Land Use Classification and Dynamic Change of Water Storage Projects Based on Support Vector Machine (Postprint)

Authors: Wang Jun, Chai Zhifu, Ma Haoyan, Zhao Zhimeng, Wu Jiabin, Fu Weiping (3)

Date: 2024-06-12T00:00:00+00:00

Abstract

To further reconstruct the historical process of land use changes before and after the construction of water storage projects and to better comprehend and forecast the direction of land use transitions, this study employs support vector machine theory to investigate the adaptability of land use type interpretation, and by examining the dynamic changes in land use, analyzes the self-adaptive regulation capacity and evolution direction of land use structures before and after the construction of water storage projects. The results indicate that: (1) Leveraging its advantages in self-learning and self-adaptation, the support vector machine achieves an overall accuracy of 91.7% and a Kappa coefficient of 0.90 for land use classification interpretation; except for relatively lower producer's accuracy for cultivated land, other land types such as water bodies and forest land exhibit high classification recognition capability. (2) Analyzing the evolution process of land use types using the Google Earth Engine (GEE) platform reveals that, influenced by the implementation of projects such as the second phase of the 'Three-North Shelter Forest' program (2001-2020), construction land and forest land areas have experienced substantial increases, with forest land area increasing nearly fivefold compared to the initial implementation period in 2000. (3) After project construction and operation, nearly two-thirds of forest land and construction land areas maintained their original state; water bodies and unused land, affected by water conservancy and urban construction projects, saw over 65% of their original types transform into other categories; the 'Three-North Shelter Forest' program accelerated the increase in forest land area and the improvement of grassland vegetation coverage, with net increases of 48.0% and 50.2% in the area of low-coverage grassland transitioning to medium- and high-coverage grassland, respectively.

ChinaRxiv

Full Text

Land Use Classification and Dynamic Changes in Water Storage Projects Using Support Vector Machine

WANG Jun[1,2], **CHAI Zhifu**[3], **MA Haoyan**[4], **ZHAO Zhimeng**[4], **WU Jiabin**[1,2], **FU Weiping**[3]

[1] China Institute of Water Resources and Hydropower Research, Institute of Water Resources for Pastoral Area, Hohhot 010020, Inner Mongolia, China

[2] Yinshanbeilu Grassland Eco-hydrology National Observation and Research Station, Beijing 100038, China

[3] Inner Mongolian Autonomous Region Research Institute on Hydraulic Sciences, Hohhot 010052, Inner Mongolia, China

[4] Central China Normal University, Wuhan 430079, Hubei, China

Abstract: To further restore and reconstruct the historical process of land use change before and after water storage project construction, and to better grasp and forecast land use transfer directions, this study investigates the applicability of support vector machine theory for land use type interpretation. By examining land use dynamic changes, we analyze the adaptive adjustment capacity and evolution direction of land use structures before and after project implementation. The results demonstrate that: (1) Relying on its advantages in self-learning and self-adaptation, the support vector machine achieved an overall classification accuracy of 91.7% and a Kappa coefficient of 0.90 for land use interpretation. Except for relatively lower producer accuracy for cultivated land, other land types such as water bodies and forest land exhibit high classification recognition capability. (2) Using the Google Earth Engine (GEE) platform to examine the evolution of land use types, the implementation of the second phase of the "Three-North Shelterbelt" project (2001–2020) caused substantial increases in construction land and forest land areas, with forest land area increasing nearly fivefold compared to the initial implementation stage in 2000. (3) Following project construction and operation, nearly two-thirds of forest land and construction land areas maintained their original state, while water bodies and unused land, influenced by water conservancy and urban construction projects, transformed from their original types to other types in over 65% of their area. The Three-North Shelterbelt project accelerated forest area expansion and improved grassland vegetation coverage, with net increases of 48.0% and 50.2% in the transition from low-coverage grassland to medium- and high-coverage grassland, respectively.

