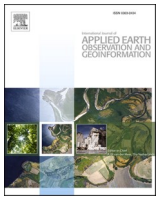
International Journal of Applied Earth Observations and Geoinformation 102 (2021) 102413

Contents lists available at ScienceDirect

International Journal of Applied Earth Observations and Geoinformation

journal homepage: www.elsevier.com/locate/jag



Long-term remote sensing monitoring on LUCC around Chaohu Lake with new information of algal bloom and flood submerging

Yi Lin [a,b], Tinghui Zhang [a,b], Qin Ye [a,b,*], Jianqing Cai [c], Chengzhao Wu [d], Awase Khirni Syed [e], Jonathan Li [e]

[a] College of Surveying and Geo-informatics, Tongji University, Shanghai 200092, China
[b] Research Center of Remote Sensing & Spatial Information Technology, Shanghai 200092, China
[c] Institute of Geodesy, University of Stuttgart, Stuttgart 70174, Germany
[d] College of Architecture and Urban Planning, Tongji University, Shanghai 200092, China
[e] Department of Geography and Environmental Management, University of Waterloo, Canada

ARTICLE INFO

Keywords:
Land use spatial pattern
Image classification
Change monitor
Flood detection
The Chao Lake Basin

ABSTRACT

Human settlements are guided by the proximity or availability of a natural resource such as river or lake basins containing set of streams. The harmonious development of human activity and natural conditions along watershed areas needs close attention and in-depth study. In this paper, the urban agglomerations and ecological spaces in the Yangtze River Delta, China, the Chao Lake Basin and its surrounding watershed ecosystem is taken as research subject for its serious environmental degradation problems during social and economic development. This paper adopted an effective machine learning algorithm (kernel-ELM) to extract land use and land /cover information, and to analyze the land use/cover pattern evolution rules of the Chao Lake Basin with long term Landsat imagery. Subsequent studies were then carried out to demonstrate the flood-affected area and its ecological impact in the basin in 2020, to reveal the occupation on land cover types. The results indicate Conclusions are drawn from the experiment results: (1) There has been significant change in cultivated land, forest land and construction land out of six key land cover types with dynamic degree of −10.17%, 4.61, 67.04% respectively. (2) Algae bloom pollution was extracted from pattern classification results and it was up to 15% of the total water area by the year 2018. (3) The occupation on land use/cover types of the flood was revealed. The results prove effective application of remote sensing technology in environmental analysis and planning for data-driven evaluation of governing policy. This work serves as a scientific basis for environmental management and regional planning in the Chao Lake Basin and can be served as a basis and a reference for evaluating an ecological policy and its impact for other economic developing watershed human settlements with ecological issues.

1. Introduction

Human existence is embedded in the availability of natural resources. The human civilization evolved using these natural resources such as water, fertile land and adequate climatic conditions for their development. Human settlements are guided by the proximity or availability of these resource. Areas of land that contain a common set of streams, rivers and lakes that all drain into a single larger body of water be it lake or river or an ocean have seen human settlements. The harmonious development of human activity and nature along these natural watershed areas such as Lakes is of interest for this research. Adrian *et al.* (2009) state that the lake environments record the changes of regional ecology on a spatio-temporal frame. Lakes are integral part of the ecosystem processing various ecosystem functions and provide habitat for a range of species and form essential components on the earth surface(Dörnhöfer and Oppelt, 2016). The lake is a water-filled area, located in a reservoir, surrounded by land, apart from any river or other source that feeds or drains the lake. The natural ecosystem and spatial pattern of the lake basin are both of critical significance to the lake itself and its perimeter. The inconsistency between urban growth and conservation of the lake is also becoming increasingly obvious (Wu et al., 2019). Monitoring and recognizing the state of the evolution of the lakes

* Corresponding author at: College of Surveying and GeoInformatics, Tongji University, Shanghai 200092, China.
E-mail addresses: linyi@tongji.edu.cn (Y. Lin), zhang_th@tongji.edu.cn (T. Zhang), yeqin@tongji.edu.cn (Q. Ye), cai@gis.uni-stuttgart.de (J. Cai), wuchzhao@qq.com (C. Wu), awase008@gmail.com (A. Khirni Syed), junli@uwaterloo.ca (J. Li).

https://doi.org/10.1016/j.jag.2021.102413
Received 5 February 2021; Received in revised form 17 June 2021; Accepted 22 June 2021

and the associated watershed are highly relevant for scientists and decision makers alike (Palmer et al., 2015).

To date, the widely used criteria for monitoring lakes and basins primarily involves the following: (a) monitoring land cover changes of the lake basin or surrounding wetlands (Aghsaei et al., 2020; Guo et al., 2008; Wolter et al., 2006); (b) spatio-temporal volumetric variation analysis of water body in area, variations in water level (De Wit and Stankiewicz, 2006; Duan and Bastiaanssen, 2013); (c) presence or occurrence of water surface vegetation and algae detection (Kutser et al., 2006; Wang and Shi, 2008). There is a strong correlation between urban development and the changes in lake basin. It is therefore, of utmost importance to accurately monitor and understand the spatio-temporal changes to the lake basin due to land use and land cover change (LUCC) of its perimeter. Sun *et al.* (2014) states that we can use remote sensing technology to accurately describe these variations and assess the impact of human activities on the neighboring ecosystem on a spatio-temporal frame. In addition, owing to the diversity of remote sensing image resolution, various types of satellite imagery have been used in lake monitoring studies of different scale, from a partial basin to an interconnected lake basin, or even to a large number of lakes in the entire area. (Dube et al., 2014; Lin et al., 2018; Liu et al., 2014; Luo et al., 2017; Stow et al., 2004; Zhao et al., 2012).

The land use cover change (LUCC) is a widely used indicator describing the strong relation between human activity and the changes in the lake ecosystem to monitor the overall physiological shift in lake conditions, including water depth, water level and water surface variations (Tong et al., 2016; Wasige et al., 2013). Along with the LUCC statistic, we may also use environmental parametric indices as additional metrics(Anderson, 1976) to monitor lake basin of varying scales to the whole catchment area (Jorgenson and Grosse, 2016; Nsubuga et al., 2017). For example, change analysis studies were conducted with Support Vector Machines (SVM) using QuickBird images of Poyang Lake wet land basin for the period of 1973 to 2013 (Han et al., 2015). Similar studies were carried out and validated by Were et al. (2013) and post-classification comparative analysis was performed to detect the changes in Eastern Mau forest reserve and Lake Nakuru drainage basin, Kenya for the period of 1973 and 2011. Lin et al. (2018) applied machine learning classification techniques to detect changes in distribution of vegetation and land cover types for estuary basin and wetlands.

The transformation of the land surface to urban use is one of the most irreversible human impacts on the ecosystem. The Yangtze River Delta cities have seen drastic urbanization due to rapid economic growth since the late 1990′s. This resulted in gradual restructuring of land use type and deeply transforming the land functions. The land functions are determined by the availability and proximity to natural resources such as water, fertile land and biodiversity. Now the economy of Yangtze River Delta in China has turned to a stage of high-quality development. This prompted the Chinese government to establish a Regional Integrated Development plan for the Yangtze River Delta in 2019. This region comprises of Yangtze River Economic Belt and East China region, Shanghai, Jiangsu, Zhejiang and Anhui provinces. The environmental effects of urban growth go far beyond urban areas themselves on several levels. Future urbanization would pose a significant challenge to the neighboring ecosystems, if it does not accommodate scientific impact assessment during planning. The region of Chao Lake Basin falls in this Economic Belt and has been designated as a national tourism region since 2015. This study focuses on a long-term monitoring of the spatial pattern of land use change in the Chaohu Lake Basin, and then on the quantitative analysis of the impact on water environment caused by the socio-economic activities around the lake. Especially, the Chao Lake Basin has suffered from frequent flooding in the past few months, impacting the ecosystem drastically.

To guarantee the classification accuracy and efficiency and describe land use pattern evolution and its influences simultaneously, the combination of information extraction, automatic imagery interpretation and change detection is the tendency of future monitoring methods. Numerous studies have been conducted in the recent years on the retrieval of land use pattern information and subsequent spatial or temporal interpretation with remote sensing imagery classification. Based on multi-temporal classification results, the LUCC between each two phases can also be calculated precisely. Most machine learning methods have been normally used to extract land cover information from remote sensing data. It is to be noted that the accuracy of information extraction is directly influenced by the accuracy of the machine learning classification algorithms used and proper machine learning methods are suitable to overcome challenges in spatial data handling and improve classification performance (Du et.al, 2020). For example, the performance of SVM classification technique with the maximum likelihood classification (MLC) technique have been applied for a rapidly changing landscape of an open-cast mine to validate the precision and efficiency, and the SVM has improved the classification accuracy on Landsat satellite images (Karan and Samadder, 2016). Many recent publications have already demonstrated the benefits of machine learning approaches for land/user cover analysis and eventual systematic implementation (Cai et al., 2018; Duro et al., 2012; Khatami et al., 2016; McIver and Friedl, 2002; Rodriguezgaliano et al., 2012). It is of great importance to choose an appropriate method to conduct classification experiments precisely and efficiently.

