APPLICATION OF CLASSIFICATION ALGORITHM OF MACHINE LEARNING AND BUFFER ANALYSIS IN TORISM REGIONAL PLANNING

T. H. Zhang, H.W. Ji, Y. Hu, Q. Ye, Y. Lin[*],

College of Surveying and Geo-informatics, Tongji University, 1239 Siping Road, Shanghai, 200092, China
(zhang_th, linyi, yeqin)@tongji.edu.cn, 313817128@qq.com

Commission III, WG III/7

KEYWORDS: the Chaohu Lake, Land Use Classification, Change Monitoring, SVM, Buffer Zone Analysis

ABSTRACT:

Remote Sensing (RS) and Geography Information System (GIS) technologies are widely used in ecological analysis and regional planning. With the advantages of large scale monitoring, combination of point and area, multiple time-phases and repeated observation, they are suitable for monitoring and analysis of environmental information in a large range. In this study, support vector machine (SVM) classification algorithm is used to monitor the land use and land cover change (LUCC), and then to perform the ecological evaluation for Chaohu lake tourism area quantitatively. The automatic classification and the quantitative spatial-temporal analysis for the Chaohu Lake basin are realized by the analysis of multi-temporal and multispectral satellite images, DEM data and slope information data. Furthermore, the ecological buffer zone analysis is also studied to set up the buffer width for each catchment area surrounding Chaohu Lake. The results of LUCC monitoring from 1992 to 2015 has shown obvious affections by human activities. Since the construction of the Chaohu Lake basin is in the crucial stage of the rapid development of urbanization, the application of RS and GIS technique can effectively provide scientific basis for land use planning, ecological management, environmental protection and tourism resources development in the Chaohu Lake Basin.

1. INTRODUCTION

The technology of Remote Sensing (RS) and Geography Information System (GIS) are applied to large-scale environmental and ecological monitoring and change analysis together with the extraction of environmental change information data.

In recent years, there has been many studies on land use and land cover change (LUCC), most of which are about LUCC information extraction. The change detection technology is mainly realized by two methods at present. One method is post-classification comparison method, where the RS images of different phases are classified separately first, and then classification results are compared to obtain the change type, quantity and location information. Another method is to compare the image pixels and directly classify the image to obtain the change information, which is direct comparison analysis (Yadav and Malanson, 2008).

In the aspect of change detection based on image transformation, PG Silva et al. (2003) used two scenes of ETM/Landsat-7 images and other supported data to detect the change of land use and vegetation cover caused by deforestation near the Amazon River in Brazil by using the change vector analysis method. Their research calculated the change intensity, characterized the change characteristics and analyzed the change area quantitatively. In the aspect of post-classification analysis, Mclver (2002) applied the prior probability to the decision tree algorithm for land cover classification with Landsat TM and AVHRR data. Unsupervised classification algorithms were used to estimate probabilities, and the various types were distinguished through Bayesian rules and auxiliary information. S.K.Karan and S.R.Samadder (2016) evaluated the performance of support vector machine (SVM) classification technique with

the maximum likelihood classification (MLC) technique for a rapidly changing landscape of an open-cast mine, and proved that the SVM technique improved the overall land use classification accuracy by almost 6 and 3 % for Landsat 5 and Landsat 8 images. In the post-classification method, the requirement of reasonable classification is higher, and SVM classification method just makes up for the deficiency. Therefore, the post-classification method can have a better effect on land use change analysis.

In related researches of ecology, buffer zone is a typical land-water ecotone, which is the connection and transition zone between river and terrestrial ecosystem. It has the ecological functions of maintaining species diversity, adjusting river microclimate, providing biological habitat, stabilizing river bank, connecting corridors, providing landscape and ornamental resources, purifying and conserving water resource(Chen, 1996). Hualin Xie(2015) proposed a method to identify the key ecological land maintaining ecological security and to put forward some measures for environmental management. Based on the GIS technology, from the view of water security, biodiversity conservation, disaster protection, natural recreation security and human disturbance, an integrated index is proposed to evaluate the ecological importance of regional space. Then a GIS-based approach for evaluating ecological importance was developed, with Xingguo County of China as a case study.

