Land Cover Change in Nanjing from 1993 to 2018

Pei-shan SHEN\textsuperscript{1,*} and Jian GAO\textsuperscript{2}

\textsuperscript{1}College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing, China

\textsuperscript{2}School of Geographic and Biologic Information, Nanjing University of Posts and Telecommunications, Nanjing, China

*Corresponding author

**Keywords:** land cover change, Google Earth Engine, CART.

**Abstract.** Land cover change is a significant indicator of regional and global environmental change, and it even has an important significance for sustainable land management. The objective of this study was analysis on the land cover change in Nanjing during the period of 1993–2018. The sample data were selected under a stratified sampling scheme. The training features were composed based on Tasseled Cap transformation, enhanced vegetation index (EVI) and normalized difference water index (NDVI). The Classification and Regression Trees (CART) method was used to obtain the land cover maps of Nanjing in 1993, 1998, 2003, 2008, 2013 and 2018. We implemented the post-classification change detection method to analyze the land cover change in Nanjing. The image composite and classification of long-term remote sensing data sequence were implemented with Google Earth Engine (GEE) platform. Results showed that: (1) the vegetation covered areas of Nanjing decreased from 79.47% in 1993 to 54.89% in 2018. The urban or exurban areas of Nanjing increased from 10.94% in 1993 to 33.92% in 2018. (2) The main feature of land cover in Nanjing was the expansion of urban and exurban, which mainly happened in the central and southern parts of Nanjing.

**Introduction**

Land cover change is defined as the loss of natural areas, particularly the loss of forests to urban or exurban development, or the loss of agricultural areas to urban or exurban development [1]. Land cover change is a significant indicator of regional and global environmental change, and even has an important significance for sustainable land management [2]. The analysis of land cover change in Nanjing can not only adjust the industrial layout but also improve the land use structure.

The common method of land cover change is the post-classification change detection method. With the development of artificial intelligence and machine learning theory, Modern land cover classification methods are mainly implemented on computer automation programs through machine learning and artificial intelligence theory [3]. For example, support vector machine, artificial neural network and decision tree.

In this paper, we mainly focused on the natural or agricultural areas and urban or exurban areas in the five different periods (1993–1998, 1998–2003, 2003–2008, 2008–2013, and 2013–2018) in Nanjing. The Tasseled Cap transformation was used for dimensionality reduction purposes in the long-term remote sensing data sequence. The Classification and Regression Trees (CART) method was applied to obtain the land cover maps of Nanjing in 1993, 1998, 2003, 2008, 2013 and 2018, on the classification feature set combined by some graphical indicators. Finally, we analyzed the land cover change in Nanjing during the period of 1993–2018.
Material and Methodology

Study Area
Nanjing is the capital of Jiangsu Province, with an area of 65.87 million km$^2$ and a population of 833.5 million. The geographical coordinates range from 31°14’ to 32°37’ north latitude and 118°22’ to 119°14’ east longitude. Nanjing is a political, economic, scientific, educational, cultural, and information center in eastern China.

Remote Data
Landsat is an ongoing mission of Earth observation satellites developed under a joint program of the USGS and NASA. Landsat TM and ETM+ have a spatial resolution of 30 m and an update interval of 16 days. In this paper, we imported Tier 1 (T1) Data, which meets geometric and radiometric quality requirements Landsat. To be specific, the numbers of Landsat 5 TM images were 5, 5 and 5 for year 1993, 1998 and 2003 respectively, while the numbers of Landsat 7 ETM+ images were 5, 5 and 6 for year 2008, 2013 and 2018 respectively.

Global Multi-Resolution Terrain Elevation Data 2010 (GMTED2010) is a significantly enhanced global elevation model developed by the US Geological Survey (USGS) and the National Geospatial Intelligence Agency (NGA) [4]. This article used the digital elevation data provided by GMTED2010.

Sample Data
In the absence of high-resolution images and ground records, sample selection by time-series Landsat images is a more general method [5]. In this paper, the study area was classified into water, woodland, cultivated land or grassland, bare land and urban or exurban. These five land cover categories have obvious identification features on Landsat images. In this paper, we adopted a stratified sampling scheme for the sample selection to reduce errors [6]. The specific process was: firstly, we randomly selected sample points from the original image of Nanjing in 2018 provided by the GEE platform. Secondly, we marked the land cover categories of the points by visual recognition and established the 2018 sample data. Thirdly, we loaded the 2018 sample data into the image in 2013 and re-examined the features of sample points. We could refresh the feature type of the sample point, delete and add some points to meet a balance. Finally, we got the 2013 sample data. We repeated this process and obtained sample data for 2018, 2013, 2008, 2003, 1998, and 1993 at last. For classification, the sample data of each year were assigned in a ratio of 4:1 at random, 80% of the samples was used for classifier training, and 20% were used for accuracy evaluation (Table 1).