In this study we used an optimized Single-hidden Layer of Feed-Forward Network (SLFN) classification technique, referred as Extreme Learning Machine (ELM) with a kernel function to perform land cover classification. For this purpose, the scope of the study is to monitor and to quantitively analyze the spatial pattern of land use evolution in the Chao Lake Basin in over twenty years with long-term series remote sensing images. These objectives are outlined here:

1) to validate the accuracy of the developed classification method that is the K-ELM algorithm through local remote sensing data in the Chaohu Lake Basin;
2) to conduct quantitatively land use pattern analysis over twenty years based on the remote sensing classification results;
3) to demonstrate the spatial distribution of two ecological events in Chaohu Lake: persistent algae bloom outbreaks on lake surface for multi years, and the 2020 mega-flood disaster in July 2020;
4) to reveal the long-term evolution of land use spatial pattern and to draw critical insights that serve as reference for future local and regional urban planning.

2. Study area and datasets

2.1. Study area

The study area selected for this research is Chao Lake, situated on the outskirts of Chaohu and Hefei in Anhui Province, China. It is the fifth largest fresh water lake in China with a water surface area admeasuring over 700 km^2 and sits on the south side of the Chao Lake rift basin. The geographical location of the whole basin 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 the Huaihe River systems. The Chao Lake area is a valley surrounded by low mountainous regions and hilly areas. The topography tends to be high in the west and gradually decrease towards the east and flat in the middle. The region has warm-temperate subtropical monsoon climate zones, where the precipitation varies considerably from year to year and precipitation distribution is irregular and mostly concentrated in summer. In the neighboring vicinity of the Chao Lake Basin is the Hefei City which serves as urban conglomeration along the Yangtze River and has direct impact on its ecosystem. There are efforts to conserve the Chao Lake Basin for ecology-based tourism zone. However, the region surrounding Chao Lake Basin has seen drastic increase in the population to twice in 2018 as compared to the population in 1995, which also attributed to high-speed transport rail and road network. Table 1 outline socio-demographic statistics of Hefei City.

Table 1
Economic and demographic statistics of Hefei City.

Statistical categories	1995	2001	2018
Population (million)	4.11	4.421	8.087
Grain Yield (megatons)	1.47	1.333	3.0135
Rate of increasing output value of leading industries (%)	/	12.3	15.6
GDP (billion Yuan)	16.75	36.34	782.291

It is of most importance to devise regulations that emphasize on sustainable development of the region bearing in mind the ecological and socio-economic indicators. Xu *et al.* (1999) states that rapid urbanization in the surrounding regions has contributed to increased pollution of Chao Lake Basin affecting the biodiversity of the ecosystem. It is to be noted that water quality level has reached pollution levels worse than Grade V (The lowest level of water pollution evaluation indicators in Environmental Quality Standards for Surface Water) (MEEPPC, 2002). The intent of this study is to identify regions with high concentrations of pollutants due to rapid urbanization in the vicinity and analyze spatiotemporal changes across the region.

Fig. 1 illustrates the Landsat OLI RGB image which is the study area. It covers the whole watershed of the Chao Lake, the Hefei City located in the north and the Chaohu City located in the east, including the polder in the southwest and the forest in the north and southeast. The main body of the Chao Lake including two islands, Mushan Island and Gushan Island, is the core resource of the whole National Tourism and Leisure Region of the basin. The lake is surrounded by the Hefei City in the northwest, the Chaohu City in the east and the Feixi, Shucheng and Lujiang Counties in the west, which are composed of abundant agriculture and landscape resource.

Furthermore, a sub-region image is used for algorithmic comparison studies covering the east side region of the Chao Lake and the Chaohu City. This contains all the land use/cover change types, namely, *water, forest, cultivated land, construction land, bare land* and aquatic vegetation cover *algae bloom*, a total of six classification types.

2.2. Data sets collection

In this study, seven cloudless Landsat optical images from 1995 to 2018 for land use pattern evolution analysis and two Landsat images in 2020 for flood detection of the Chao Lake Basin were selected, which were separately acquired from Landsat Thematic Mapper ™, Enhanced Thematic Mapper Plus (ETM +) and Operational Land Imager (OLI) (See Table 2). The Landsat instrument captures data at 30 m spatial resolution. It has 11 bands distributed across the electromagnetic spectrum compared to 3 or 4 NAIP imagery has. It offers improved spectral and spatial characteristics that serve as valuable base for a wide range of applications. The study takes into consideration planting season and algae bloom period for classification over a period of April to October months each year for land use cover change analysis. Each Landsat image has a spatial resolution of 30 m and is 3000 × 2400 pixels in size. In total, study area admeasures around 6480 km^2. For the purpose of detailed classification studies, the sub-region image with 496 × 472 pixels in size and covers an area of 210 km^2 approximately was used.

3. Method

3.1. Preprocessing

The radiometric calibration and atmosphere correction have been processed on all RS images with ENVI 5.3 software (https://www.harrisgeospatial.com/) to remove the influence of atmosphere radio mutation and to achieve real surface reflectance. For the TM and ETM + images, geometry correction has been also carried out. Ten ground control points were selected on each image and the RMSE (root mean square error) of each point is computed and is less than one pixel.

The feature space consists of one optimal bandwidth combination, normalized difference vegetation index (NDVI), Modified Normalized Difference Water Index (MNDWI), and the wetness dimension of the

Table 2
Data source and description of Landsat images.

Applications	Satellite Sensor	Acquision Date (YYYY/MM/DD)
Land cover pattern analysis	Landsat-5 TM	1995/09/02
	Landsat-7 ETM+	2001/07/24
	Landsat-5 TM	2006/07/30
	Landsat-5 TM	2009/06/04
	Landsat-5 TM	2011/04/23
	Landsat-8 OLI	2015/10/11
	Landsat-8 OLI	2018/10/03
Flood detection	Landsat-8 OLI	2020/06/02
	Landsat-8 OLI	2020/07/20

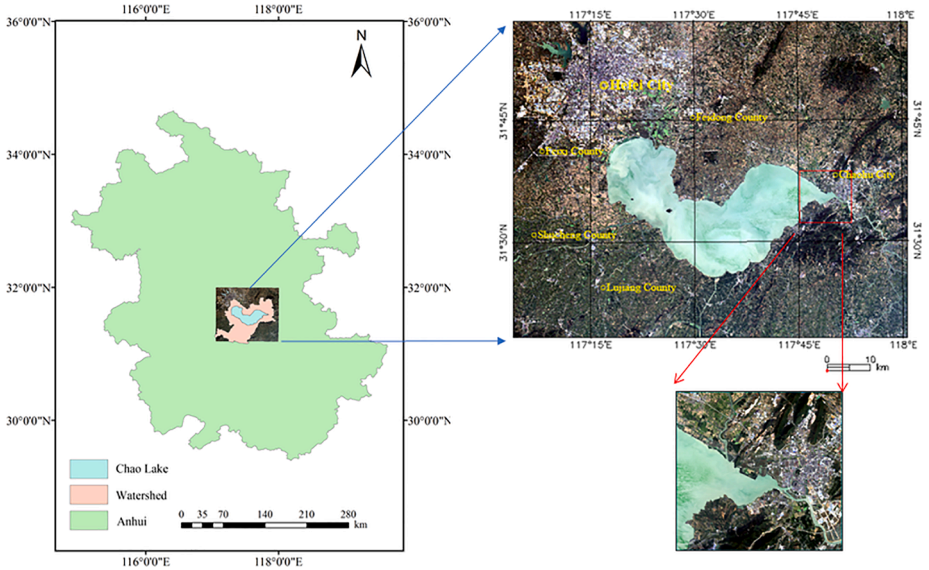


Fig. 1. Study area of the Chao Lake Basin in this paper and the sub-area for accuracy comparison experiment (Oct.2015 Landsat 8/OLI RGB image).

Kanth–Thomas (K-T) transformation. The optimum index factor (OIF) (Dwivedi and Rao, 1992), which weighs the variance of each individual band, was utilized to select the optimal bandwidth combination. According to the results of the OIF, bands 3, 4 and 5 (red, near-infrared, and short-wave infrared) for TM/ETM + data and bands 3, 5, and 6 (green, near-infrared, and short-wave infrared bands) for OLI data were chosen as a feature vector.

3.2. Land cover classification

3.2.1. Extreme learning Machine

The most critical role in the classification system development process is to construct a network model for particular experimental problems. Many researchers use the technique of acquiring prior information from existing data samples to construct a neural network model. These approaches, such as neuron network algorithms, have gone through a mechanism ranging from basic to complex, from specific to comprehensive and from single method to multi-combined. However, there are some issues that need to be tackled in the artificial neural network. For example, it is simple to collapse into the local minimum for non-linear optimization problems, training efficiency is not high enough, several parameters need to be set manually, and the activation function needs to be separated. To solve the problems above, a machine learning approach of single-hidden layer of feed-forward networks (SLFNs) called the Extreme Learning Machine (ELM) was proposed (Huang et al., 2004), which assigns the input weights and number of hidden layers randomly.

Working for SLFNs other than normal SLFNs, the hidden layer in ELM need not be tuned, if only the activation function$g(x)$is infinitely differentiable. Thus, the input weight can be randomly initialized, and it has little effect on the outputs (Huang et al., 2006).

