurious change can be identified by the geo-eco zoning rule base. The confusion matrix was used to analyze the accuracy of change detection after removing spurious changes to determine whether the accuracy of change detection can be improved by using geo-eco zoning for the removal of spurious changes. The method involved selecting several sample points of changed and unchanged patches from the original land cover change map and the change detection result map after removing false changes. Then, Google Earth images were utilized to visually interpret the correctness of the results.

In the experimental area, 325 sample points were randomly sampled from the changed and unchanged spots, totaling at 650 sample points and accounting for about 10% of the total number of change spots in the experimental area. Table 8 shows the confusion matrix of the original change detection in the experimental area, and Table 9 shows the confusion matrix after removing the false change patches in the experimental area. Among the 325 sample points, 213 were unchanged, 112 were changed, and user accuracy was only 34.46%. This result indicates that over-detection was evident in the change detection of subtracting two phase land cover maps. Of the 325 unchanged sample points, only 14 were change points with an accuracy of 95.69%; the overall accuracy was 65.08%. As shown in Table 9, after eliminating the false changes identified by the geo-eco zoning rule base, the number of unchanged spots decreased to 200 out of 325 sample points, which was 13 spots less than that before the spurious change removal. User accuracy was improved by 3%, and the total accuracy increased by 2.46%.

	Changed	Unchanged	Sum	Producer's accuracy
Changed	112	213	325	34.46%
Unchanged	14	311	325	95.69%
Sum	126	524	650	
User's Accuracy	88.89%	59.35%		
Overall Accuracy	65.08%			

Table 8 Original change detection confusion matrix

	Changed	Unchanged	Sum	Producer's Accuracy
Changed	125	200	325	38.46%
Unchanged	11	314	325	96.62%
Sum	136	514	650	
User's Accuracy	91.91%	61.09%		
Overall Accuracy	67.54%			

Table 9 Confusion matrix after removing spurious change patches

4. CONCLUSION

To address the problems in remote sensing image classification and change detection, a method based on a geo-eco zoning rule base was proposed to identify spurious changes. Experiments showed that the accuracy of change detection improved by 2.46% after eliminating spurious changes, which proves that the developed method can improve the accuracy of change detection and has practical significance for the updating of land cover maps.

REFERENCES

- Bailey R G. Delineation of ecosystem regions[J]. *Environmental Management*, 1983, 7(4):365-373.
- Bailey R G. Explanatory (Suppl.)to Ecoregions Map of the Continents[J]. *Environmental Conservation*, 1989, 16(4):307-309.
- Bailey R G, Hogg H C. world ecoregions map for resource reporting[J]. *Environmental Conservation*, 1986, 13(3):195-202.
- Bailey R G. Identifying Ecoregion Boundaries[J]. *Environmental Management*, 2004, 34(1):S14-26.
- Chen, J., Cao, X., Peng, S., & Ren, H., 2017. Analysis and applications of Globeland30: a review. *ISPRS International Journal of Geo-Information*, 6(8), 230.
- Herbertson A J. The Major Natural Regions: An Essay in Systematic Geography[J]. *Geographical Journal*, 1905, 25(3):300-312.
- Holdridge L R. Life zone ecology.[M]. 1967.
- Olson D M, Dinerstein E, Wikramanayake E D, et al. Terrestrial ecoregions of the world : a new map of life on Earth[C]// 2001:933-938.