Research on regional light pollution risk level measurement based on neural network model

Yuyang Ma[1], Yifei Song[2], Ruiyu Zhang[2]

[1]Tourism School, Xi'an International Studies University, Xi'an, Shaanxi, 710128, China
[2]Business School, Xi'an International Studies University, Xi'an, Shaanxi, 710128, China

Abstract: *Light pollution is a new type of environmental pollution, long-term exposure to light pollution not only harms human health, but also restricts urban development and even threatens biodiversity [1]. In order to evaluate the risk level of light pollution in different regions, this paper first selects 9 indicators leading to light pollution risk, takes 10,000 regions in the world as samples, constructs a neural network model, and the output of the network is the probability value of the light pollution risk level in the region. To improve the accuracy of the output, the parameters are then updated by back propagation to constrain the predicted regional risk level to be close to the actual risk level. Finally, four real areas are selected, and the indicator data brought into the area is used to obtain the risk level.*

Keywords: *light pollution risk level, neural network model, back propagation algorithm*

1. Introduction

Light pollution refers to the harm caused by the application of a large number of artificial light to our live. Light pollution has the characteristics of active infringement, unperceived, damage involvement and so on [2]. With the continuous improvement of lighting technology, the phenomenon of light pollution caused by artificial light is becoming more and more serious, which has a negative impact on the environment, as well as our health and safety. For example, long-term exposure to light pollution not only damages human health, but also restricts urban development and even threatens biodiversity [3].

In 2021, Hao Qingli studied the risk assessment of birds under light and noise pollution stress, and this ecological risk assessment method can provide scientific and standardized research and technical examples for the rapid assessment of the ecological risk of light and sound pollution in urban macroregions [4]. Yan Ziyan studied the changes in light pollution in protected areas from 1992 to 2021 in 2023 and found that the increase in light pollution is closely related to urban expansion, and specific issues need to be further explored.

To solve those questions, team collected academic literature related to light pollution, and synthesized the main factors affecting light pollution, and obtained the main factors affecting light pollution through the literature, average temperature, population size, green area, GDP, climate type, GDP per capita, green area per capita, and biological species per unit of green area as the metrics according to the importance of their influence.

In addition, the team selected 10,000 regions in the world and took the data of each metric in 2022 to construct a $10,000 \times 9$ matrix, and obtained the data of light pollution risk level of the region to construct a $10,000 \times 1$-dimensional target vector. The PyTorch framework is used to construct a neural network, and the output of this network is the probability value of the light pollution risk level in the region, and the level corresponding to the maximum probability value is selected as the light pollution level.

At the same time, the loss value between the output probability vector and the real light pollution level of the region is calculated according to the cross-entropy objective function, and the parameter values of the neural network are updated by the back propagation algorithm using Adam optimizer, and the expert evaluation network is obtained after the training convergence, which can extract the statistical properties and distribution between the input data and simulate the function mapping between the nine indicators and the light pollution level. The model is used to evaluate the light pollution risk level specifically in combination with the specific conditions of the four regions.

2. Assessment of Light Pollution Risk Levels

2.1 Development of Metrics

Combined with the title, we know that light pollution is becoming more and more serious, and the number of current studies on urban dynamic light pollution is small, and there is a lack of quantitative research on the impact of dynamic light pollution, so the scientific basis is not enough to classify urban dynamic light pollution [5]. In order to address this situation and reduce the harm caused by light pollution to our lives, we need to select a set of applicable criteria to judge the current level of light pollution risk in an area. The International Commission on Lighting (CIE) defines light pollution as "light behavior that causes annoyance, discomfort, mental distraction or reduces the ability to recognize important information in the environment due to the amount, direction or spectral distribution of light in a specific environment"[6]. It can be seen that light pollution is generated in an environment with a large amount of light, and population, green space, rapidly developing economic GDP, per capita GDP and per capita green space create environmental and technological influences for the formation of light pollution, thus constituting the influencing factors of light pollution.

At the same time, because the use of artificial light is mainly concentrated in the night after sunset, the lack of sunlight leads to the demand for artificial light [7]. Therefore, the average sunshine duration, average temperature and climate type constitute the prerequisite factors for the generation of light pollution, and the number of species of organisms and the species of organisms per unit of green space provide objects for the influence of light pollution. To sum up, the average annual sunshine duration, average temperature, population number, green area, GDP, climate type, per capita GDP, per capita green area, and number of species per unit of green area are selected as the criteria for judging the light pollution risk level.