Keywords: land use; support vector machine; state transfer; water storage project

1.1 Study Area Overview

The study area encompasses the upstream, midstream, and downstream regions of the Chaganmuren River system affected by the Derisubaoleng Reservoir, with the reservoir's watershed serving as the natural boundary. Considering the river's runoff generation and convergence relationships, the study area was comprehensively defined covering 1501.35 km^2 [Figure 1: see original paper]. The Derisubaoleng Reservoir is located in Bairin Right Banner, Chifeng City, Inner Mongolia, with its dam site situated on the Chaganmuren River, a first-order tributary of the Xilamuren River in the West Liao River mainstream. The reservoir has a designed total storage capacity of 98.82 million m^3, accounting for 73.3% of the entire watershed area of 8427.4 km^2. It is a medium-sized multipurpose reservoir primarily for ecological protection, flood control, and industrial water supply, while also serving irrigation, aquaculture, and tourism. The project commenced water storage in 2010. Field investigations indicate that reservoir construction significantly reduced flood disaster impacts, effectively safeguarding downstream lives and property while substantially improving the surrounding ecology and climate. Additionally, according to surveys, Chifeng City, where the study area is located, has been included in the Three-North Shelterbelt System Construction since 1978, undergoing Phase I (1978–2000) and Phase II (2001–2020), which greatly increased regional forest coverage.

1.2 Research Methods

Support Vector Machine (SVM) is a binary classification model built upon statistical learning theory. Its fundamental model defines a linear classifier with maximum margin in feature space, essentially making it a nonlinear classifier [15]. The basic approach seeks to find a separating hyperplane that correctly classifies the training dataset while maximizing the geometric margin [16]. As illustrated in [Figure 2: see original paper], black solid circles and white hollow circles represent two classes of training samples. In high-dimensional space, $\omega \cdot x + b = 0$ (where ω is the normal vector) represents the separating hyperplane, while $\omega \cdot x + b = 1$ and $\omega \cdot x + b = -1$ are parallel planes passing through the points closest to the separating hyperplane from each class. The vertical distance between these two parallel planes, $2/\|\omega\|$, is called the classification margin or interval. For linearly separable datasets, infinitely many such hyperplanes exist (as in perceptron algorithms), but the hyperplane with maximum geometric margin is unique. Larger margins between parallel hyperplanes result in smaller total classifier error.

1.3 Data Acquisition and Processing

Land use type data were obtained from Landsat imagery downloaded from the United States Geological Survey (USGS). Remote sensing interpretation primarily utilized three systems: Landsat, Google Earth Engine (GEE), and Geographic Information System (ArcGIS). ArcGIS was used to select and create training samples, while GEE performed image classification processing, including remote sensing spectral index calculation, image composition, and classification using the SVM method. To fully compare changes before and after project construction, the study used 2010 (reservoir operation commencement) as the 分界点, further delineating pre- and post-construction periods with study years selected as 2000 and 2020 .

1.4 Accuracy Assessment Indices

To evaluate land cover classification results, confusion matrices were employed to statistically compare predicted and actual results for each class, using overall accuracy, producer accuracy, user accuracy, and Kappa coefficient as evaluation metrics. Producer accuracy represents the probability that ground-truth reference data of a given class are correctly classified, while user accuracy indicates the proportion of verification points classified as a particular category on the classification map that are correctly classified.

The Kappa coefficient is calculated as follows:

$$OA = \frac{1}{N}\sum_{i=1}^{r} x_{ii}$$

$$Kappa = \frac{p_o - p_e}{1 - p_e}$$

where x_{ii} is the number of correctly classified samples, N is the total number of samples, p_o is the sum of correctly classified samples for each category divided by the total number of samples (i.e., overall classification accuracy), and p_e is the sum of the product of actual and predicted quantities for each class divided by the square of the total number of samples.

2.1 Land Use Classification Accuracy Assessment

Overall accuracy and Kappa coefficient were used to evaluate classification accuracy for land use in the study area for 2000 and 2020. For 2000, 89 sample points were used for accuracy assessment [Figure 3: see original paper], with 80 correctly classified points, yielding an overall accuracy of 89.4% and Kappa coefficient of 0.88. This accuracy is consistent with results from Han Wenting et al. [22], indicating reliable classification results. Among different land use types, forest land and construction land showed high classification accuracy, while cultivated land had relatively lower accuracy and was easily confused with grassland.

