Academic Journal of Science and Technology
ISSN: 2771-3032 | Vol. 5, No. 3, 2023

Research on Light Pollution Based on Bp Neural Network

Jiawang Wang#, Hehong Huang#, Zilei Ji#

School of Mechanical Engineering, Tianjin University of Science and Technology, Tianjin, 300222, China
#These authors contributed equally.

Abstract: This study focuses on the impact of local development level, population, biodiversity, geography, climate and other factors on local light pollution, establishes a widely applicable evaluation model, and reasonably evaluates the impact of various factors on light pollution. Firstly, twelve factors that may be related to light pollution are selected, grayscale processing is performed on the night remote sensing image, the color value of the grayscale image is read and quantified as the light pollution value, and the selected factors are averaged and normalized. Then, the gray correlation model is used to obtain the gray correlation degree and select the factors with strong correlation. Secondly, the light pollution evaluation model is established by the BP neural network, and the light pollution risk level is divided into four grades: A, B, C, and D, of which A grade represents the low light pollution risk level (A: 0~1, B:2~4, C:5~7, D:8~9, and the number represents the light pollution value). Based on this, this study established a set of evaluation index system to determine the level of light pollution risk level of a site. Finally, different types of locations are selected to test the established prediction model to verify the wide applicability of the established BP neural network evaluation system. The results show that the light pollution evaluation model in this study can effectively predict light pollution for different locations and has wide applicability.

Keywords: Light pollution, Remote sensing imagery, Gray associative model, BP neural networks.

1. Introduction

With the rapid development of science and technology and economy, people's quality of life has been improved, urban planning and construction have also made continuous progress, and creating a perfect and comfortable nighttime living environment has become an important part of development. In order to create a perfect and comfortable nighttime living environment, night artificial lighting design is becoming increasingly important. However, the problem of light pollution caused by excessive or poor use of artificial light, just like food, water, air, and even nuclear pollution, has attracted more and more attention [1].The negative effects of light pollution include disturbing the natural life patterns of animals and plants and disrupting the balance of ecosystems; For humans, it can interfere with the biological clock, disrupt endocrine balance, and cause physical and psychological problems; Excessive night lighting also consumes unnecessary electricity, resulting in a great waste of energy, etc., and glare caused by artificial light may lead to some motor vehicle accidents[2].

In this context, while community officials or local groups can implement intervention strategies to mitigate the negative effects of light pollution, such as developing policies to reduce the use of artificial light, it may then provide favorable conditions for offenders to increase crime. At present, in the field of light environment research, there are no accurate quantitative and evaluation standards for the safety, comfort, environmental performance and health requirements of lighting, which makes it impossible to evaluate the risk level of light pollution in a certain place, and then it is impossible to propose targeted strategies to deal with light pollution in different regions. Therefore, this study aims to establish a widely applicable light pollution evaluation index system to determine the risk level of light pollution in a location, which is helpful to help the local government formulate corresponding intervention strategies to mitigate the impact of light pollution.

In this study, the luminous remote sensing map was processed to obtain the degree of light pollution in each region. Then, SPSSPRO is used to analyze the correlation between the influencing factors and the light pollution degree value, and the grey correlation degree is obtained, and the influencing factors with large correlation are selected. Secondly, the BP neural network algorithm is used to find the mapping relationship between the influencing factors and the light pollution degree value, and the BP neural network prediction model for predicting the light pollution degree of the region given the influencing factor value in the region is established. Finally, different types of sites are selected to test the model and analyze whether the model is reasonable. The model creation process is shown in Figure 1.

2. Materials and Methods

2.1. Data acquisition and preprocessing

In this paper, a model is developed to predict the degree of light pollution based on given indicators. In 2018, China launched the world's first professional luminous remote sensing satellite, Luojia 1 (LJ 1-01). The satellite provides a luminous remote sensing image with a resolution of 130 m, which greatly improves the resolution compared with DMSP and VIIRS, and makes it possible to carry out fine night-time light pollution monitoring at the urban scale [3]. This article from http://59.175.109.173:8888/app/login.html for wuhan university no yoga a satellite data night lights. As shown in Figure 2. Get the light pollution map from the Light Pollution map (lightpollutionmap.info) as shown in Figure 2. The grayscale processing and quantization of the night remote sensing map were carried out to obtain the light pollution degree of each region. In this paper, we obtained the pollutant nitrogen oxide emissions, pollutant particulate emissions, crop disaster, environmental emergencies, tourism, pollutant sulfur dioxide emissions, earthquake frequency, forest coverage rate, annual average temperature, population density, green area, air quality in different regions through http://www.stats.gov.cn/. GDP value as a possible correlation factor with light pollution.