| Year  | Water | Woodland | Cultivated land or grassland | Bare land | Urban or exurban | Sum   |
|-------|-------|----------|-----------------------------|-----------|------------------|-------|
| 1993  | 147   | 342      | 450                         | 7         | 145              | 1091  |
| 1998  | 146   | 330      | 420                         | 9         | 186              | 1091  |
| 2003  | 145   | 329      | 368                         | 9         | 237              | 1088  |
| 2008  | 145   | 320      | 275                         | 8         | 309              | 1057  |
| 2013  | 145   | 320      | 282                         | 8         | 309              | 1064  |
| 2018  | 145   | 320      | 286                         | 11        | 316              | 1078  |

GEE and Data Pre-Processing
The Google Earth Engine (GEE), the Google Earth Engine, is a cloud-based geographic data processing platform jointly developed by Google, Carnegie Mellon University and the US Geological Survey [7]. The GEE platform provides publicly available remote sensing images and weather and elevation data of the order of PB, optimized algorithms for analyzing geographic data, API interfaces based on JavaScript and Python languages, and visual code editors based on JavaScript API interfaces. Compared with traditional remote sensing image processing systems,
Google Earth Engine has unprecedentedly improved the computational efficiency of geospatial data, which can be used for global scale research, and is also easier for users to learn [8].

By writing code on the GEE platform JavaScript API interface, the cloud score range and maximum depth were selected for the Landsat time series satellite image to obtain the minimum cloud size and least shadow image of the target year. The processed images were tailored according to the Nanjing vector boundary [9].

**Research Method**

**Tasseled Cap Transform**

Tasseled cap transform was put forward by Kauth and Thomas in 1976 [10]. This transform is used widely in analyzing and mapping the vegetation or urban development changes. This transform can also make a dimensionality reduction in a long-term remote sensing data sequence. The first three components after the transformation are named “brightness”, “greenness” and “humidity”, which respectively reflect the moisture information of soil, vegetation and water.

In this paper, the B1, B2, B3, B4, B5 and B7 bands of TM and ETM images in the study area were subjected to the tasseled cap transform. The bands obtained by tasseled cap transform (the brightness, greenness, humidity, fourth component, fifth component and sixth component) had better separability in urban or exurban type.

**Classification Feature**

Through experimental comparison, it was clear that the water and cultivated land or grassland can be classified more sensitively by adding Enhanced Vegetation Index (EVI) and Normalized Water Body Index (NDWI). The combination of brightness, greenness, humidity, fourth component, fifth component, sixth component, EVI, NDWI and the digital elevation data were used as the classification feature set for each target year.

**CART Method**

Classification and Regression Trees method belongs to a decision tree. The Gini coefficient is used to select the optimal segmentation feature, and the training data set is divided into two parts to generate the decision tree model.

In CART, the generated tree needs to be pruned by cross-validation to avoid over-fitting the training number. The final analysis result is an optimal binary tree with both complexity and error rate. The CART method is simple, flexible, and highly accurate, and can process large amounts of data and high-dimensional data quickly and efficiently.

Using the CART method for the classification feature set of each target year, the land cover classification map of Nanjing from 1993 to 2018 was obtained in Figure 1.

(a)1993 (b)1998 (c)2003 (d)2008 (e)2013 (f)2018

Figure 1. Land cover maps of Nanjing City.
Results and Analysis

Classification Accuracy Assessment

The precision of Land cover change detection depended on the classification accuracy assessment in this study. Kappa coefficients are used for consistency testing and can also be used to measure classification accuracy. The overall accuracy (OA) is the ratio of the correct number of tests in a classification. In every classified land cover map, Kappa coefficients and OA were calculated and figured in Figure 2.

![Figure 2. Accuracy assessment of classification results.](image)

Overall, the classification accuracy of each year's classification results was relatively high and stable. The overall accuracy was between 83.5% and 89.32%, the average was 86.99%, the standard deviation was 2.34%; the KAPPA coefficient was between 76.55% and 86.57%, the average was 81.93%, and the standard deviation was 3.44%.

Temporal Trend of Changes

We counted the number of pixels of each category in the land cover map for each target year. The sum number of pixels in the woodland and cultivated land or grassland was an alternative to the number of pixels in the vegetation coverage area. We compared it with the number of pixels in the urban or exurban to find the temporal trend of land cover change in Nanjing during the past 25 years (Table 2).