In this study, SVM classification algorithm is used to monitor the LUCC information and quantitatively evaluate the tourism area of Chaohu lake watershed. And the ecological buffer analysis is conducted to set the buffer zone of the Chaohu Lake basin.

---

[*] Corresponding Author

2. STUDY AREAS AND DATASET

2.1 Area

Chaohu Lake basin is located in the middle of Anhui Province with an area of about 4000km^2. The geographical location is about east longitude 116°24′30″ - 118°0′0″ and north latitude 30°58′40″ - 32°6′0″, located between the Yangtze River and Huaihe River systems. Surrounded by low mountains and hills, the terrain of research area is high in the west, low in the east, and flat in the middle. In the subtropical and warm-temperate transitional subtropical monsoon climate regions, the precipitation varies greatly from year to year, and the distribution of precipitation is uneven, mostly concentrated in summer.

Figure 1. Research area of the Chaohu Lake basin (bands of 321)

2.2 Dataset

In this study, Landsat5, Landsat8 and domestic high-resolution Mapping Satellite-1 image data are utilized to analyze the LUCC of Chaohu Lake basin, combined with the corresponding elevation and slope data. Landsat-8 has a total of 11 bands. The spatial resolution of band 1-7 and 9-11 is 30meters, and band 8 is 15miters resolution of panchromatic band. The global coverage period of the satellite is 16 days. The parameters of Landsat data are shown in Table 1.

Experimental data used in the ecological buffer zone analysis is Landsat-8 Satellite RS images in August 2016. Digital Elevation Model (DEM) data and vector data file of Chaohu Lake basin is also introduced in this research.

Time	Satellite	Sensors	Spatial-resolution （m）	bands
1992.07.23	Landsat 5	TM	30	7
1995.09.18	Landsat 5	TM	30	7
2000.09.15	Landsat 5	TM	30	7
2006.07.30	Landsat 5	TM	30	7
2008.05.16	Landsat 5	TM	30	7
2013.05.14	Landsat 8	OLI	30	11
2015.10.11	Landsat 8	OLI	30	11

Table 1. Landsat data for experiment

3. METHODOLOGY

3.1 Land use classification and change monitoring

3.1.1 Data preprocessing and Establishment of interpretation marks library: The RS image data in the year of 1992, 1995, 2000, 2006, 2008, 2013 and 2015 of the

Chaohu Lake basin is selected in this study. The selection principles are as follows: the time is from May to October as far as possible with an interval of approximately 5 years, and the amount of cloud is relatively small. And the preprocessing mainly includes radiation calibration, atmospheric correction, geometric correction and image mosaic.

In order to analyze the spectral characteristics of the ground objects and interpret the RS image of the study area accurately, a library of interpretation marks suitable for the Chaohu Lake basin is established. The details are shown in Table 2.

Land feature	Image	Description
Forest land		Bottle-green, uneven texture and hue, mainly distributed in nature reserve
Cultivated land		Green, regular geometrical shape, clear boundary
Bare land		Earthy yellow, uneven distribution
Water body		Black blue, distributed in lake/water area
Wetland		Light green, distributed vegetation
Algae		Cyan, distributed in water surface, high brightness
Construction		White, black and red interlaced presentation，uneven distribution
Waterlog area		Light blue, distributed among cultivated land

Table 2. Interpretation marks library

3.1.2 Classification method: The classification method in this study is Support Vector Machine (SVM) (Ding et al., 2011). This method has the following main features:

(1) Non-linear mapping is the basic theory of SVM algorithm. Instead of non-liner mapping to high-dimensional space, this algorithm uses inner-product kernel functions.

(2) The optimal hyperplane for feature space partition is the goal of SVM, and the idea of maximizing the margin of classification is the core of SVM method.

(3) SVM is a novel small sample learning method with a solid theoretical basis, and it basically does not involve probability measurement and the laws of charge numbers.

(4) The final decision function of SVM is only determined by a few support vectors. The computational complexity depends on the number of support vectors instead of the dimensions of sample space. This feature helps avoid "dimensional disasters" to some degree.

