# Research on the Light Pollution Risk Index Model Based on Entropy Weight Method and TOPSIS

Hongyang Li [1, *, #], Liu Bao [1, #], Jiaman Zhu [2, #]

[1] Department of Applied Mathematics and Informatics, Belarusian State University Minsk

[2] Department of Mechanics and Mathematics, Belarusian State University Minsk

* Corresponding Author Email: a870406667@gmail.com

**Abstract.** This paper applies entropy weight method and TOPSIS ideal solution method, multiple linear and linear regression, the least squares method and sensitivity analysis to the light pollution problem, and firstly, it discusses how to measure light pollution risk levels using multiple factors, and then explores different influencing factors for specific areas by controlling for variables and conducting a regional analysis of four urban areas in Beijing, while verifying the validity of the indicator model in an example. Next, this paper develops and tests three intervention strategies for light pollution risk levels, establishes a correlation model, and uses case studies to analyze what reasonable and effective interventions people should take. Finally, the successful implementation of this strategy may increase people's knowledge and environmental awareness of light pollution through leaflets to educate the population and thus reduce the possibility of light pollution flooding, in order to construct light pollution risk indicators, the evaluation criteria were determined based on the TOPSIS method combined with urban light pollution, normalized matrices, and a normalization method was used to standardize the score of each indicator into a relative score with weighted weights.

**Keywords:** Light pollution index; entropy weight method; TOPSIS ideal solution.

## 1. Introduction

Light pollution is a complex issue that affects the environment, health and safety. The effects of light pollution may vary depending on the characteristics of the site, and the potential impact of any intervention strategy should be tailored to the specific site [1]. While reducing the negative impacts of light pollution, a balance must be found between the positive and negative impacts of artificial light. A broadly applicable metric should be developed to determine the level of light pollution risk at a site [2].Based on the entropy weight method, the optimal and inferior solutions are obtained through the ideal solution method, so as to select the method to evaluate the light pollution risk indicators and provide a theoretical basis for subsequent interventions and metrics. The light pollution indicators are applied to four different types of locations and the results are interpreted [3]. the correlation analysis of least square method and entropy weight method is carried out, and the following conclusions are obtained through data modeling: light pollution risk level model [4] It is more effective when applied to different regions with great differences, and after removing the influence of geographically related factors [5], the risk index of light pollution in different regions located in similar geographical locations will be affected by light intensity and population density. More significant [6].

## 2. Constructing a light pollution risk index model based on topsis and entropy weight method

The TOPSIS method was used to construct a light pollution risk index model by combining urban light pollution TOPSIS operations An evaluation system consisting of n evaluation objects and m evaluation indicators isestablished X:

$$\begin{bmatrix} x_{11} & x_{12} & x_{13}\cdots\cdots x_{1n} \\ x_{21} & x_{22} & x_{23}\cdots\cdots x_{2n} \\ x_{31} & x_{32} & x_{33}\cdots\cdots x_{3n} \\ \cdots & & \\ x_{m1} & x_{m2} & \cdots\cdots \quad x_{mn} \end{bmatrix} \tag{1}$$

Standardization of the evaluation matrix For the matrixX = (xij)n×m, let the minimum value of the yi-th column be and the maximum value be then the maximum-minimum normalization of the matrix is:

$$x'_{ij} = \frac{x_{ij} - a_j}{b_j - a_j}\left(1 \le i \le m, 1 \le j \le n\right) \tag{2}$$

The evaluation index system zij is established according to the given conditions of the problem

$$z_{ij} = \frac{x_{ij}}{\sum_{i=1}^{n} x_{ij}^2} \tag{3}$$

Calculate the information entropy. For each indicator j, calculate its information entropy $E_j$

$$E_j = -\sum_{i=1}^{n} 1 \frac{x_{ij} \ln\left(x'_{ij}\right)}{\ln(n)} \tag{4}$$

Where, x′denotes the standardized score of the i-th evaluation object on the j-th evaluation index, n denotes the number of evaluation objects [7], and ln denotes the natural logarithm Then, the weights of each indicator are calculated as follows:

$$w_j = \frac{1 - E_j}{m - \sum_{j=1}^{m} E_j} \tag{5}$$

Calculate the weighted normalization matrix. Multiply the normalization matrix x′with the weight vector W = (w1, w2, ⋯,wm) to obtain the weighted normalization matrixY = (yij)n×m:

The positive and negative ideal solutions are calculated to measure the closeness of each evaluation object to the ideal solution [8]. The maximum uj + and minimum values u j − are calculated respectively. The formulae for the positive ideal solution A + and the negative ideal solution A − are as follows [9]

$$A^+ = \left(u_1^+, u_2^+, \cdots, u_m^+\right)$$
$$A^- = \left(u_1^-, u_2^-, \cdots, u_m^-\right) \tag{6}$$

Calculate the distance between the evaluation object and the positive ideal solution $d_i^+$ and the negative ideal solution $d_i^-$ , respectively

$$d_1^+ = \sqrt{\sum_{j=1}^{m} 1 * w_j \left(u_{ij} - u_j^+\right)^2}$$
$$d_1^- = \sqrt{\sum_{j=1}^{m} 1 * w_j \left(u_{ij} - u_j^-\right)^2}$$

(7)

Calculate the relative proximity si of each evaluation object (i=1, 2... n).

$$S_i = \frac{d_i^-}{d_l^- + d_i^+} \tag{8}$$

## 3. Evaluation criteria based on TOPSIS method combined with urban light pollution

A basic statistical analysis of the data given in the file "data.xlsx" reveals that the data affecting light pollution are divided into positive and negative correlations [8], including some data values that have actual significance but are identified as meaningless by the prediction model, while these variable factors are considered as redundant elements, thus affecting the The specific accuracy of light pollution indicators [9].

The paper use the TOPSIS method to determine the evaluation criteria in combination with urban lightpollution, and select the appropriate evaluation indicators according to the specific situation of the problem.

Based on these evaluation criteria, an evaluation matrix X can be constructed [10], where each row represents an evaluation object andeach column represents an evaluation index.

$$\begin{bmatrix} x_{11} & x_{12} \cdots\cdots x_{1m} \\ x_{21} & x_{22} \cdots\cdots x_{2m} \\ \vdots & \vdots \cdots\cdots \vdots \\ \vdots & \vdots \cdots\cdots \vdots \\ x_{n1} & x_{n2} \cdots\cdots x_{nm} \end{bmatrix} \tag{9}$$

Where $x_{ij}$ denotes the original score of the i-th evaluation subject on the j-th evaluation index

### 3.1 Assessing the impact level of light pollution based on entropy weighting method

Calculate the weighted normalization matrix. Multiply the normalization matrix x′with the weight vector W = ($w_1$, $w_2$, ⋯, $w_m$) to obtain the weighted normalization matrix Y = (yij) n×m:

Where W denotes the number of evaluation indicators as shown in Table 1 and Fig 1

**Table 1.** Weights of Each Indicator.

| Indicators | Weights (%) |
|---|---|
| **LI** | 12.522 |
| **PD** | 13.212 |
| **GDP** | 23.631 |
| **SH** | 8.724 |
| **ACF** | 16.863 |
| **AA** | 3.408 |
| **RD** | 5.259 |
| **GS** | 7.782 |
| **SQM** | 8.599 |

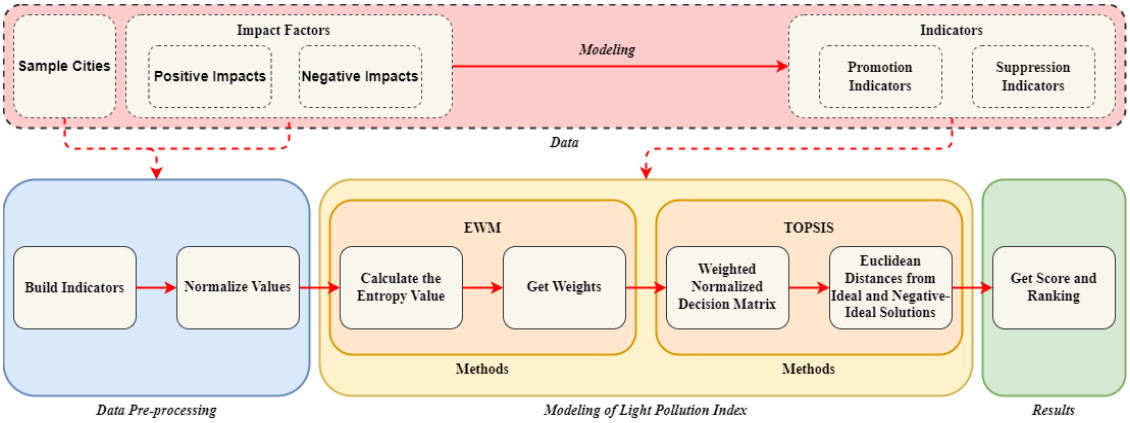



**Fig 1.** Model process flow char.

### 3.2 Obtain and rank the composite index values

The sorting results are shown in the table 2 below.