For 2020 classification, 86 sample points were used for accuracy assessment [Figure 3: see original paper], with 81 correctly classified points, achieving an overall accuracy of 94.2% and Kappa coefficient of 0.93, again confirming reliable classification results. Water bodies, construction land, and unused land showed high classification accuracy, while cultivated land producer accuracy remained relatively low and was easily confused with grassland. Combining accuracy assessments for both years, 175 of 191 sample points were correctly classified, with an average overall accuracy of 91.7% and average Kappa coefficient of 0.90 .

2.2 Land Use Type Change Characteristics Based on Support Vector Machine

Building upon accuracy evaluation and analysis, this study utilized the GEE platform to examine land use type evolution in the study area using Landsat imagery, analyzing land use type changes from 2000 to 2020. Due to the implementation of the Three-North Shelterbelt Project and other initiatives, construction land and forest land areas showed substantial increases . Construction and other construction land areas increased nearly fivefold, while forest land area increased by nearly 23.7 km^2 by the end of Phase II (2020) compared to 2000.

Averaging land use type areas before (2000) and after (2020) project construction revealed that, except for grassland types, other land use categories did not show large fluctuations [Figure 4: see original paper]. Notably, cultivated land area did not increase substantially as might be expected, with its maximum extent occurring in 2010 within 5% variation. Grassland, as the most important land type in the study area, consistently maintained over 50% area proportion throughout the study period. Medium-coverage grassland showed an increasing trend, with more significant growth after reservoir construction, indicating an objective improvement in ecological vegetation in the Chaganmuren River basin where the Derisubaoleng Reservoir is located [Figure 5: see original paper].

2.3 Transfer Direction of Different Land Use Types Before and After Water Storage Project Construction

To further clarify land use structure transfer directions before and after reservoir construction, this study employed a state transfer matrix model to analyze spatial changes in land use types from 2000 to 2020 [Figure 6: see original paper]. Analysis of spatial transfer results shows that nearly two-thirds of forest land and construction land areas maintained their original state after project construction. With reservoir development, water bodies and unused land experienced substantial transformations, with over 65% of water bodies converting to other land use types. In the upstream area of the Derisubaoleng Reservoir dam site, grassland rapidly transformed into reservoir water bodies after dam construction, while downstream river channels narrowed, water area decreased, and some areas converted to cultivated land and grassland.

Furthermore, the Three-North Shelterbelt ecological restoration project accelerated forest area expansion and grassland vegetation coverage improvement, with more cultivated land and grassland converting to forest land. Low-coverage grassland transformed into medium- and high-coverage grassland, with net increases of 232.3 km^2 and 64.9 km^2, respectively . Among all land types, forest land showed the highest average annual increase rate at 23.2%, followed by construction land at 10.8%. Water bodies and unused land decreased to varying degrees, with average annual reductions not exceeding 3%. Although grassland area decreased, aggregating high-, medium-, and low-coverage grassland areas showed annual variation not exceeding 1%. Compared with Li Zhenzhen's [25] research on land use change dynamics in Northeast China, this study area shows consistent conclusions: cultivated land, forest land, and construction land had positive dynamic degrees, while grassland and water bodies had negative dynamic degrees. The study also analyzed temporal evolution of land use transfer in Northeast China from 2000–2020, finding that the most intense and frequent transfers occurred among cultivated land, forest land, and grassland, which aligns with the results in .

3 Discussion

From 2000 to 2020, cultivated land, forest land, and construction land areas in the study area showed increases of varying degrees, while grassland, water body, and unused land areas decreased. Specifically, nearly two-thirds of forest land and construction land maintained their original appearance, while water bodies and unused land, influenced by water conservancy and urban construction projects, transformed from original types to other types in over 65% of their area. Additionally, the Three-North Shelterbelt Project and other initiatives accelerated forest area expansion and grassland vegetation coverage improvement, with net increases of 48.0% and 50.2% in transitions from low-coverage grassland to medium- and high-coverage grassland, respectively.