For N arbitrary distinct training samples$(\mathbf{x}_i,\ \mathbf{y}_i)$, where $\mathbf{x}_i = [x_{i_1}, x_{i_2}, \ldots, x_{i_n}]^T \in \mathbf{R}^n$and $\mathbf{y}_i = [y_{i_1}, y_{i_2}, \ldots, y_{i_m}]^T \in \mathbf{R}^m$. Mathematically, a standard SLFN with L hidden layer neurons can be modelled as

$$\sum_{i=1}^{L} \beta_i g_i(\mathbf{x}_j) = \sum_{i=1}^{L} \beta_i g(\mathbf{w}_i \cdot \mathbf{x}_j + b_i) = \mathbf{o}_j,\ j = 1, 2, \cdots, N, \tag{1}$$

where $\mathbf{w}_i = [\mathbf{w}_{i_1}, \mathbf{w}_{i_2}, \ldots, \mathbf{w}_{i_n}]^T$is the weight vector connecting the ith hidden neuron and all the input neurons, $\beta_i = [\beta_{i_1}, \beta_{i_2}, \cdots, \beta_{i_m}]^T$is the weight vector connecting the ith neuron in the output layer and all the hidden neurons and b_i is the bias for the ith hidden neuron. As it is proved, a standard SLFN with L hidden neurons, which are activated by the activation function$g(x)$,can approximate the N training samples with zero error. Mathematically, it implies that $\Sigma_{j=1}^{L} ||\mathbf{o}_j - \mathbf{y}_j|| = 0$, i.e., there exist a set of $(\beta_i, \mathbf{w}_i, b_i)$such that

$$\sum_{i=1}^{L} \beta_i g(\mathbf{w}_i \cdot \mathbf{x}_j + b_i) = \mathbf{y}_j,\ j = 1, 2, \cdots, N.$$

$$\mathbf{H}(\mathbf{w}, \mathbf{b}, \mathbf{x}) = \begin{bmatrix} g(\mathbf{w}_1 \cdot \mathbf{x}_1 + b_1) & \cdots & g(\mathbf{w}_L \cdot \mathbf{x}_1 + b_L) \\ \vdots & \ddots & \vdots \\ g(\mathbf{w}_1 \cdot \mathbf{x}_N + b_1) & \cdots & g(\mathbf{w}_L \cdot \mathbf{x}_N + b_L) \end{bmatrix}_{N \times L},\ \beta_i = \begin{bmatrix} \beta_{i_1} \\ \beta_{i_2} \\ \vdots \\ \beta_{i_m} \end{bmatrix}$$

$$\mathbf{w}_i = [\mathbf{w}_{i_1}, \mathbf{w}_{i_2}, \ldots, \mathbf{w}_{i_n}]^T$$

$$\mathbf{y}_i = [y_{i_1}, y_{i_2}, \ldots, y_{i_m}]^T \in \mathbf{R}^m$$

$$\mathbf{x}_i = [x_{i_1}, x_{i_2}, \ldots, x_{i_n}]^T \in \mathbf{R}^n$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_L^T \end{bmatrix}_{L \times m} \text{ and } \mathbf{Y} = \begin{bmatrix} \mathbf{y}_1^T \\ \vdots \\ \mathbf{y}_N^T \end{bmatrix}_{N \times m}. \tag{2}$$

The above N equations can be written compactly as

$$\underset{N \cdot L\ L \cdot m}{\mathbf{H}}\ \beta = \underset{N \cdot m}{\mathbf{Y}}, \tag{3}$$

$\mathbf{H}$ is called the hidden layer output matrix of the neural network, the i^{th} column of $\mathbf{H}$ is the i^{th} hidden node output with respect to inputs $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N$.

Traditionally, throughout the training procedure, it is aimed at finding a specific set of $(\widehat{\beta}_i, \widehat{\mathbf{w}}_i, \widehat{b}_i)$which minimizes the cost function

$$E = \sum_{j=1}^{N} \left(\sum_{i=1}^{L} \beta_i g(\mathbf{w}_i \cdot \mathbf{x}_j + b_i) - \mathbf{y}_j \right)^2 \tag{4}$$

The stochastic gradient descent training algorithm is generally applied to search the minimum of the cost function by iteratively updating all the parameters$(\widehat{\beta}_i, \widehat{\mathbf{w}}_i, \widehat{b}_i)$. It is particularly time-consuming and sometimes it will get trapped in local minima, when an improper learning rate is selected.

Whereas to the stochastic gradient descent training algorithm, the proposed ELM (Huang et al., 2006) has much faster training speed. In ELM, the input weight and the bias for the hidden layer do not need to be estimated, the input weight and the bias can be randomly assigned. Once these parameters are initialized and fixed, the hidden layer (feature mapping) matrix $\mathbf{H}$ remains unchanged. Training an ELM is to simply find a least-squares solution$\widehat{\beta}$of the linear system (3) $\mathbf{H}\beta = \mathbf{Y}$, which should minimize the residuals, it means

$$\|\mathbf{H}\widehat{\beta} - \mathbf{Y}\| = \min_{\beta} \|\mathbf{H}\beta - \mathbf{Y}\| \tag{5}$$

By using Moore-Penrose generalized inverse (Rao and Mitra, 1972), the least-squares solution$\widehat{\beta}$can be estimated as

$$\widehat{\beta} = \mathbf{H}^{\dagger}\mathbf{Y}. \tag{6}$$

Generally, when we have more distinct training samples than hidden neurons ($N > L$), it is an over-determined case and if the normal matrix$\mathbf{H}^{\mathrm{T}}\mathbf{H}$is nonsingular. The $\mathbf{H}^{\dagger}$ is generated by

$$\mathbf{H}^{\dagger} = (\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}}$$

i.e.,

$$\widehat{\beta} = \mathbf{H}^{\dagger}\mathbf{Y} = (\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{Y}. \tag{7}$$

If the normal matrix$\mathbf{H}^{\mathrm{T}}\mathbf{H}$is singular or near-singular, it is necessary to apply the regularization method for the estimation of Eq. (3)

$$\widehat{\beta} = \mathbf{H}^{\dagger}\mathbf{Y} = \left(\mathbf{H}^{\mathrm{T}}\mathbf{H} + \frac{\mathbf{I}}{C}\right)^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{Y}. \tag{8}$$

3.2.2. Extreme learning Machine with kernel function

In some actual application cases, to solve the linear inseparable problems in initial space, the activation function can be replaced by kernel function. In this section the kernel model is introduced, which uses N arbitrary distinct training samples$\mathbf{x}_i$, where $\mathbf{x}_i = [x_{i_1}, x_{i_2}, \ldots, x_{i_n}]^T \in \mathbf{R}^n$for basis function design. The kern model is also linear in term of parameters (Sugiyama, 2015)

$$\sum_{i=1}^{L} \beta_i K(\mathbf{x}, \mathbf{x}_j) = \mathbf{y}_j,\ j = 1, 2, \cdots, N \tag{9}$$

As a kernel function, the Gaussian kernel would be the most popular choice considering both the accuracy and efficiency:

$$K(\mathbf{x}, \mathbf{c}) = exp\left(-\frac{||\mathbf{x} - \mathbf{c}||}{2h^2} \right), \tag{10}$$

h and $\mathbf{c}$ are called the Gaussian bandwidth and the Gaussian centre

respectively.

Thus, the estimation for the kernel model Eq. (8) can also be obtained in the same way, by replacing the hidden layer output matrix **H** of the neural network with kernel matrix **K** in Eq. (6):

$$\mathbf{K} = \begin{bmatrix} K(\mathbf{x}_1, \mathbf{x}_1) & \cdots & K(\mathbf{x}_1, \mathbf{x}_N) \\ \vdots & \ddots & \vdots \\ K(\mathbf{x}_N, \mathbf{x}_1) & \cdots & K(\mathbf{x}_N, \mathbf{x}_N) \end{bmatrix}_{N \times N} \tag{11}$$

$$\widehat{\beta} = \left(\mathbf{K}^{\mathrm{T}}\mathbf{K}\right)^{-1}\mathbf{K}^{\mathrm{T}}\mathbf{Y}. \tag{12}$$

If the normal matrix$\mathbf{K}^{\mathrm{T}}\mathbf{K}$is singular or near-singular, it is necessary to apply the regularization method for the estimation of Eq. (11)

$$\widehat{\beta} = \left(\mathbf{K}^{\mathrm{T}}\mathbf{K} + \frac{\mathbf{I}}{C}\right)^{-1}\mathbf{K}^{\mathrm{T}}\mathbf{Y}. \tag{13}$$

In K-ELM function, instead of the hidden layer output function **H**, it is the kernel function which need to be given to solve the output results. And the **H** can remain unknown to the users. Therefore, we only need to define the type of kernel function instead of the activation function and node number of the hidden layer. The node number of the input layer and output layer are equal to the dimension of the feature space and the number of classes respectively.

3.2.3. Classification system and feature space combination

Based on the land use and land cover distribution in the Chao Lake Basin, the classification system is divided into five different land cover types, including *water, algae bloom, forestland, cultivated land, construction land and bare land* and one aquatic vegetation, *algae bloom*. Thus, the total number of classification type items is six (See Table 3). By analyzing the spectral features of each land cover type in the remote sensing images, the dataset of training and testing samples suitable for the Lake Basin were used for classification analysis.

In order to evaluate the precision of K-ELM method in RS imagery classification, the classification result of a partial Landsat image was compared with the results calculated by maximum likelihood classification (MLC) method, support vector machine (SVM) and standard ELM method. To compare the precision of different algorithm, the accuracy evaluation was performed in terms of overall accuracy (OA), kappa coefficient, producer's accuracy (PA) and user's accuracy (UA).