(5) Because the final result is mainly determined by a few support vectors, a large number of redundant samples are removed, which means that the algorithm is relatively simple and robust.

The weighted kernel function plays an important role that it can linearize the linear algorithm and solve the problem of

linear inseparable pattern recognition in the original input space. Moreover, compared with SVM, the classification method of Extreme Learning Method (ELM) performs faster and more accurate in some classification questions (Huang et al., 2012).

Therefore, in the module of land use classification, the image to be classified was input into an SVM model, and the classification result of the image in the research period was obtained.

3.2 Buffer zone analysis

In buffer analysis, methods of SVM classification, post-processing and classification accuracy evaluation are used in RS analysis. In combination with geospatial analysis techniques, the central Chaohu Lake basin is taken as research objects, and ArcGIS is utilized to conduct hydrological analysis. The width of the forest land and the wetland in the buffer zone are obtained by synthesizing the input data of the relative distance, elevation and slope. And finally, the buffer zone of the forest land and wet land in each catchment area is calculated. The main steps are as follows:

(1) First of all, SVM was used to classify the image of the Chaohu Lake basin by forestland, cultivated land, construction, water body, algae and bare land based on RS image data in August 2016. Then DEM data was input into ArcGIS and filled in to determine the flow direction of the grid cells. And the area of catchment unit was generated and the basin distribution line was divided. In this way, the river network of the basin and distribution of sub-basin was obtained.

(2) According to the DEM data and recognition results of study area, the terrain data such as relative distance, elevation and slope of each catchment is extracted. And the possibility of contamination was comprehensively analyzed.

(3) The setting of the buffer zone in the Chaohu Lake basin was analyzed combined with RS and GIS technology. The specific process is shown in Figure 2.
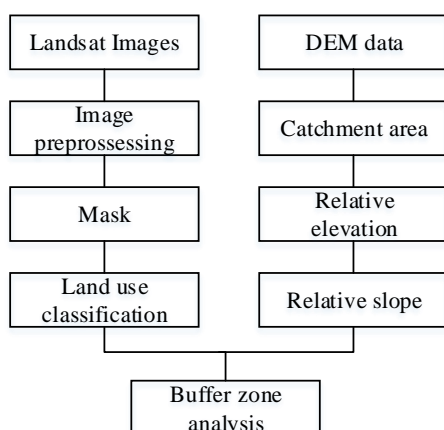

Figure 2. Buffer zone analysis process

4. RESULTS AND ANALYSIS

4.1 Land use classification and change monitoring in Chaohu Lake basin

According to the classification of the land use in the Chaohu Lake basin from 1995 to 2015, the spectral characteristics of eight ground objects including construction land, water body, forest land, cultivated land, bare land, water land, wet land and algae are analyzed. And the area statistics and change situation of classification results are also analyzed. The temporal and spatial variation between each adjacent two years of the wet land are evaluated to obtain the temporal and spectral variation results of every ground object in the Chaohu Lake basin. The area variation curves are displayed in Figure 3, and the result of change monitoring is shown in Figure 4.
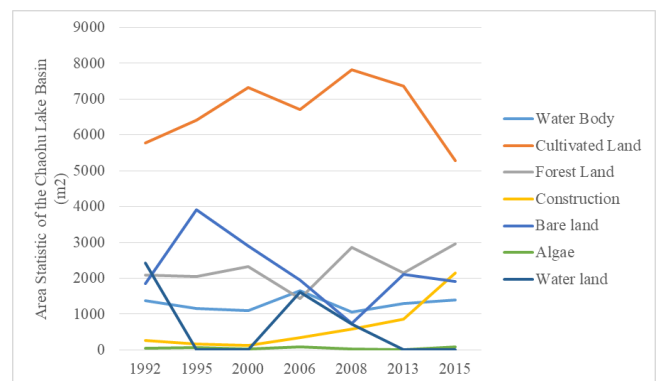

Figure 3. Area variation statistics data (from 1992 to 2015)

With the passage of time, the area of construction land increased gradually. The changes of cultivated land and forest land were of similar variation trends, which increased firstly and then decreased. The area of water body basically remained stable in 20 years. In addition, the area of bare land went up in the first year, went down subsequently and then went up back. The algae shoed an overall trend of growth. And the area of wetland was in a decreasing tendency.