Table 2. Land cover change of Nanjing city from 1993 to 2018.

| Year | Vegetation coverage area | Percentage/% | Urban or exurban | Percentage/% |
|------|--------------------------|--------------|------------------|--------------|
| 1993 | 5235009534               | 79.47        | 720336188        | 10.94        |
| 1998 | 5096149851               | 77.36        | 817593404        | 12.41        |
| 2003 | 4770433889               | 72.42        | 1121212953       | 17.02        |
| 2008 | 4436109831               | 67.34        | 1488088808       | 22.59        |
| 2013 | 4253466209               | 64.57        | 1682862360       | 25.55        |
| 2018 | 3615864991               | 54.89        | 2234575820       | 33.92        |

As can be seen from Table 2, the vegetation coverage area in Nanjing decreased from 79.47% in 1993 to 54.89% in 2018, with a decrease of 24.58%. At the same time, the urban or exurban area of Nanjing increased from 10.94% in 1993 to 33.92% in 2018, with an increase of 22.98%. From 1993 to 2013, the decreasing amplitude of vegetation coverage area in Nanjing was relatively stable, as well as the increasing amplitude of urban or exurban area. Among them, from 1993 to 1998, the decrease of vegetation coverage area reached to the valley value of 2.11%; from 2003 to 2008, the increase of urban or exurban area reached to the peak value of 5.57%. From 2013 to 2018, the urban or exurban area experienced a substantial increase, increasing by 8.37%. At the same time, the vegetation coverage area also fell sharply, down 9.68%.
Spatial Trend of Changes

In this context, we compared the information in land cover maps between adjacent target years to learn the spatial trend of land cover change in Nanjing during the past 25 years in Figure 3.

![Figure 3. Land cover change results of Nanjing.](image)

It can be seen that the expansion of urban or exurban was the main characteristic of land cover change in Nanjing. This kind of expansion mainly happened in the central and southern parts of the city and transformed from cultivated land or grassland. From 1993 to 1998, cultivated land or grassland in the surrounding areas of the main city region transformed into urban or exurban. From 1998 to 2003, cultivated land or grassland in the southern part of the city converted into woodland. From 2003 to 2013, cultivated land or grassland in the central part and Yangtze River coastal in Nanjing transformed into cities and suburbs. From 2013 to 2018, the urban or exurban area in the southern part of city transformed from woodland and cultivated land or grassland.

Conclusions

In this study, we used GEE and Landsat images, the stratified sampling method was used to construct the classification feature set, and the CART method was used to obtain the results of the land cover classification of Nanjing in 1993, 1998, 2003, 2008, 2013 and 2018, and the forest land was extracted and calculated. The total area of cultivated land and grassland, as well as the total area of urban and suburban areas, using post-classification detection methods to analyze land cover changes in Nanjing. The following conclusions can be drawn from the study:

1. The vegetation covered areas of Nanjing decreased from 79.47% in 1993 to 54.89% in 2018.
2. The urban or exurban areas of Nanjing increased from 10.94% in 1993 to 33.92% in 2018.

References

[1] Sealey, Kathleen Sullivan, P-M. Binder, and R. King Burch. "Financial credit drives urban land-use change in the United States." Anthropocene 21, 2018, pp. 42-51.
[2] Lu, L., Kuenzer, C., Guo, H., Li, Q., Long, T., and Li, X. "A novel land cover classification map based on a MODIS time-series in Xinjiang, China." Remote Sensing 6.4, 2014, pp. 3387-3408.
[3] Cheng, Gong, Junwei Han, and Xiaoqiang Lu. "Remote sensing image scene classification: Benchmark and state of the art." Proceedings of the IEEE 105.10, 2017, pp. 1865-1883.
[4] GMTED2010: Global Multi-resolution Terrain Elevation Data 2010. Earth Engine Data Catalog, http://developers.google.com/earth-engine/datasets/catalog/USGS_GMTED2010.
[5] Li, Xuecao, Peng Gong, and Lu Liang. "A 30-year (1984–2013) record of annual urban dynamics of Beijing City derived from Landsat data." Remote Sensing of Environment 166, 2015, pp. 78-90.

[6] Xi, Yantao, Nguyen Thinh, and Cheng Li. "Spatio-Temporal Variation Analysis of Landscape Pattern Response to Land Use Change from 1985 to 2015 in Xuzhou City, China." Sustainability 10.11, 2018, p. 4287.

[7] Google Earth Engine API, https://developers.google.com/earth-engine/.

[8] Mutanga, O.; Kumar, L. “Google Earth Engine Applications.” Remote Sensing 11, 2019, p. 591.

[9] Nanjing City Electronic Map Data, http://www.dsac.cn/DataProduct/Detail/201820.

[10] R.J. Kauth and G.S. Thomas, "The tasseled Cap—A Graphic Description of the Spectral-Temporal Development of Agricultural Crops as Seen by LANDSAT." Proceedings of the Symposium on Machine Processing of Remotely Sensed Data, Purdue University of West Lafayette, Indiana, 1976, pp. 4B-41 to 4B-51.