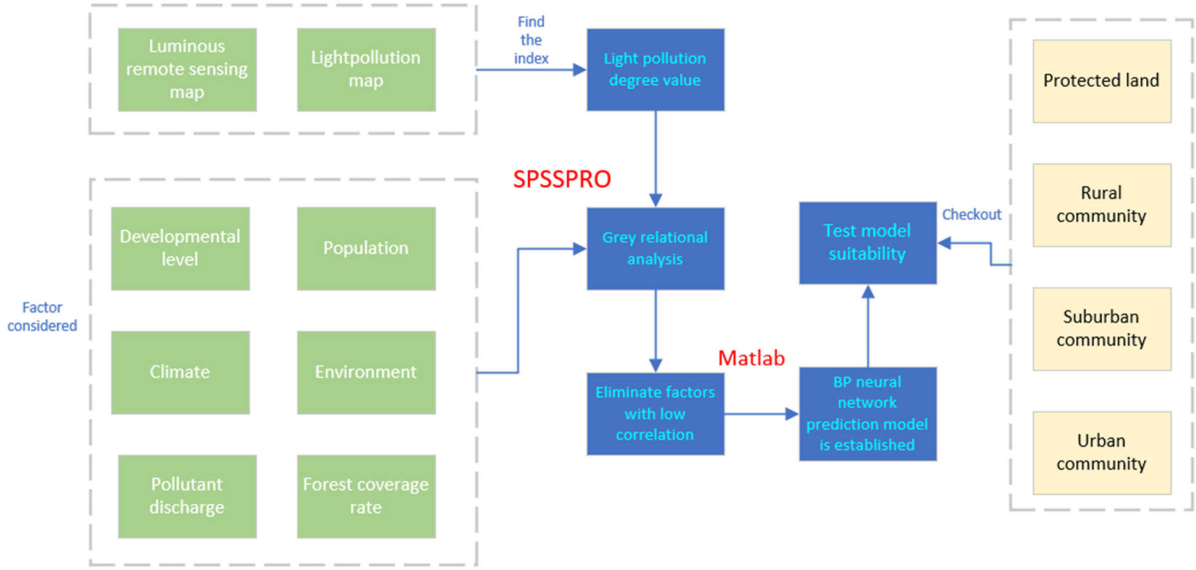

Figure 1. Model creation flowchart

Dimensionless processing is carried out for data of light pollution-related factors. The dimensionless processing adopts mean value and initial value, which means to uniformly divide the sequence data by the initial value. Since there is little difference in the magnitude of the sequence of the same factor, these values can be sorted to the magnitude of 1 by dividing the initial value. Since the larger order of magnitude has a larger mean, it can be normalized to the order of 1.

Initial value processing:

$$x_i(k) = \frac{x_i(k)}{x_i(1)}, k = 1,2 \dots n; i = 0,1,2 \dots m \qquad (1)$$

Averaging:

$$x_i(k) = \frac{x_i(k)}{\bar{x}_i}, k = 1,2 \dots n; i = 0,1,2 \dots m \qquad (2)$$

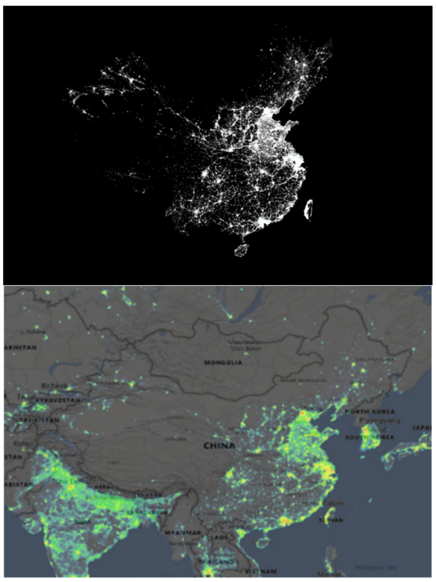

Figure 2. Night remote sensing image (left) and Light pollution image(right)

2.2. Method introduction

2.2.1. Grey correlation model

Grey correlation analysis is the correlation coefficient analysis in essence. Firstly, the correlation coefficient of the ideal scheme composed of each scheme and the best index is obtained from the correlation coefficient, and then the correlation degree is sorted and analyzed according to the degree of correlation, and the conclusion is drawn [4].