With the passage of time, the construction land increased year by year from $263.85km^2$ in 1992 to $2149.03km^2$ in 2015. The increase of construction land around the Chaohu Lake basin mainly occurred in the surrounding area of Hefei, which is located on northwest of Chaohu. And it is consistent with the social and economic development level of Hefei and the Chaohu Lake basin in recent years.

The area of cultivated land was firstly increasing and then decreasing. The trend is due to the fact that part of bare land was reclaimed as cultivated land before 2008. After 2013, there was a trend of rapid decline, which was related to the increase of the occupied area of construction land.

In the past 20 years, the area of water body has remained basically stable. And the flood in 2006 resulted in the increase of water body area. It was also related to the large variation of precipitation in the Chaohu Lake basin.

The change trend of forest land was basically similar to that of cultivated land. In July 2006, there was a flood in the Chaohu Lake basin and some forest land was confusedly identified as water body, so a significant decrease is reflected in the curve of forest land. However, the overall trend is in a

state of increasing.

The overall appearance of bare land was increasing first, then decreasing and finally increased. The significant decrease between 1995 and 2008 was due to the phenomenon that bare land was reclaimed as forest land, cultivated land and construction land during this period. It indicated that human activity plays a crucial role in the land use change. Since 2008, there has been an increasing trend of bare land because of the seasonal factors. Some crops had been harvested and the cultivated land has been confusedly identified as bare land.

The change of algae showed an increasing trend from 47.43km^2 in 1992 to 83.52km^2 in 2015. It is closely related to the environmental pollution caused by local human activities. In recent years, with the rapid development of the Chaohu Lake basin, the issue of water pollution has become serious.

With the deterioration of Chaohu lake water environment, from 1992 to 2013, the wetland area was in a decreasing trend. However, the area increased suddenly in 2015. Through visual interpretation, the reason is the excessive reclamation of humanity. The wet land near the Chaohu Lake basin was reclaimed as cultivated land, and in the RS image, it was mistakenly identified as wetlands. Although the classification result was numerical increase, the actual situation should be reduce.

As to the change of waterlog land, it is shown in the classification result that there were floods occurred in 1992, 2006 and 2008 in varying degrees.
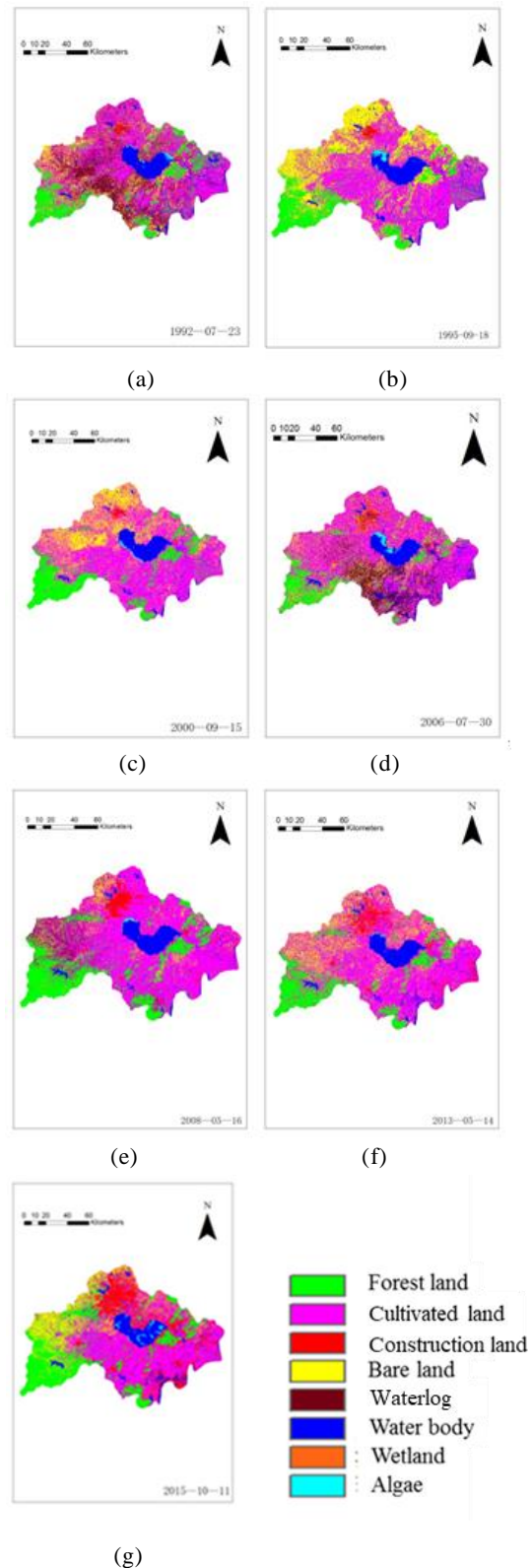