For the precision comparison experiment, a total of 1130 pixels were selected as a dataset for the whole six land use types. The land cover types on high-resolution imagery from Google Earth of similar period were taken as reference for the sample selection. The training sample pixels were randomly selected from the dataset in proportion by pixels of each LUCC types, and rest pixels in the dataset were regarded as testing data sample.

Subsequent spatiotemporal analysis was conducted using images from 1995 to 2018, a sample from 2015 by unit of pixel with a total number of 1600 for six classes was selected as a sample dataset. Separated index between each two classes was greater than 1.9 in ENVI 5.3. In the process of formal classification experiments, the training sample was stratified sampled from the whole sample dataset weighted by the proportion of each class in dataset, and the rest sample in dataset was automatically settled as test sample. The sample allocation process was randomly conducted in MATLAB.

Table 3
Land use and land cover classification system with Landsat data in this paper (Anderson, 1976).

Land use/cover Types	Description
Construction land	Urban or built-up land, including residential, commercial and services, industrial, transportation, roads, mixed urban, and other urban land area
Cultivated land	Agricultural area, including cropland and pasture, dike paddy folder, groves, vegetable lands and fallow lands
Forest land	Deciduous forest, mixed forest lands, evergreen forest and other forest
Bare land	Exposed soils, landfill sites, and areas of active excavation
Water body	Rivers, canals, permanent open water, lakes, ponds, and reservoirs
Aquatic vegetation	Description
Algae bloom	Algae bloom on the water body surface, including lakes, reservoirs and rivers

3.3. Land use and land cover change monitoring

According to the statistical results, a quantitative change for each land cover type was calculated by the area variation between every two images from neighbor period. The index called LUCC dynamic degree D was developed to measure the rate of changes, which can lead to a better description of the quantitative change of each period (Lin et al., 2018). And the expression is given by:

$$D = \frac{U_2 - U_1}{U_1} \times \frac{1}{T_2 - T_1} \times 100\% \tag{14}$$

where U_1 and U_2 are the area of a given class in the year of T_1 and T_2 respectively. The dynamic degree can intuitively reflect the increase and decrease for each land cover type in the study period. And the dynamic spatial–temporal changes are analyzed according to the result of each year's LUCC dynamic degree. The whole workflow of this study is shown in Fig. 2.

4. Results and discussion

4.1. Algorithm comparison and precision evaluation

To examine the accuracy of the K-ELM method, a comparison experiment with other three methods: Maximum Likelihood Classification (MLC), SVM, and original ELM was generated with the image from

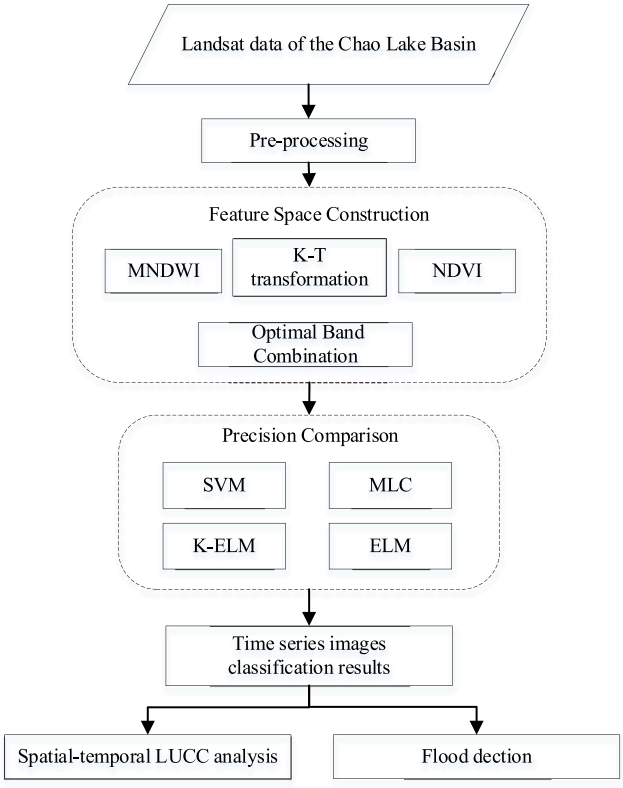


Fig. 2. Workflow of method comparison and spatial–temporal change analysis.

2015. The classification results based on 2015 Landsat image for four different classification methods are shown in Fig. 3. Each land cover type is represented by an individual color. Table 4 outlines the accuracy and index comparison of each classification method.

According to the classification results, images illustrated in Fig. 3, standard ELM illustrated in Fig. 3(c) did not perform so well on distinguishing water and algae bloom. There are also quite a few cultivated land pixels misclassified to forest land by ELM. We can find that K-ELM method could detect algae bloom more accurately, while in the results of MLC some of the algae pixels were misclassified to cultivated land. And a fraction of paddy pixels was misclassified to bare land in the classification result of MLC, see Fig. 3(a).

Of all the methods, K-ELM method outperforms the other four methods with an overall accuracy of 95.19%. SVM follows with an overall accuracy of 93.29%, while standard ELM has the lowest accuracy of 86.67%. From the accuracy comparison result we can draw a conclusion that the SVM method performs better than standard ELM and MLC, but cannot reach the accuracy produced by K-ELM method. Obviously, between standard ELM and K-ELM, standard ELM fails to distinguish forest land and cultivated land as accurately as K-ELM. Furthermore, K-ELM also performs better on classifying algae bloom and water than standard ELM both from the classification results figures and evaluation index. By using K-ELM could increases the classification accuracy significantly, especially on distinguishing algae bloom from water and classifying forest land and cultivated land.

Overall, K-ELM is the most beneficial of these classification algorithms when applied to Chao Lake Basin landcover classification. Hence, in this study, it was used to perform the long-term monitoring of land use spatial pattern change.

4.2. Spatiotemporal analysis

4.2.1. Land use pattern monitoring from 1995 to 2018

This section sets out with a view to perform through spatiotemporal analysis of land use spatial pattern changes in the Chao Lake Basin. Using K-ELM method, a total of seven images were analyzed for land cover variation for the duration of 1995 to 2018. Fig. 4 illustrates the temporal classification results.

The cultivated land has grown from 1995 to 2006 with the highest area of 4441.6 km^2 which covers more than 60% of the total area in 2006. After that, a gradual fall started until 2018, with an occupation rate of 45%. During the growing season, the forest land area was great affected, hitting a minimum area of 484.7 km^2 in 2009. From 2009 onwards, the forest region started to steadily regenerate and grow on the basis of the original site of the forest. As far as land size is concerned, the forest area has recovered to the extent of more than 1300 km^2 as the area in 1995, accounting for nearly 20% of the total area by 2018.

The area remained secure for the development of land and bare land from 1995 to 2006. From 1995 to 2001, bare land was primarily situated in the boundary region between the cultivated land and the forest land., reflecting the phase of encroachment of cultivated land on forested land caused by human activities. In 2006, bare land emerged on the periphery of the metropolitan city and the development of land started to increase and the urban area began to develop steadily. Construction land area has typically seen a growing trend over the last 20 years and has peaked in 2015 with an area of 1022.8 km^2. There was also a small decline in trends from 2009 to 2011 and from 2015 to 2018 related to the introduction of the urban land reclamation program. Much of the bare land has been restored to cultivated land or built for building land in decades to come. Fig. 4(e) also indicates that the bare land is scarcely dispersed in urban and rural areas in the majority of cases and within large parcels of agricultural land. Evidently, the findings suggest that the spatial distribution of bare land was primarily affected by economic development and changes in the status of building and cultivated lang.

The findings of the classification show in Fig. 4(b) of 2001 and Fig. 4 (c) of 2006 show that there has been a limited change in the water area from September 1995 (858.2 km^2) to a maximum of 981.4 km^2 in October 2018. This increase is attributed to the construction of a new

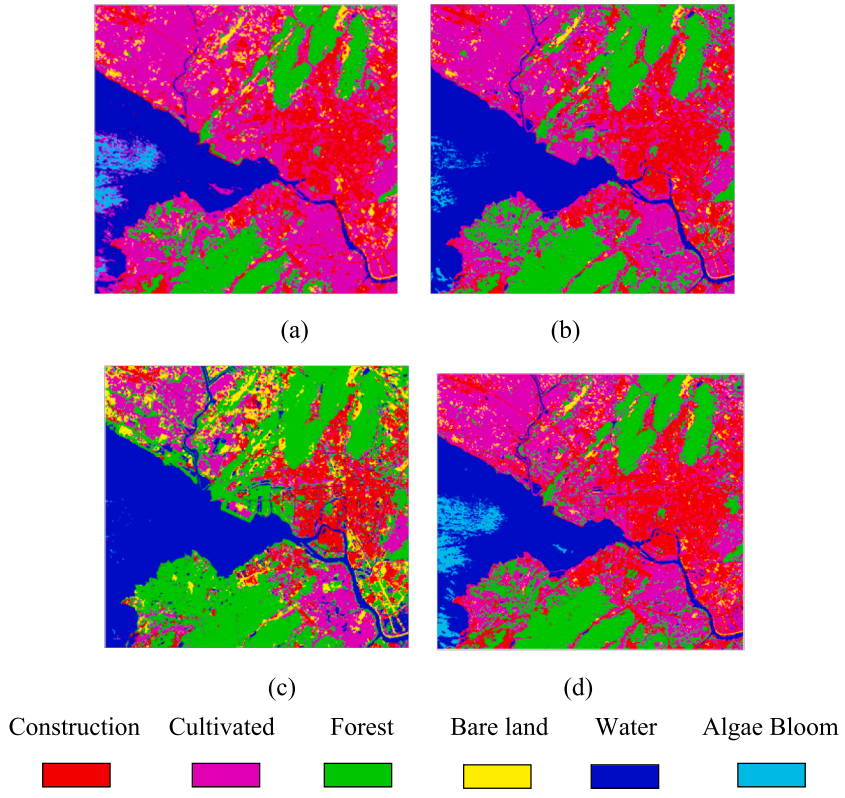


Fig. 3. Classification results of four different methods: (a) MLC (b) SVM (c) ELM (d) K-ELM.