Grey correlation model (GM) is a mathematical model used to analyze data. Grey correlation analysis is a quantitative description and comparison method for the development and change of a system. Its basic idea is to determine whether the relationship between reference data columns and several comparative data columns is close by determining the similarity degree of their geometric shapes. Generally, this method can be used to analyze the influence degree of various factors on the results, and can also be used to solve the problems of comprehensive evaluation over time. The core of this method is to establish the parent sequence over time according to certain rules, take the changes of each evaluation object over time as the sub-sequence, find the correlation degree between each sub-sequence and the parent sequence, and draw a conclusion according to the correlation degree. In data analysis, this method can be used to study the interactions and effects of different variables, as well as to analyze future trends and forecasts.

2.2.2. BP neural network

BP neural network model is a widely interconnected neural network composed of multiple neurons, which can simulate the interactive response of biological nervous system in the real world and objects. Artificial neural network information processing is to train neural network through information samples, so that it has human brain memory, recognition ability, complete various information processing functions. It has good abilities of self-learning, adaptive, associative memory, parallel processing and non-linear transformation, avoids complex mathematical derivation, and can still guarantee stable output in the case of sample defect and parameter drift. Artificial neural network, which simulates human brain intelligence, has been successfully applied in many fields [5].

BP neural network can be divided into two parts, BP and neural network. BP is short for Back Propagation, meaning back propagation. BP network can learn and store a large number of input-output mode mapping relationships without revealing the mathematical equation describing this mapping relationship in advance. Its learning rule is to use the fast descent method, through back propagation to constantly adjust the weight and threshold of the network, so as to minimize the sum of squares of error of the network. Its main characteristics are: the signal is propagated forward, and the error is propagated back. Neural networks, a common mathematical model in machine learning, process information by building structures similar to synaptic connections in the brain. In the application of neural network, information processing units are generally divided into three categories: input units, output units and implicit units.

2.3. Model evaluation index

The evaluation indexes of the model in this paper are SSE squared error, MAE mean absolute error, MSE mean square error, RMSE root mean square error and MAPE mean percentage error.

$$SSE = \Sigma(y_i - \bar{y})^2 \quad (3)$$

$$MSE = (1/n) * \Sigma(i = 1 \text{ to } n)\ (yi - \hat{y}i)^2 \quad (4)$$

$$MAE = (1/n) * \Sigma|i = 1 \text{ to } n|\ |\ yi - xi\ | \quad (5)$$

RMSE = \sqrt{\frac{1}{n} \sum_{i = 1}^{n}(y_i − \hat{y_i})^2} (6)

$$MPE = (1/n) * \Sigma[(A\_i - F\_i)/A\_i] * 100\% \quad (7)$$

3. Model Establishment and Solution

3.1. The establishment of simulation model

The aim of this study is to develop a broadly applicable indicator for determining the level of risk at a site. The grayscale processing was carried out on the night remote sensing monitoring image, and the color value of the grayscale processing image was quantified, and the quantified value represented the degree of light pollution. The gray correlation model was used to analyze whether the selected factors were associated with light pollution, and the factors with poor correlation were eliminated. Then the BP neural network model was used to establish the mapping relationship between the factors related to light pollution and the light pollution level. Finally, the light pollution risk level is classified.

3.1.1. The establishment of grey relational model

After pretreatment, relevant factors such as pollutant nitrogen oxide emissions, pollutant particulate emissions, crop disaster, number of environmental emergencies, tourism, pollutant sulfur dioxide emissions, number of earthquakes, forest coverage rate, annual average temperature, population density, green area, air quality and GDP are taken as characteristic sequence variables. The quantized luminous remote sensing data were taken as the degree of light pollution, as the parent sequence variable, and the region as the index item. Find the correlation degree between each subsequence and parent sequence, and draw a conclusion according to the correlation degree. The establishment process of grey correlation model is shown in Figure 3.