Figure 4. Classification results of the Chaohu lake basin from 1992 to 2015: (a) 1992, (b) 1995, (c) 2000, (d) 2006, (e) 2008,　(f) 2013, (g) 2015

4.2 Ecological buffer zone analysis

The hydrographic map of the Chaohu Lake was generated to determine the range of watershed in this study. And the hydrological analysis was performed on the sub-basin of the Chaohu Lake. The results of catchment distribution is displayed in Figure 5 and the black line represents the division boundary of each catchment. The calculation result of relative elevation and slope are shown in Figure 6 and the analysis result of buffer zones for each catchment calculated from those results is displayed in Figure 7.
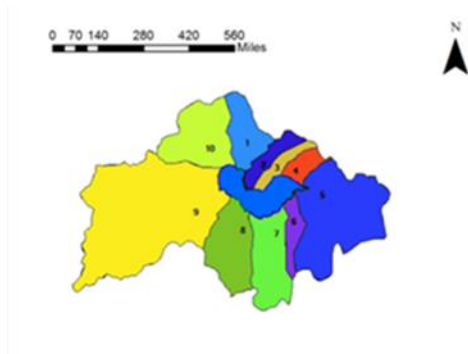

Figure 5. Division of catchment area in the Chaohu Lake basin
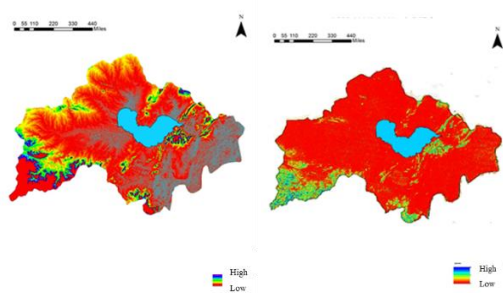

Figure 6. Relative elevation (left) and slope (right) of the Chaohu Lake basin

By dividing the catchment in the study area, the relative area were analyzed comprehensively, and buffer zones were established to provide analysis means for the control of pollution around the Chaohu Lake. The buffer zone as divided into two parts: forest land and wetland. In the case of the surrounding site conditions, priority was given to the establishment of forest buffer zone. In the condition the land use of lake surrounding was unsuitable for establishing forest buffer zone, wetland buffer zone was set up inward by taking the lakeshore line as a boundary. Generally, a mixed arrangement mode of combining two buffer zones was adopted. And the width settings of buffer zone in the Chaohu Lake basin is displayed in Table 3.