Table 4
Accuracy and index comparison of each classification method.

Methods	Index (%)	Land use/cover types						Precision Index		Time (s)
		Water	Forest	Cultivated land	Bare land	Construction	Algae	OA(%)	Kappa	
MLC	PA	96.00	92.36	90.57	96.55	94.40	73.81	91.33	0.8955	3.44
	UA	99.65	99.77	78.14	85.02	95.01	98.41			
ELM	PA	96.19	90.98	83.91	56.25	95.56	61.54	86.67	0.8423	2.24
	UA	95.28	98.23	74.49	81.82	87.76	66.67			
SVM	PA	99.02	99.11	85.86	60.00	95.65	84.62	93.29	0.9146	6.15
	UA	96.19	98.23	86.73	81.82	89.80	91.66			
K-ELM	PA	99.78	100	86.64	95.4	94.4	96.13	95.19	0.9422	9.37
	UA	97.19	93.45	95.25	93.26	93.79	99.69			

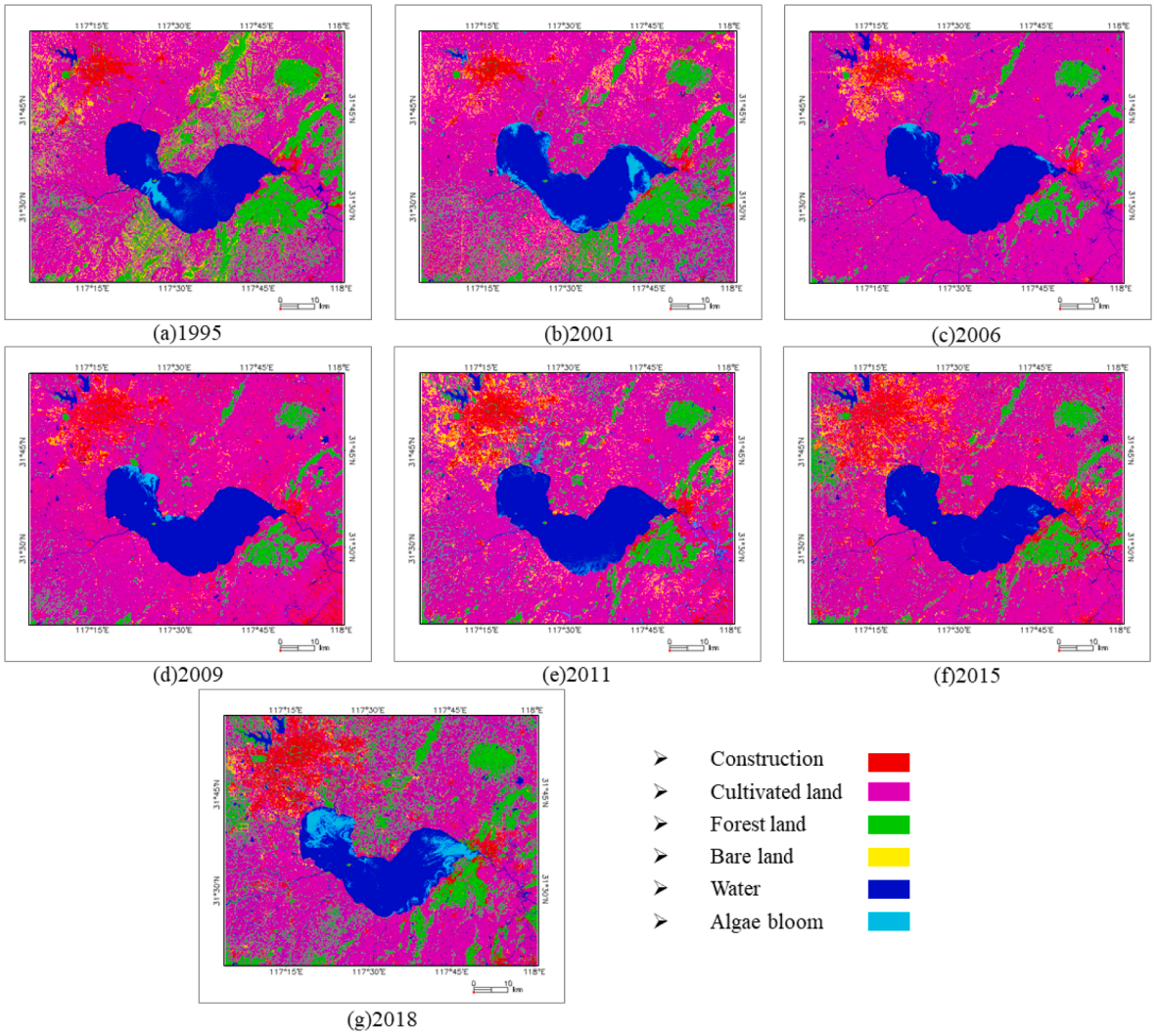


Fig. 4. Classification results of K-ELM from 1995 to 2018 in the Chao Lake Basin.

reservoir developed in the northwest of Hefei City between the years 2001 and 2006. The construction of Dafangying Reservoir began in December 2001, and the closing of the river was finished in October 2002. Having met he intermediate approval at the end of 2003 and closed the impoundment gate. In 2004, Dafangying Reservoir and Dongpu Reservoir, both in the northwest of Hefei District were connected together. Both of the large-scale reservoirs rely on flood control and a mix of urban water supply, which contributes a considerable amount to the hydrological ecological balance of urbanized areas in the Chao Lake Basin.

4.2.2. Spatial-temporal change statistic and analysis

This section details the statistical analysis to derive the degree of variance of each land use type. Table 5 and 6 present statistical insights for land cover type for the period of 1995 to 2018 in the Chao Lake Basin. Fig. 5 illustrates a bar graph that shows comparisons between various land cover types. The complex degree of change (Table 6) of spatiotemporal land use encompasses trends estimated from area statistics and their effect is visible over time.

First of all, from the area and change statistics (Tables 5 and 6) and graph (Figs. 5 and 6), the area of agricultural land increased from 1995 to 2006, with a volatile rate of 2.07% from 1995 to 2001 and 3.84%

Table 5
Area statistics for land cover types from 1995 to 2018 in the Chao Lake Basin.

Year	Area (km^2)					
	Forest	Cultivated land	Construction land	Bare land	Algae	Water (Algae included)
1995/09	1322.8	3315.5	558.3	425.2	45.4	858.2
2001/07	1108.8	3726.9	408.5	354.9	150	880.9
2006/07	487.1	4441.6	522.4	87.1	32	941.7
2009/06	484.7	4086.4	748.9	226.9	41.2	933.1
2011/04	693.5	3733	692.4	411.8	88.1	949.3
2015/10	742.3	3525.4	1022.8	286.3	27.8	903.3
2018/10	1383.8	2978.4	932.6	203.8	155.4	981.4

from 2001 to 2006 separately. It started to decline after 2006. Meanwhile, the pattern of forest variation in the region is in contrast to that of agricultural land. The forest area had a diminishing pattern from 1995 to 2006 and remained largely unchanged from 2006 to 2009. The trend is growing from 2009 to 2018, with a volatile pace of 21.54% from 2009 to 2011, 1.76% from 2011 to 2015 and 28.81% from 2015 to 2018, respectively. The variation outcomes of agricultural land and forest land in each increasing and declining cycle are shown in Figs. 7 and 8. Changes also occurred in the entire basin, especially throughout the wilderness zone or the moderately wild land cover in the north and south.

As far as construction land area is concerned, there is a general variance in the pattern of rising over the entire study duration from 1995 to 2018. As shown in classification results (Fig. 4), the construction land is the major component of rural land scape, township and urban area. Firstly, the rural landscape mainly bases on the extensive cultivated land with fragments of artificial construction and forests. In rural landscape, there are also scattered residential zones performed as pixels of construction land in land use and land cover types. Secondly, the township region includes mid-sized counties and towns, which are presented as small-scale construction land area scattered along traffic rail and roads on classification result figures. Finally, urban regions manifest as large polygons of construction land type and fractions of natural environment elements such as cultivated land, forest or water body in the background. The urban area mainly consists of downtown the Hefei City in northwest and the Chaohu city in east to the Chao Lake.

LUCC statistics (Tables 5 and 6) indicate that construction land area has risen by 67.04% since the 1990 s, reflecting a major shift in the rate of rural settlements and dramatic urban growth pattern over the last 20 years. As shown in Fig. 6(a), the steady increase occurred predominantly in the Hefei City and Chaohu City regions situated in the northwest and east to the Chao Lake separately. Fig. 6(a) and (b) indicate that the growth of construction land and the reduction of cultivated land are simultaneous in the Hefei City and Chaohu City areas, which means that, since the expansion of the city, the adjacent agricultural land has been encroached. The gradual increase in construction land is also evident from the transport networks linking major cities and towns such Wuhu-Hefei Expressway in 2000 and the high-speed rail line passing via Hefei City-Chaohu City-Lujiang County. The drastic scale of expansion is consistent with the rapid development and urbanization of China (Li et al., 2018; Zhang et al., 2018), especially for the Yangtze River Delta Region.