The study also found that forest land area showed substantial increases from 2000–2020, with an average annual growth rate of 23.2%. This period coincides with the second phase (Phase IV and V) of China's Three-North Shelterbelt System Construction Project [27], aligning with construction priorities in Northeast China's western region within the shelterbelt system—namely, building farmland shelterbelts as the basic framework, combining multiple forest types and tree species, integrating networks, belts, and patches with trees, shrubs, and grasses, and creating a regional shelterbelt system that integrates agriculture, forestry, and animal husbandry across connected counties. The land use change monitoring results based on support vector machines can provide data support for evaluating the construction effectiveness of the Three-North Shelterbelt System Construction Project.

4 Conclusions

This study conducted adaptive research on land use type interpretation based on support vector machines, objectively examining dynamic land use changes before and after water storage project construction, and further analyzing the adaptive adjustment capacity and evolution direction of land use structures using state transfer matrices. The main conclusions are as follows:

(1) Leveraging its self-learning and self-adaptation advantages, this study performed adaptive research on interpreting land use types before and after Derisubaoleng Reservoir construction in the typical Chaganmuren River basin. Through comparative analysis with sample data, 175 of 191 sample points were correctly classified, with an average overall accuracy of 91.7% and average Kappa coefficient of 0.90. Except for relatively low producer accuracy for cultivated land, other land types such as water bodies and forest land demonstrated high classification recognition capability.

(2) The GEE platform was used to examine land use type evolution processes and differences before and after project construction. Over the past 20 years, influenced by the second phase of the Three-North Shelterbelt Project (2001–2020), construction land and forest land areas increased substantially, with construction and other construction land areas increasing nearly fivefold and forest land area increasing nearly five times compared to the initial implementation stage in 2000.

(3) State transfer matrix model analysis of land use type spatial transfer directions before and after water storage project construction revealed that over two-thirds of forest land and construction land areas maintained their original state after project construction, while water bodies and unused land, affected by water conservancy and urban construction projects, transformed from original types to other types in over 65% of their area. The Three-North Shelterbelt Project accelerated forest area expansion and grassland vegetation coverage improvement, with net increases of 48.0% and 50.2% in transitions from low-coverage grassland to medium- and high-coverage grassland, respectively.

References

[1] Dong Zengchuan, Liang Zhongmin, Li Dayong, et al. Influences of Three Gorges Project on water resources and ecological effects in Poyang Lake[J]. Journal of Hohai University (Natural Sciences), 2012, 40(1): 13-18.

[2] Kang Yan, Gao Xuan, Li Lingjie, et al. A bilayer optimization method of water stor supplying sequence and operation diagram for water supply reservoir group[J]. Journal of Hydraulic Engineering, 2022, 53(10): 1240-1250.

[3] Hua Ding, Hu Shi, Mo Xingguo. Impacts of water conservancy projects on agricultural water use efficiency and crop productivity in the Nianchu River Basin of China[J]. Transactions of the Chinese Society of Agricultural Engineering, 2022, 38(14): 98-107.

[4] Dang Chiheng. Runoff Simulation and Risk Assessment on Reservoir Operation in the Jing River Basin Under Climate Change[D]. Xi'an: Chang'an University, 2021.

[5] Hu Changtong, Yang Tao, Wan Xuhao, et al. Distribution characteristics of

heavy metals in river sediments and their relationship with land use types in Xi'an City[J]. Arid Zone Research, 2022, 39(4): 1270-1281.

[6] Ma Yaoyao, Shi Peijun, Xu Wei, et al. Remote sensing monitoring of the ecological environment of hydropower station construction and operation in arid areas[J]. Arid Zone Research, 2023, 40(9): 1498-1508.

[7] Su Yingqing, Zhang Enyue, Liu Yuan, et al. Land use change and ecological environment effects on Fenhe River Basin[J]. Arid Zone Research, 2022, 39(3): 968-977.