2.2 Data sourcing and Data Preprocessing

In the establishment of the expert evaluation network model, combining the important factors affecting light pollution, we finally selected nine factors as measuring standards, including average annual sunshine duration, average temperature, population number, green space area, GDP, climate type, per capita GDP, per capita green space area and number of biological species per unit green space area. Meanwhile, in order to improve the reliability of the results. We selected 10,000 regions as samples, analyzed the relationship between their light pollution levels and the nine factors, and finally built the model. A large amount of data ensured that the final model results were more persuasive.

Our data are mainly obtained from: Light Pollution Map (to confirm the intensity of light pollution in the area), NOAA (to obtain weather temperature and climate related data), National Bureau of Statistics (GDP, population, area), NUMBEO query about crime rate) and both big data capture to search for related information.

Data preprocessing: After obtaining the light pollution level indicators of 10,000 regions and the specific values of the corresponding 9 indicators, we removed the repeated blank data, normalized and standardized each data, so that the indicators were between 0 and 1, and selected significantly representative cities all over the world for further analysis.

We used the crawler algorithm to crawler the light pollution levels of 10,000 regions on the light pollution map website. The light pollution levels of the United States and its surrounding areas are shown as Figure 1:
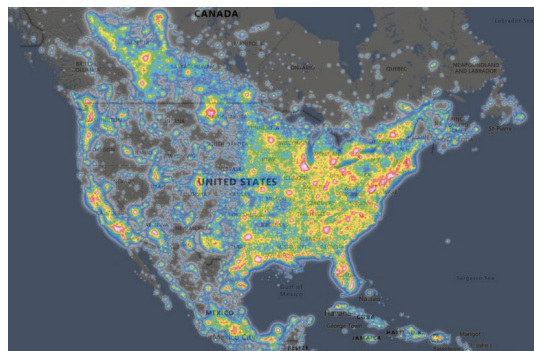

Figure 1: Map of light pollution levels in and around the United States

2.3 Model of Light Pollution Risk Level Assessment

For each region, we search for 9 evaluation indicators of 10000 regions and construct the matrix 10000×9 dimensional matrix x, let the feature vector of the ith region$x_i \in R^9$, and$x_i \in X$, the risk level of the region is an integer$y_i \in R^{1.}$ The value range of the rank is 0-9.

For this problem, we use neural networks and establish an expert evaluation network, as shown in Fig. 2. Since there exists a specific function mapping between the nine evaluation indicators we searched and the light pollution risk level$f(x): x \in X \to y \in Y$, according to the approximation theorem of neural networks it is known that a multi layer feed forward network containing enough hidden layer neurons can approximate any continuous function with arbitrary accuracy. Therefore, a god will model the network to function mapping$f(x)$ The simulation approximation is performed. The steps and details of the neural network establishment are as follows.

Expert evaluation network Basic detail setting:

Our neural network architecture is chosen to have an input layer of 9 neurons corresponding to the 9 feature vectors of the i-th regionx_i, the hidden layer is built as a fully connected structure with 20 neurons in hidden layer 1, 100 neurons in hidden layer 2, 30 neurons in hidden layer 3, 9 neurons in hidden layer 4, and the output layer uses$Softmax$ Normalize the predicted values of 9 levels of light pollution in a single region to the interval$[0,1]$ that$Softmax$ The calculation function is as follows.

$$p(y|x) = \frac{\exp(f_y)}{\sum_{c=1}^{9} \exp(f_c)} = Softmax(f)_y \tag{1}$$

In the calculation function, f(y) indicates the calculated value output at the last hidden layer. Softmax turns the output of the neural network into a probability distribution, and the difference between the predicted probability distribution and the actual light pollution level can be measured by the cross entropy objective function. The nonlinear activation of each fully connected layer is selected as the ReLU function. Optimizer select Adam optimizer, the learning rate is set to 0.0001, and the Dropout technique is used to prevent over fitting of the model. For this classification prediction network, cross entropy is selected as the loss function, and the specific calculation formula is as follows.

$$Loss(\hat{y}|y) = -\sum_i p(\hat{y}_i)\log(q(y_i)) \tag{2}$$

Among them, the $\hat{y}$ denotes output forecast categories corresponding to one - hot vector, y denotes formal category label corresponding to one - hot vector.

In the training process, we randomly selected 7000 regional corresponding indexes and light pollution levels as the training set, 3000 regional corresponding indexes and light pollution levels as the test set, set the training algebra as 30, batch size as 128, and selected the model parameter with the highest accuracy of the test set as the final parameter. We froze the trained model parameters and named them Expert Evaluation network for subsequent use.