Catchment Number	Forest land buffer (m)		Wetland buffer (m)	
	part	whole	part	whole
1	180	100	100	100
2	0	80	130	160
3	0	80	70	160
4	0	80	70	320
5	0	240	0	80
6	0	240	0	240
7	0	320	0	160
8	0	320	0	160
9	0	200	60	120
10	70	180	0	320

Table 3. Buffer zone width setting
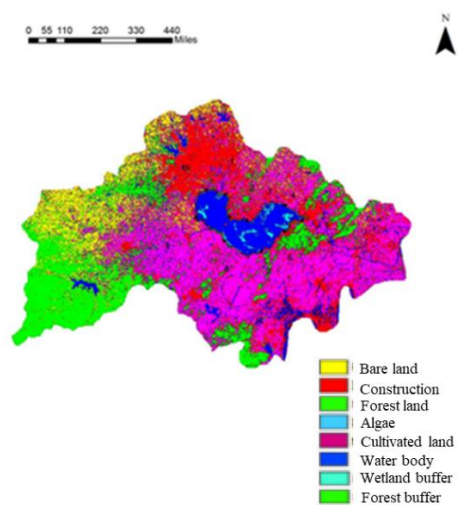
(a)
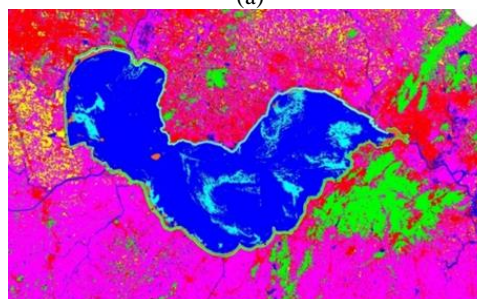
(b)

Figure 7. Buffer zone analysis (a) and local zoom (b)

5 CONCLUSIONS AND DISCUSSION

This study makes an analysis of the land use and change situation from 1992 to 2015 in the Chaohu Lake basin. Based on the result compared with the profile of local statistics, we can draw a conclusion that RS and GIS technology can be applied to obtaining information of large scale regional ecological environment. Moreover, it is effective to apply the information of land use and the change information acquired by RS imagery and to analyze the ecological buffer zone by GIS tools. The method in this study can also be applied to RS quantitative analysis and evaluation in other ecological regions.

In addition, the construction of the Chaohu Lake basin is in the crucial stage of the rapid development of urbanization in China. It is an important topic to control the process of

construction and establish a scientific development mode for other tourism leisure regions. Meanwhile, it should be effective on the scientific implement of regional planning and the efficient utilization of resources. The technology of RS and GIS integration technology can help objectively grasp the resource utilization, temporal-spatial distribution, changes in ecological environment and other information of a certain regional development. The method of spatial information extraction and analysis in this study is a powerful way to realize a coordinated regional planning, which can provide technical support and decision basis for the management of urban planning and construction and the regional sustainable development.

ACKNOWLEDGEMENTS

REFERENCES

Chen, J., 1996. Riparian Vegetation Characteristics and Their Functions in Ecosystems and Landscapes. *Chinise Journal of Applied Ecology*(04): 439-448. (in Chinese)

Ding, S., Qi, B. and Tan, H., 2011. An Overview on Theory and Algorithm of Support Vector Machines. *Journal of University of Electronic Science and Technology of China*(01): 2-10. (in Chinese)

Huang, G.B., Zhou, H., Ding, X. and Zhang, R., 2012. Extreme learning machine for regression and multiclass classification. *IEEE Trans Syst Man Cybern B Cybern*, 42(2): 513-29.

Karan, S.K. and Samadder, S.R., 2016. Accuracy of land use change detection using support vector machine and maximum likelihood techniques for open-cast coal mining areas. *Environmental Monitoring and Assessment*, 188(8).

McIver, D.K. and Friedl, M.A., 2002. Using prior probabilities in decision-tree classification of remotely sensed data. *Remote Sensing of Environment*, 81(2): 253-261.

Silva, P. G., et al. Change vector analysis technique to monitor selective logging activities in Amazon. *Geoscience and Remote Sensing Symposium, 2003. IGARSS '03. Proceedings. 2003 IEEE International IEEE*, 2003:2580-2582 vol.4.

Xie, H., Yao, G. and Liu, G., 2015. Spatial evaluation of the ecological importance based on GIS for environmental management: A case study in Xingguo county of China. *Ecological Indicators*, 51: 3-12.

Yadav, V. and Malanson, G., 2008. Spatially explicit historical land use land cover and soil organic carbon transformations in Southern Illinois. *Agriculture, Ecosystems & Environment*, 123(4): 280-292.