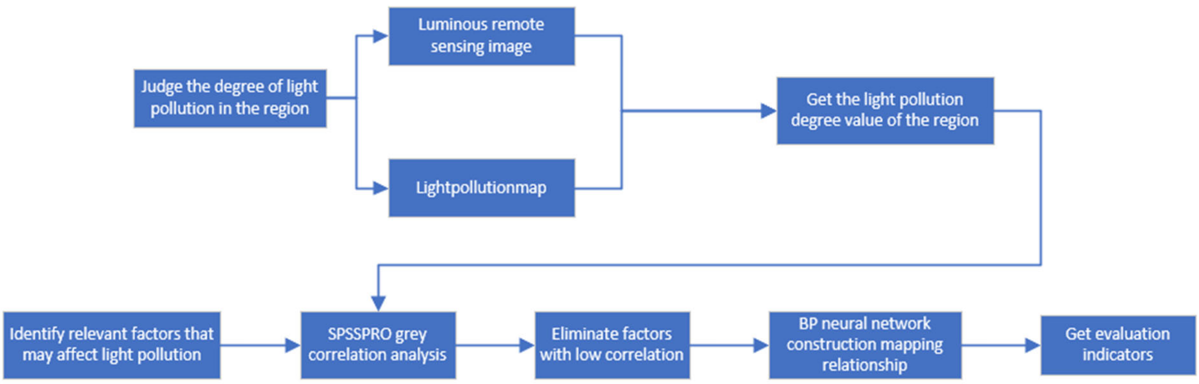

Figure 3. Flow chart of grey correlation model

3.1.2. The establishment of BP neural network model

Firstly, the relevant factor data after elimination is sorted out, the MATLAB code is written to read the data and load the input data. Then, part of the data is taken as the training data and part is taken as the prediction data. At the same time, the number of nodes in the input layer, the number of nodes in the hidden layer and the number of nodes in the output layer are set. Input weight and threshold, configure training times, learning rate, training target minimum error, etc., and then conduct neural network training, normalization of test samples, and finally perform BP neural network prediction.

Through the above steps, BP neural network code is run through MATLAB to determine the best hidden layer node number and various error indicators, obtain the prediction error of test samples, and compare with the actual value, and finally determine the quantitative relationship between relevant factors and the degree of light pollution.

3.2. Solution of model

3.2.1. The solution of grey correlation model

Using SPASSPRO, grey relation values between parent sequence and feature sequence are solved. As shown in Figure 4.

$$\xi_i(k) = \frac{\underset{i}{min}\underset{k}{min}|y(k)-x_i(k)| + \rho\underset{i}{max}\underset{k}{max}|y(k)-x_i(k)|}{|y(k)-x_i(k)| + \rho\underset{i}{max}\underset{k}{max}|y(k)-x_i(k)|} \quad (8)$$

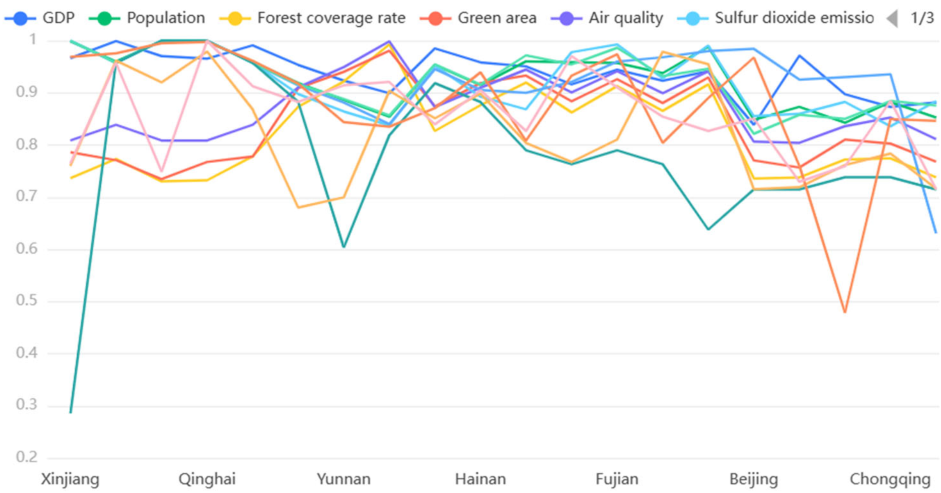

Figure 4. Grey correlation coefficient

Solve the grey relational degree value.

$$r_i = \frac{1}{n}\sum_{k=1}^{n}\xi_i(k), k = 1,2,\dots n \tag{9}$$

Table 1. Correlation degree result table

Evaluation	Degree of association	Ranking
GDP	0.938	1
Population	0.927	2
Forest coverage rate	0.927	3
Green area	0.926	4
Sulfur dioxide emission	0.922	5
Nitrogen oxd emission	0.881	6
Air quality	0.874	7
PM2.5	0.859	8
Environmental emergency	0.446	9
Crop disaster situation	0.427	10
Tourist trade	0.324	11
Number of earthquakes	0.284	12

Table 1 shows the correlation degree between 12 related factors and the degree of light pollution, among which GDP is the most correlated with the degree of light pollution, and the number of earthquakes is the least correlated with the degree of light pollution. The correlation degree of sudden environmental events is 0.446, that of agriculture as a disaster situation is 0.427, that of tourism is 0.324, and that of the number of earthquakes is 0.284, all of which are less than 0.5, indicating a weak correlation. In the subsequent BP neural network is used to establish correlation factors and light pollution degree mapping relationship, environmental emergencies, crop damage, tourism and earthquake frequency are excluded.