The basic framework of the neural network established by us is shown in the figure 2:
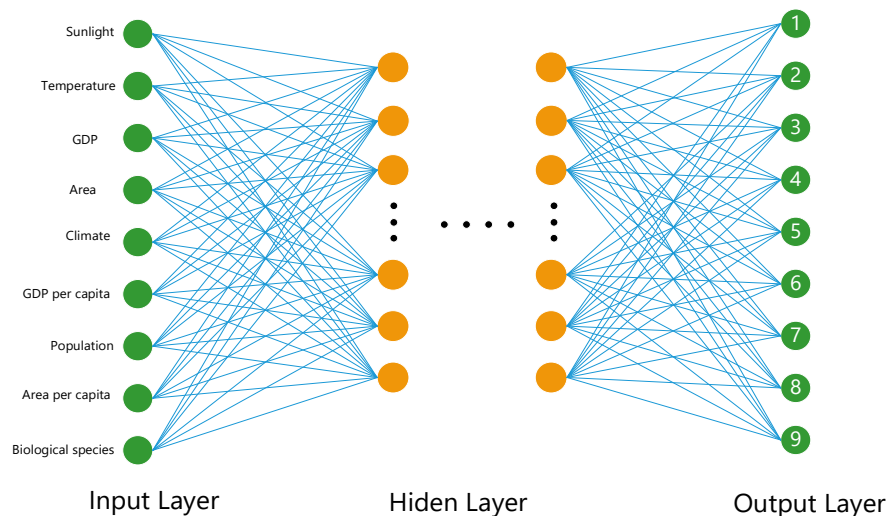

Figure 2: Basic framework of expert scoring network

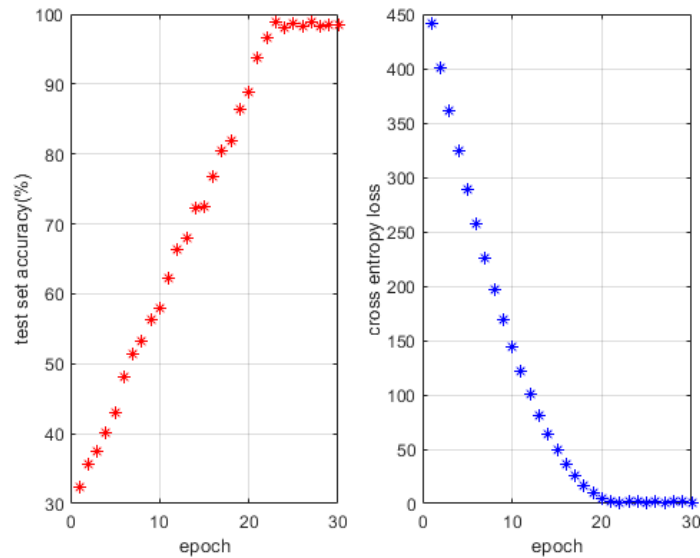In the figure 3, we show the changing process of loss value and test set accuracy during model training.

Figure 3: Variation trend of loss value and test set accuracy with training algebra

Finally, the optimal accuracy of our expert evaluation model I on the test set is 98.9481%, and the minimum loss value is 0.8715. The above results fully prove that our model performs well under the given index distribution and has relatively strong generalization ability.

3. Evaluate the Specific Application of the Model

According to the requirements of the topic, we construct the expert evaluation network I formed by the nervous system, import the data, and use the cross entropy loss function to output the loss. At the same time, experts evaluated the risk level corresponding to the maximum Softmax by Neti score to determine the risk level in the following four areas. In the case of the same climate, sunshine duration and other factors, combined with the population, area, GDP and other indicators of the region as the basis for analysis, we selected the most statistically representative 5 conservation areas, 5 rural communities, 5 suburbs, and 5 urban areas from the existing data set, and obtained the light pollution risk levels as figure 4.
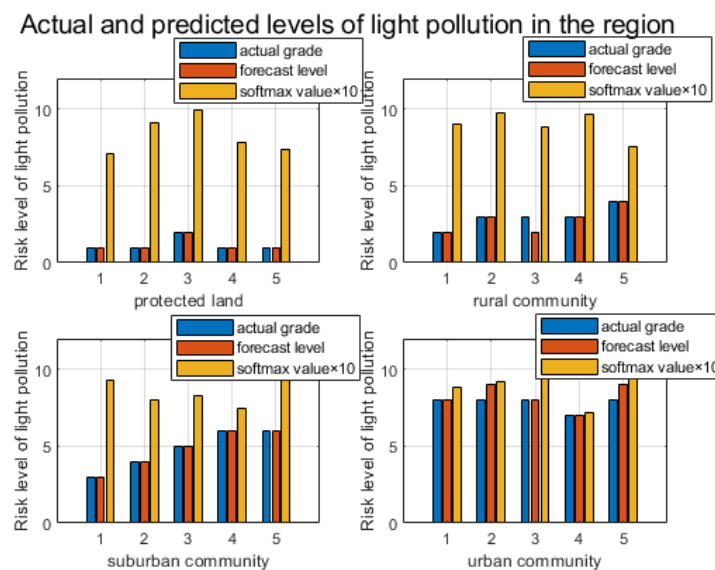