In particular, as an evidence of the shifts in the land use trends, the bare land area reveals a disparity in observable variations. From 1995 to 2006, there was a steady decrease of 2.76% and 15.09% of bare land respectively at the same period with the intermittent shifts in cultivated land, forest land and construction land along the Chao Lake Basin.

Based on the above analyses it can be inferred that the land use and land cover trends in the Chao Lake Basin have undergone drastic variation for over 20 years. Land use trend variation was driven by policy and interacted with each other. It should be noted, for example, that the findings show that the rise in construction land had a substantial effect on agricultural land to a level of −10.17%. This forest cover is located mostly in and near the urban region and waterfront areas around the lake, such as the southwest of Hefei District, the north, and the east coast of Chao Lake in Chaohu City. This can be attributed to governing policy

Table 6
Dynamic degree of different land cover from 1995 to 2018.

Periods	Dynamic degree (%)					
	Forest	Cultivated land	Construction	Bare land	Algae	Water
1995–2001	−2.70	2.07	−4.47	−2.76	38.40	0.44
2001–2006	−11.21	3.84	5.58	−15.09	−15.73	1.38
2006–2009	−0.16	−2.67	14.45	53.50	9.58	−0.30
2009–2011	21.5	−4.32	−3.77	40.74	56.92	0.87
2011–2015	1.76	−1.39	11.93	−7.62	−17.11	−1.21
2015–2018	28.81	−5.17	−2.94	−9.61	153.00	2.88
1995–2018	4.61	−10.17	67.04	−52.07	/	/

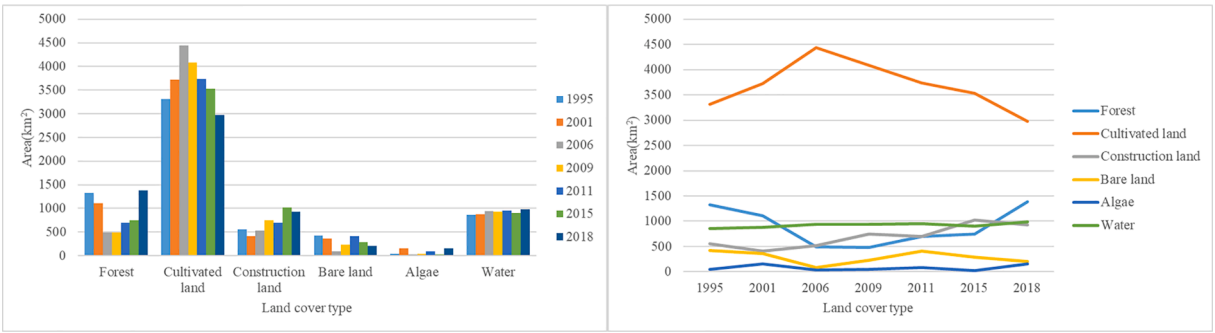


Fig. 5. Land use/cover area statistics. (a) Bar graph for each land cover area (b) Multi-temporal LUCC line chart.

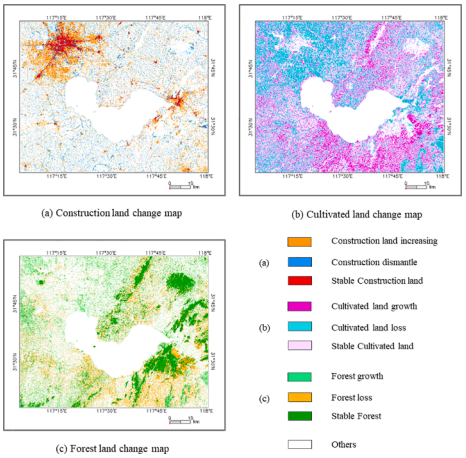


Fig. 6. Construction land, cultivated land and Forest land cover change between 1995 and 2018.

directive to earmark this region for ecological tourism. In contrast to this, as a result of the Grain for Green Project Policy has led to a significant decrease in cultivated land from 2006. This trend is due to a rise in the return of agriculture land to forest between 2006 and 2018. The aforementioned analysis outlines the significant impact due to human activity on the ecological space. It also reveals that data-driven decision making is of interest to biodiversity at large. One such example, is the validation of Grain for Green Project, where reversing urbanization practices by converting cultivable land into forest increase the forest land cover footprint area, thereby improving the phenological attributes of the ecosystem.

In addition to this, land-cover shifts over the years have led to a rapid increase in algae contamination in the Chao Lake over the period of more than 20 years. The largest algae bloom was recorded in 2018, account for 15% of the overall body of water, admeasuring an area of 155.4 km^2. The results validate the correlation between human activities and its impact on the ecology.

4.2.3. Algae bloom outbreaks on water surface

From the above long-term sequence image classification results and data statistics (as Fig. 4 and Table 5), the continuous outbreak of Algae bloom in Chaohu Lake could be detected and analyzed. Over the years, the water body of the Chao Lake has been in a severe stage of eutrophication for decades, and algae bloom has occurred several times at set intervals (Qiu et al., 2015). From 1995 to 2018, the classification findings conclusively identified the geographical distribution and specific area of each algae bloom outbreak. For the year 1995, the algae bloom was 5% of the total water area. For the subsequent year 1996 and 1997 it remained constant, and gradually increased by 6% as compared to 1995. In 2001 the algae bloom had an outbreak mainly on eastern lake surface and the southwest shoreline along the polder field (Fig. 4(b)). Following from 2006 to 2011, there were small-scale algae blooms along the northwest shore close to urban region and south shore close to rural area. For example, algae bloom is found on the south surface of lake surface with sparce and wide distribution and also on the surface of reservoirs and river water bodies in the lakeside of Hefei City in 2011, see Fig. 4(e). In 2015, algae bloom sporadically distributed on the lake surface. By the year of 2018, algae bloom accounted for 15.8% of the total water area admeasuring 155.4 km^2, which is concerned with urbanization and intense human activities reflected by rapid construction land expansion and frequent cultivated land change in Fig. 6. This constant increase in the algae bloom is attributed to rapid urbanization and the pollution of water sources due to Nonpoint source pollution.

4.2.4. Regional flood hazard detection

In the summer of 2020, the area around the Chao Lake was hit by a once-in-a-century flood disaster. The flat terrain, complex river network and uneven spatio-temporal distribution of precipitation has the potential to create floods across the Chao Lake Basin impacting the ecosystem at large. With the catchment area being a sub-tropical monsoon climate zone with high seasonal variability of precipitation. The likelihood occurrence of flood is inevitable, especially during summer or pre-harvest. These floods tend to have a significant impact on the crop yields. It is of interest for us to estimate the regional area where flooding occurred in July 2020 surrounding the Chao Lake. In this section, Landsat OLI images were used to conduct flood detection by post-classification method based on k-ELM and compare between normal seasonal time frame on June, 02, 2020 and the flooded temporal frame on July, 20, 2020.

The analytical insights drawn from the studies outline that there has been a significant increase in the water area by 458.6445 km^2 compared to its original state. This 50.79% increase in the water area is identified by land submerged due to flooding. Through the flood spatial distribution in Fig. 9, the flooded area mainly lies in Lujiang County and Shucheng Count to the southwest of Chao Lake Basin.

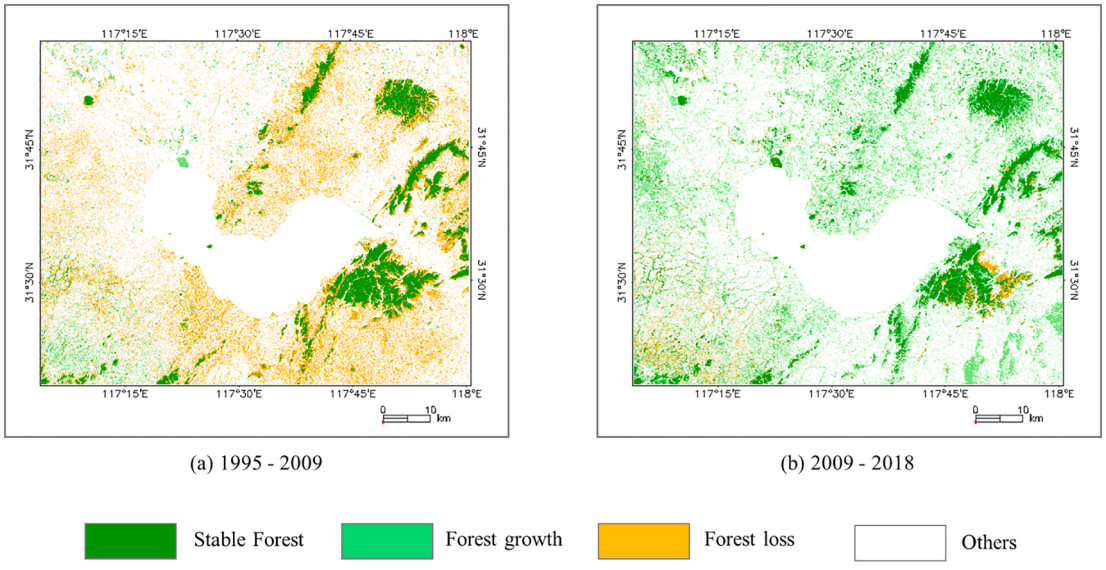


Fig. 7. Forest land cover change between 1995 and 2009, 2009 and 2018.