3.2.2. The solution of BP neural network model

When BP neural network is used for training, when the optimal number of hidden nodes is 9, the minimum mean square error is 0.020384. Therefore, the number of hidden nodes is 9 in subsequent calculation to ensure accuracy. The prediction error of the test sample of the trained neural network model fluctuates around the actual value, as shown in Table 2, whose sum of error squares SSE is 16.4838, mean absolute error MAE is 1.0661, mean square error MSE is 1.4985, and root mean square error RMSE is 1.2241. The average percentage error MAPE is 17.1451%, the correlation coefficient is 0.78294, and the prediction accuracy is 82.8549%, reaching 80%. The established model meets the requirements of wide application. The light pollution risk level is divided into four levels: A, B, C, and D. A indicates low light pollution risk level, B indicates medium light pollution risk level, C indicates higher light pollution risk level, and D indicates higher light pollution risk level. (A: 0~1, B: 2~4, C: 5~7, D: 8~9, the number represents the light pollution value)

Table 2. Table of error indicators

Error indicator	Result
SSE	16.4838
MAE	1.0661
MSE	1.4985
RMSE	1.2241
MAPE	17.1451%
Forecast accuracy	82.8549%
Coefficient of association	0.7829

3.3. Test of model

Considering that light pollution is related to local development, in order to ensure the reliability of test results, multiple types of sites were selected during the test, including four types of sites: nature reserve, rural community, suburban community and urban community, and multiple sets of data were selected for each type of site. The verification process is shown in Figure 5.

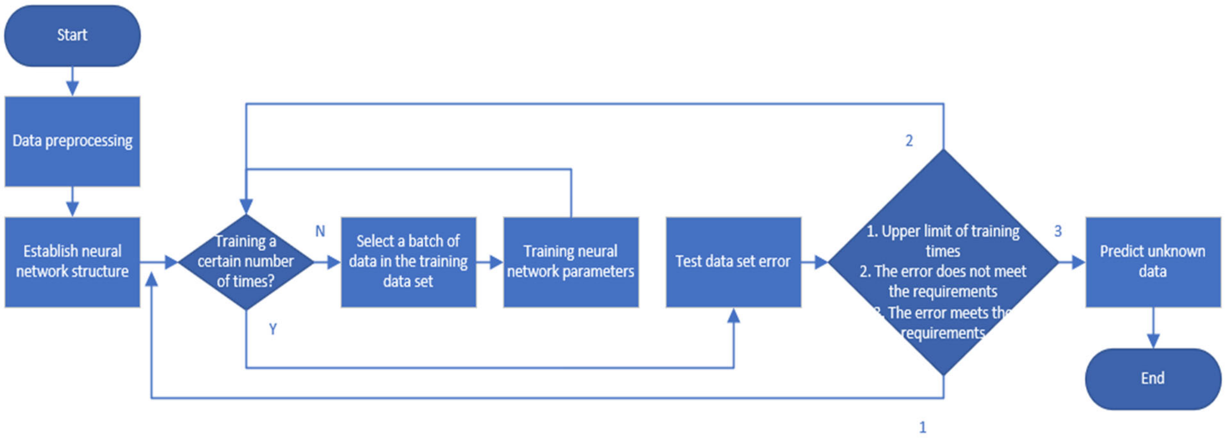

Figure 5. Model test flow chart

The relevant factor data of the selected locations are processed and substituted into the BP neural network model established above for prediction. The predicted data is analyzed and processed, and then the average value is taken as the light pollution degree of this type of location to determine the light pollution level. Finally, the analysis is made according to the comparison between the actual situation and the predicted results.