Figure 4: Bar chart comparing the predicted light pollution level with the actual level

Protected land locations: Combined with the model evaluation, we found that protected land locations had the lowest light pollution risk level of 1-2. Protected land, which is protected by legal administration, is mainly used for plant afforestation and ecological restoration. Because it is far away from residential and commercial areas, it lacks the environment and needs formed by light pollution, such as a large population and economic development needs. The land is protected and undeveloped from the interference of external light sources, so the risk level of light pollution in this region is low.

Rural communities: According to the evaluation results of the model, the light pollution risk level of rural communities is graded 3-4, which belongs to mild light pollution. Rural communities with lower economic development cover a large area and are mainly engaged in agriculture. Meanwhile, it has small population and relative scattered housing. Despite of weak artificial light intensity, there still exists light pollution due to lighting needs at night.

Suburban communities: We applied light pollution situation of suburban communities to our model and found that the light pollution level of suburban communities ranged from 3-5, which was higher than that of rural communities. As the suburbs urbanize, the upper and middle classes migrate to the suburbs, providing a demand force for light pollution. Suburban communities are mainly residential areas, where a large number of people gather at night. As a result, the large area of light at night brings harm to people's health and life, resulting in serious light pollution.

Urban communities: According to the evaluation results of the model, the light pollution level of urban communities is the most serious, ranging from 5 to 9. Combined with the characteristics of urban communities, such as rich economic activities at night, wide geographical scope, large number of dense population, combined with the high level of urban economic development, low greening, and convergence of night lights, a large area of light pollution is formed. Meanwhile, the large number of victims is difficult to solve. The light pollution risk level of urban communities is high.

4. Conclusion

Based on authoritative data and comprehensive analysis, using the back propagation algorithm to optimize and integrate the data to build a neural network to assess the light pollution risk level. Based on the analysis of measurement criteria, the risk level of light pollution in different locations in a region is mainly affected by population, area and GDP when the sunshine time is the same.

According to the results obtained by the expert online evaluation model, the protected land location has fewer human intervention factors and the minimum degree of light pollution. Rural communities are mainly engaged in agriculture, and the degree of light pollution is mild. The urbanization phenomenon of suburban communities is obvious, and the phenomenon of light pollution is more serious. The economic development level of urban communities is high, and the level of light pollution is the most serious, ranging from level 5-9.

References

[1] Zhang Shijun. Light pollution-a new type of environmental pollution [J]. Urban problem, 2004 (06): 31-34+42.

[2] Chen Chaonan. Research on light pollution and countermeasures in urban night lighting [J]. Low Carbon World, 2021, 11(5): 2.

[3] Liu Tianyuan. Light pollution control: a two-way examination of domestic practice and foreign experience [J]. Journal of Northwestern University for Nationalities: Philosophy and Social Science Edition, 2022(1): 8.

[4] Hao Qingli, Ren Zhuofei, Liu Gang, etc. Under the coercion of light and noise pollution, the risk evaluation of urban ecological plaques and birds [J]. Journal of Ecology, 2022, 42 (06): 2186-2201.

[5] Xiao-Ting W, Ai-Hong L, Pei-Long S. Research on Regional Financial Risk Measurement and Evaluation Based on Macro Balance Sheet:Taking Shanxi Province as an Example[J].On Economic Problems, 2019.

[6] Li Z, Huang S, Xu HW, Wu Duxiong, Zhou Gang, & Wu Chunhai. (2020). A study on the status of light pollution standards for urban night scenes. China Lighting Appliances.

[7] Jiang Yixin, Shi Yingying, Gao Shuo, etc. The impact of artificial noise, artificial light and road kill on amphibians [J]. Bio diversity, 2023, 31 (03): 220-232.