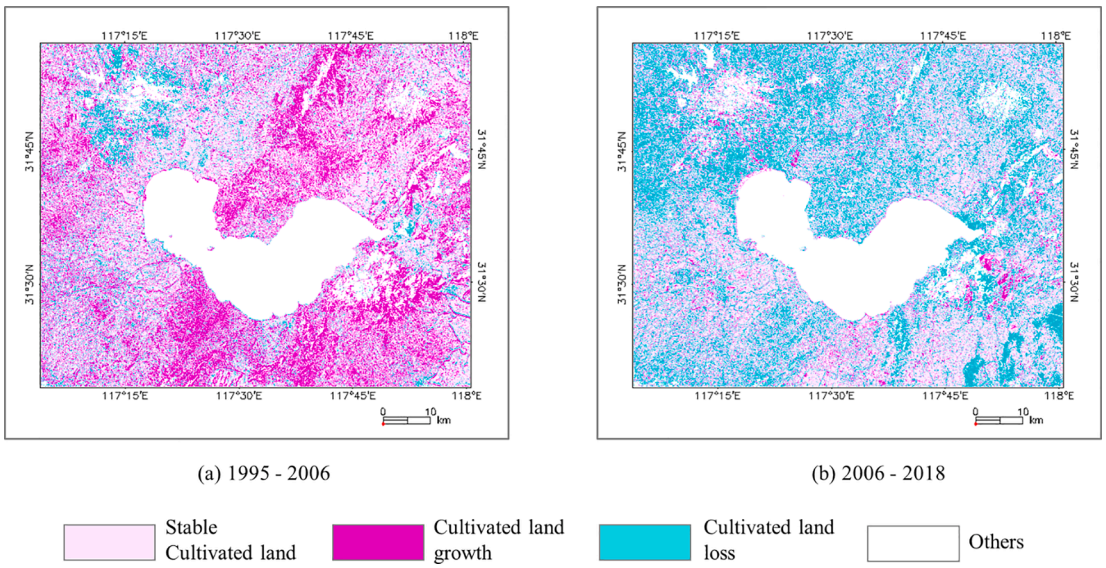


Fig. 8. Cultivated land cover change between 1995 and 2006, 2006 and 2018.

For further analyzing, Table 7 outlines flooded area statistic of each land cover type and the percentage of the cultivable land impacted. This would help us in predicting socio-economic loss incurred to these floods. In terms of affected area, it is to be noted that cultivable and construction land had significant impact to local residents. The affected area of cultivated land was 317.5 km^2, accounting for 69.2% of total flood area and 10.5% of original cultivated land area. The affected area of construction land was 107.5 km^2, making up 23.4% of total flood area and 7.2% of original area. According to percentage of original area, among the various land cover types, cultivated land and bare land suffered more from flood hazard compared with buildings and forests.

As the images were taken during heavy rainfall, a fraction of error is attributed to the climatic conditions such as dense clouds in the southeast during which the images were captured. But we strongly believe that still could be used a general frame of reference for comparative analysis.

Table 7
Flood area statistics of each land cover type and the percentage of the whole flood area and its original area.

Original land cover type	Forest	Cultivated land	Construction land	Bare land
Flood affected area (km^2)	18.2655	317.502	107.4861	4.1058
Original land cover area (km^2)	974.2	3012.1	1488.2	34
Percentage of flood (%)	3.98	69.23	23.44	0.90
Percentage of original area (%)	1.87	10.54	7.22	12.08

5. Conclusions

In this paper, we used the K-ELM algorithm to derive six type of land cover patterns for the duration of 1995 to 2018 using long term Landsat images in the Chao Lake Basic and provided crucial insights into the

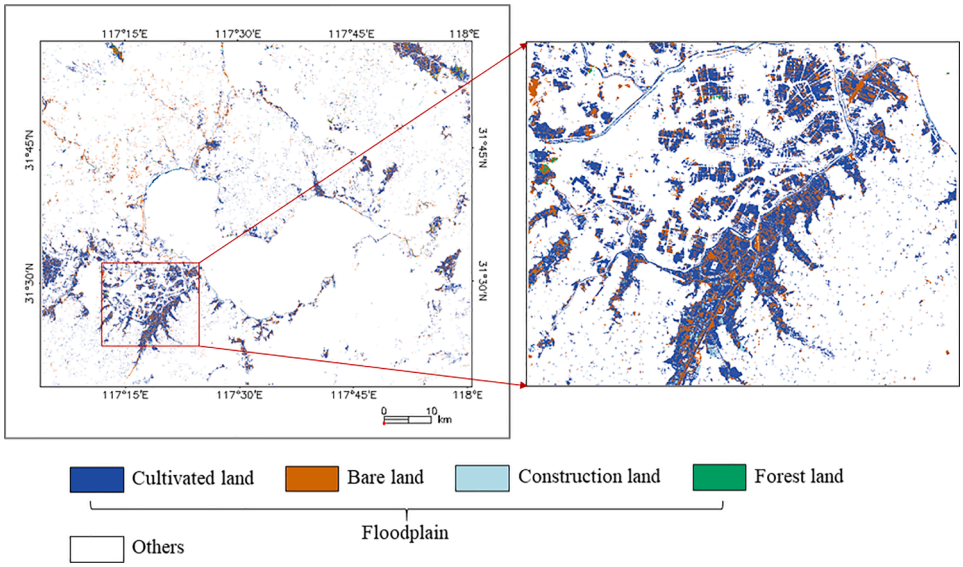


Fig. 9. Flood-affected area and local zoom from water body change detection between June 2 and July 20.

detection of upstream lake basin conservation areas estimated about 6480 km^2. The spatio-temporal change analysis was carried out with dynamic degree and change statistics based on the post classification change monitoring results, which could help us enforce regulatory measures to prevent algae bloom outbreaks and flood disaster. A thorough review of the results allowed the following conclusions to be drawn:

1) The extreme learning machine with kernel function was successfully applied to classify the land cover types in the Chao Lake Basin from 1995 to 2018. The accuracy of K-ELM system was first tested and compared with three other conventional classification systems. The results show that K-ELM was able to attain the highest average precision at 95.19%, which demonstrates the accuracy and stability of the K-ELM method and the reliability of follow-up multi-temporal classification results.
2) Using the land use data based on classification findings, we have successfully evaluated land use spatial pattern trends for Chao Lake Basin from 1995 to 2018. It is to be noted that the lake and river system and their surrounding basin land cover are interconnected together as a complete ecosystem in the watershed catchment area. The land use and land cover pattern have significant variation for the following categories: *cultivated land, forest land* and *constructed land* with dynamic degree of change of −10.17%, 4.61, 67.04% respectively. Along with economic and population growth, there is an increasing trend towards urbanization in the Chao Lake Basin. Some of the major ecological concerns identified during the study period were the ongoing algae bloom outbreaks with total water area proportion of 15% by 2018.
3) Furthermore, this study has clearly outlined flood affected areas through change detection results based on Landsat OLI images acquired before and during the 2020 flood of Chao Lake Basin. The flood disaster in July 2020 has a significant impact on the ecology of the affected region with a submerged area of 458.6445 km^2, and most of the affected areas were farmland and village buildings. The flood detection results would help to predict and evaluate the socio-economic impact and ecological degradation on the Chao Lake Basin caused by floods over the next years.

One of the main shortcomings of this research is its inability to conduct comprehensive land cover analysis to classify vegetation type or industrial development regions due to limited spatial resolution to the Landsat results (30 m spatial resolution). In order to identify and track phenological changes, we suggest the use of high resolution QuickBird or GF Satellite data for localized region where a thorough analysis is to be carried out. Investigation of algae bloom causes should go in-depth combined with local pollution emission statistics and chemical analysis in the future at the level of watershed, for example, the quantitively correlation among metrological factor, non-point source pollution, and residential or industrial emission and algae outbreaks in the Chao Lake Basin. The authors also suggest that further researches should use multi-source data, such as weather stations, field surveys, satellite images, hydrological data, and in-situ geospatial environmental sensor data, to perform a comprehensive study to assess the impact of human activity on the ecosystem.

CRediT authorship contribution statement

Yi Lin: Project administration, Conceptualization, Supervision, Formal analysis. **Tinghui Zhang:** Methodology, Data curation, Writing - original draft, Software, Visualization. **Qin Ye:** Validation, Methodology. **Jianqing Cai:** Methodology, Formal analysis. **Chengzhao Wu:** Resources, Conceptualization. **Awase Khirni Syed:** Writing - review & editing, Formal analysis. **Jonathan Li:** Validation, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The authors would like to thank editor and reviewers for their insightful comments improving this paper. This work was supported by the National Natural Science Foundation of China (Project No. 41771449) and the DAAD Thematic Network (Project No. 57421148).

References:

Aghsaei, H., Mobarghaee Dinan, N., Moridi, A., Asadolahi, Z., Delavar, M., Fohrer, N., Wagner, P.D., 2020. Effects of dynamic land use/land cover change on water resources and sediment yield in the Anzali wetland catchment, Gilan. Iran. Sci. Total Environ. 712, 136449 https://doi.org/10.1016/j.scitotenv.2019.136449.