Figure 6 shows that in the second round of test, the mean square error is 0.0055593, which is almost negligible and has the best effect. Therefore, the training results of the second round are used in the evaluation. As can be seen from Table 3, the final evaluation results show that the nature reserve is at low light pollution risk level, the rural community is at medium light pollution risk level, the suburban community is at high light pollution risk level, and the urban community is at high light pollution risk level. According to the actual situation, the nature reserve has low population density, wide green area, large forest coverage, low pollutant discharge, and almost no artificial light, which should be low light pollution risk level, consistent with the evaluation; Rural communities have lower population density, lower GDP, lower pollutant emission, and less artificial light use, but the population density, pollutant emission, and artificial light use are higher than nature reserves, so the risk level of light pollution should be medium, consistent with the evaluation results. Suburban communities have higher population density, higher GDP, higher pollutant emission, lower forest coverage and more artificial light use, so suburban communities should have higher light pollution risk level, which is consistent with the evaluation. Urban communities have high population density, high GDP, high emission of pollutants, low forest coverage and high use of artificial light, so urban communities should be the risk level of high light pollution, which is consistent with the evaluation, indicating that the BP neural network model established is suitable for predicting the risk level of light pollution in various types of locations, and has wide applicability.

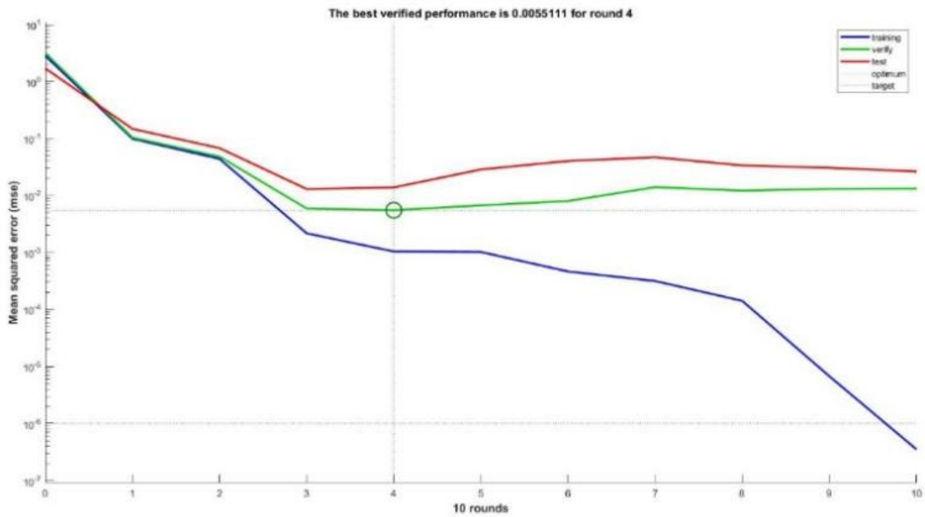

Figure 6. Mean-square error chart of the second round forecast

Table 3. Regional light pollution risk level

Region	Light pollution degree value	Light pollution risk level
Protected land	0.5996	Low
Rural community	3.3759	Middle
Suburban community	6.6338	High
Urban community	8.5802	Higher

4. Conclusions

Research on light pollution at home and abroad is mainly qualitative research, and quantitative research on light pollution is few. At the same time, research mainly focuses on a single area, and there is a lack of a set of widely applicable quantitative evaluation system for light pollution. In this study, grey correlation and BP neural network were used to establish a set of widely applicable quantitative evaluation system. Through continuous training of the model through machine learning, the current model has reached the prediction accuracy of 82.8549%. If the training continues, it will reach a higher accuracy, which can provide favorable technical support for the construction and evaluation of light environment. It provides more basis for optimizing the design of light environment and support for reducing the risk level of light pollution. Promote the improvement of light-polluted roads.

References

[1] SU X M. Comprehensive evaluation of residential light pollution [D]. Tianjin University,2011.

[2] Zhang Yue, Xu Yongming, Xiong Wencheng, et al. Analysis of night light pollution monitoring and control measures by remote sensing [J]. Environmental Monitoring and Warning,2019,11(5):108-112.

[3] Li Jiayi, Xu Yongming, Cui Weiping, Wu Yuyang, Wang Jing, Su Boyang, Ji Meng. Monitoring of night light pollution in Nanjing City based on Luojia No.1 luminous remote sensing data [J]. Remote Sensing of Natural Resources,2022, Vol. 34 (2): 289-295

[4] ZHOU Xiuwen. Research and Application of Grey Relational Degree [D]. Jilin University,2007.

[5] Yin Chunhua, Chen Lei. Based on the research and application of BP neural network model of population prediction [J]. Journal of population, 2005 (02): 44-48, DOI: 10.16405 / j.carol carroll nki. 1004-129 - x. 2005.02.009.