[8] Yuan Jingwen, Wu Chen, Du Bo, et al. Analysis of landscape pattern on urban land use based on GF-5 hyperspectral data[J]. Journal of Remote Sensing, 2020, 24(4): 465-478.

[9] Luo Jiancheng, Hu Xiaodong, Wu Tianjun, et al. Research on intelligent calculation model and method of precision land use/cover change information driven by high resolution remote sensing[J]. National Remote Sensing Bulletin, 2021, 25(7): 1351-1373.

[10] Feng Quanlong, Niu Bowen, Zhu Dehai, et al. Review for deep learning in land use and land cover remote sensing classification[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(3): 1-17.

[11] Lin Nan, Jiang Qigang, Yang Jiajia, et al. Classifications of agricultural land use based on high spatial resolution ZY1-02C remote sensing images[J]. Transactions of the Chinese Society for Agricultural Machinery, 2015, 46(1): 278-284.

[12] Zhou Ke, Yang Yongqing, Zhang Yanna, et al. Review of land use classification methods based on optical remote sensing images[J]. Science Technology and Engineering, 2021, 21(32): 13603-13613.

[13] Zhang Yinhui, Zhao Gengxing. Classification methods of land use/cover based on remote sensing technologies[J]. Journal of China Agricultural Resources and Regional Planning, 2002(3): 21-25.

[14] Han Wenting, Guo Congcong, Zhang Liyuan, et al. Classification method of land cover and irrigated farm land use based on UAV remote sensing in irrigation[J]. Transactions of the Chinese Society for Agricultural Machinery, 2016, 47(11): 270-277.

[15] Ren Haijuan, Dong Jianjun, Li Xiaoyuan, et al. Extraction artificial alfalfa grassland information using Landsat8 remote sensing data[J]. Chinese Journal of Grassland, 2015, 37(2): 81-87, 120.

[16] Auria L, Moro R A. Support Vector Machines (SVM) as a Technique for Solvency Analysis[M]. Berlin: Social Science Electronic Publishing, 2008.

[17] Cervantes J, Garcia-Lamont F, Rodríguez-Mazahua L, et al. A comprehensive survey on support vector machine classification: Applications, challenges and trends[J]. Neurocomputing, 2020, 408: 189-215.

[18] Thamaga K H, Dube T, Shoko C. Evaluating the impact of land use and land cover change on unprotected wetland ecosystems in the tropical areas of South Africa using the Landsat dataset and support vector machine[J]. Geocarto International, 2022, 37(25): 7821-7840.

[19] Li Daoji, Guo Haitao, Lu Jun, et al. A remote sensing image classification procedure based on multilevel attention fusion U-Net[J]. Acta Geodaetica et Cartographica Sinica, 2020, 49(8): 1051-1064.

[20] Wang Shengli, Zhang Lianpeng, Zhu Shouhong, et al. Multi-invariant feature combined photovoltaic power plants extraction using multi-temporal Landsat8 OLI imagery[J]. Bulletin of Surveying and Mapping, 2018(11): 46-52.

[21] Cherkassky V, Ma Y. Practical selection of SVM parameters and noise estimation for SVM regression[J]. Neural Networks, 2004, 17(1): 113-126.

[22] Liu Zhigang. Key Problems of Applying Support Vector Machines to the Classification of Spectral Remote Sensing Imagery[D]. Wuhan: Wuhan University, 2004.

[23] Yang Q, Li X, Shi X. Cellular automata for simulating land use changes based on support vector machines[J]. Computers & Geosciences, 2008, 34(6): 592-602.

[24] Li Zhenzhen. Study on the Impact of Land Use Change on Ecosystem Services in Northeast China[D]. Harbin: Heilongjiang University, 2023.

[25] Dao Rina. Effects of Climate and Land Use Change on Vegetation NDVI in the Three North Region[D]. Hohhot: Inner Mongolia Normal University, 2019.

[26] Li Shidong, Feng Deqian. World famous ecological project: China Three-North Shelterbelt system Construction Project[J]. Zhejiang Forestry, 2021(9): 9-11.

Note: Figure translations are in progress. See original paper for figures.