Anderson, J.R., 1976. A land use and land cover classification system for use with remote sensor data. US Government Printing Office.

Cai, Y., Guan, K., Peng, J., Wang, S., Seifert, C.A., Wardlow, B.D., Li, Z., 2018. A high-performance and in-season classification system of field-level crop types using time-series Landsat data and a machine learning approach. Remote Sens. Environ. 210, 35–47.

De Wit, M., Stankiewicz, J., 2006. Changes in surface water supply across Africa with predicted climate change. Science (80-.) 311, 1917–1921. https://doi.org/10.1126/science.1119929.

Dörnhöfer, K., Oppelt, N., 2016. Remote sensing for lake research and monitoring - Recent advances. Ecol. Indic. 64, 105–122. https://doi.org/10.1016/j.ecolind.2015.12.009.

Du, P., Bai, X., Tan, K., Xue, Z., Samat, A., Xia, J., Li, E., Su, H., Liu, W., 2020. Advances of Four Machine Learning Methods for Spatial Data Handling: a Review. J. Geovisualization Spat. Anal. 4 https://doi.org/10.1007/s41651-020-00048-5.

Duan, Z., Bastiaanssen, W.G.M., 2013. Estimating water volume variations in lakes and reservoirs from four operational satellite altimetry databases and satellite imagery data. Remote Sens. Environ. 134, 403–416. https://doi.org/10.1016/j.rse.2013.03.010.

Dube, T., Gumindoga, W., Chawira, M., 2014. Detection of land cover changes around Lake Mutirikwi, Zimbabwe, based on traditional remote sensing image classification techniques. African J. Aquat. Sci. 39, 89–95.

Duro, D.C., Franklin, S.E., Dube, M.G., 2012. A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. Remote Sens. Environ. 118, 259–272.

Dwivedi, R.S., Rao, B.R.M., 1992. The selection of the best possible Landsat TM band combination for delineating salt-affected soils. Int. J. Remote Sens. 13, 2051–2058.

Guo, H., Hu, Q., Jiang, T., 2008. Annual and seasonal streamflow responses to climate and land-cover changes in the Poyang Lake basin. China. J. Hydrol. 355, 106–122.

Huang, G. -B., Zhu, Q.-Y., Siew, C.-K., 2004. Extreme learning machine: A new learning scheme of feedforward neural networks, in: Proceedings of IEEE International Conference on Neural Networks, pp. 985–990. https://doi.org/10.1109/IJCNN.2004.1380068.

Han, X., Chen, X., Feng, L., 2015. Four decades of winter wetland changes in Poyang Lake based on Landsat observations between 1973 and 2013. Remote Sens. Environ. 156, 426–437. https://doi.org/10.1016/j.rse.2014.10.003.

Huang, G.-B., Zhu, Q.-Y., Siew, C.-K., 2006. Extreme learning machine: theory and applications. Neurocomputing 70, 489–501.

Jorgenson, M.T., Grosse, G., 2016. Remote sensing of landscape change in permafrost regions. Permafr. Periglac. Process. 27, 324–338.

Karan, S.K., Samadder, S.R., 2016. Accuracy of land use change detection using support vector machine and maximum likelihood techniques for open-cast coal mining areas. Environ. Monit. Assess. 188, 486. https://doi.org/10.1007/s10661-016-5494-x.

Khatami, R., Mountrakis, G., Stehman, S.V., 2016. A meta-analysis of remote sensing research on supervised pixel-based land-cover image classification processes: General guidelines for practitioners and future research. Remote Sens. Environ. 177, 89–100.

Kutser, T., Metsamaa, L., Strömbeck, N., Vahtmäe, E., 2006. Monitoring cyanobacterial blooms by satellite remote sensing. Estuar. Coast. Shelf Sci. 67, 303–312. https://doi.org/10.1016/j.ecss.2005.11.024.

Li, G., Sun, S., Fang, C., 2018. The varying driving forces of urban expansion in China: Insights from a spatial-temporal analysis. Landsc. Urban Plan. 174, 63–77. https://doi.org/10.1016/j.landurbplan.2018.03.004.

Lin, Y., Yu, J., Cai, J., Sneeuw, N., Li, F., 2018. Spatio-temporal analysis ofwetland changes using a kernel extreme learning machine approach. Remote Sens. 10, 1–15. https://doi.org/10.3390/rs10071129.

Liu, Y., Huang, X., Yang, H., Zhong, T., 2014. Environmental effects of land-use/cover change caused by urbanization and policies in Southwest China Karst area - A case study of Guiyang. Habitat Int. 44, 339–348. https://doi.org/10.1016/j.habitatint.2014.07.009.

Luo, L., Wang, X., Liu, J., Guo, H., Zong, X., Ji, W., Cao, H., 2017. VHR GeoEye-1 imagery reveals an ancient water landscape at the Longcheng site, northern Chaohu Lake Basin (China). Int. J. Digit. Earth 10, 139–154.

McIver, D.K., Friedl, M.A., 2002. Using prior probabilities in decision-tree classification of remotely sensed data. Remote Sens. Environ. 81, 253–261. https://doi.org/10.1016/S0034-4257(02)00003-2.

MEEPPC, 2002. Environmental Quality Standards for Surface Water. URL http://www.mee.gov.cn/ywgz/fgbz/bz/bzwb/shjbh/shjzlbz/200206/t20020601_66497.shtml.

Nsubuga, F.W.N., Botai, J.O., Olwoch, J.M., Rautenbach, C.J.D., Kalumba, A.M., Tsela, P., Adeola, A.M., Sentongo, A.A., Mearns, K.F., 2017. Detecting changes in surface water area of Lake Kyoga sub-basin using remotely sensed imagery in a changing climate. Theor. Appl. Climatol. 127, 327–337.

Palmer, S.C.J., Kutser, T., Hunter, P.D., 2015. Remote sensing of inland waters: Challenges, progress and future directions. Remote Sens. Environ. 157, 1–8. https://doi.org/10.1016/j.rse.2014.09.021.

Qiu, L., Dijk, M.P.V., Wang, H., 2015. Water pollution and environmental governance of the Tai and Chao Lake Basins in China in an international perspective. J. Water Resour. Prot. 07, 830–842. https://doi.org/10.4236/jwarp.2015.710067.

Rao, C.R., Mitra, S.K., 1972. Generalized inverse of a matrix and its applications, in: Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability. Theory of Statistics vol. 1.

Rodriguezgaliano, V.F., Chicaolmo, M., Abarcahernandez, F., Atkinson, P.M., Jeganathan, C., 2012. Random Forest classification of Mediterranean land cover using multi-seasonal imagery and multi-seasonal texture. Remote Sens. Environ. 121, 93–107.

Stow, D.A., Hope, A., McGuire, D., Verbyla, D., Gamon, J., Huemmrich, F., Houston, S., Racine, C., Sturm, M., Tape, K., 2004. Remote sensing of vegetation and land-cover change in Arctic Tundra Ecosystems. Remote Sens. Environ. 89, 281–308.

Sugiyama, M., 2015. Introduction to statistical machine learning. Morgan Kaufmann.

Tong, X., Pan, H., Xie, H., Xu, X., Li, F., Chen, L., Luo, X., Liu, S., Chen, P., Jin, Y., 2016. Estimating water volume variations in Lake Victoria over the past 22 years using multi-mission altimetry and remotely sensed images. Remote Sens. Environ. 187, 400–413. https://doi.org/10.1016/j.rse.2016.10.012.

Wang, M., Shi, W., 2008. Satellite-observed algae blooms in China's Lake Taihu. EOS Trans. 89 (22), 201–202. https://doi.org/10.1029/2008EO220001.

Wasige, J.E., Groen, T.A., Smaling, E., Jetten, V., 2013. Monitoring basin-scale land cover changes in Kagera Basin of Lake Victoria using ancillary data and remote sensing. Int. J. Appl. Earth Obs. Geoinf. 21, 32–42.

Were, K.O., Dick, T.B., Singh, B.R., 2013. Remotely sensing the spatial and temporal land cover changes in Eastern Mau forest reserve and Lake Nakuru drainage basin. Kenya. Appl. Geogr. 41, 75–86. https://doi.org/10.1016/j.apgeog.2013.03.017.

Wolter, P.T., Johnston, C.A., Niemi, G.J., 2006. Land use land cover change in the US Great Lakes basin 1992 to 2001. J. Great Lakes Res. 32, 607–628.

Wu, J., Luo, J., Tang, L., 2019. Coupling relationship between urban expansion and lake change-A case study of Wuhan. Water (Switzerland) 11. https://doi.org/10.3390/w11061215.

Zhang, Z., Liu, F., Zhao, X., Wang, X., Shi, L., Xu, J., Yu, S., Wen, Q., Zuo, L., Yi, L., Hu, S., Liu, B., 2018. Urban expansion in China Based on remote sensing technology: A review. Chinese Geogr. Sci. 28, 727–743. https://doi.org/10.1007/s11769-018-0988-9.

Zhao, Y., Zhang, K., Fu, Y., Zhang, H., 2012. Examining land-use/land-cover change in the lake dianchi watershed of the Yunnan-Guizhou plateau of Southwest China with remote sensing and GIS techniques: 1974–2008. Int. J. Environ. Res. Public Health 9, 3843–3865. https://doi.org/10.3390/ijerph9113843.