**Table 2.** Sorting results.

| *City* | *Score* | *Ranking* |
|---|---|---|
| **Los Angeles** | 0.65514473 | 1 |
| **Tokyo** | 0.61129207 | 2 |
| **Buenos Aires** | 0.52720932 | 3 |
| **Milan** | 0.47218403 | 4 |
| **Moscow** | 0.46721745 | 5 |
| **London** | 0.45149145 | 6 |
| **Singapore** | 0.41793095 | 7 |
| **Melbourne** | 0.40450584 | 8 |
| **Montréal** | 0.40435437 | 9 |
| **Bogotá** | 0.3623023 | 10 |
| **Cape Town** | 0.32153022 | 11 |
| **Auckland** | 0.3152001 | 12 |

### 3.3 Contrast Model

By eliminating the indicators LI, SQM, which are derived from direct measurements of regionallight intensity and night sky brightness by observational instruments[11], the risk level of light pollution in the city is modeled using the same method and the results of the two models are compared as shown in Table 3

**Table 3** two models are compared

| *City* | *Score* | *Ranking* |
|---|---|---|
| **Tokyo** | 0.64120238 | 1 |
| **Los Angeles** | 0.63917958 | 2 |
| **London** | 0.46945016 | 3 |
| **Buenos Aires** | 0.44839136 | 4 |
| **Singapore** | 0.41237713 | 5 |
| **Moscow** | 0.40973047 | 6 |
| **Melbourne** | 0.39173973 | 7 |
| **Bogotá** | 0.38831616 | 8 |
| **Milan** | 0.36667933 | 9 |
| **Cape Town** | 0.35538781 | 10 |
| **Montréal** | 0.32558106 | 11 |
| **Auckland** | 0.24456196 | 12 |

### 3.4 Model ideas and analysis

From the comparison of the two models the following conclusions are drawn in this question.The entropy weighting method shows that in the regional light pollution risk level model, the economic indicator GDP has the largest weight, which means that the more developed the economic level of a region is, the greater the light pollution it may cause [12]. The top three cities are still the top three in the group in terms of GDP under the influence of other indicators, which shows that in today's world, the development of economic level brings a series of environmental problems, including but not limited to the deepening risk of being exposed to light pollution.

By comparing the processed light intensity with the above two groups of light pollution risk indices, this question can be seen from the figure that it is unwise to use only the intensity of city lights monitored by remote sensing satellites as a measure of light pollution levels, and it is clear that the true light pollution levels are influenced by many additional factors. It can also be found from the image that Europe and North America are in the top of the light intensity index[13], as this index is the most intuitive reflection of the intensity of urban lights at night, which indicates that developed areas in Europe and America generally have a wide range of lights on at night, high intensity and long hours, the reason for this phenomenon may be that the electricity consumption habits in these areas are influenced by the cultural and living habits from.

## 4. Analysis of Regression Models in Regional Classification Features

In order to further verify the validity of the risk level model. The regression model was added to combine four different regions located within the same city added the following three indicators. Meanwhile, since the four regions to be compared are located in the same city, the previous indicators of absolute carbon footprint, sunshine hours and the percentage of cloudy days in a year are highly similar and therefore are no longer considered as indicators. Instead, the new indicator fully considers the selected characteristics of the regions, and the average number of lights at night in the designated region is estimated from the lighting situation of the region, as an important criterion to judge whether it has excessive lighting[14]. The question also investigates the number of traffic accidents occurring in each region, and explores the possible impact of dazzling light caused by artificial light on driving safety. It is worth noting that the number of traffic accidents caused by dazzling light could not be completely filtered, so the question adopted a narrowed-down sequential countingmethod to arrive at a specific number. In order to verify the effect of artificial light on the crimeindex, the number of criminal cases was also used to investigate the level of light pollution risk,which will be used in the subsequent model as one of the criteria to measure the extent to whichlighting devices and related regulations should be improved.

### 4.1 Model ideas and analysis

The correlation coefficient is a statistical indicator used to determine the closeness of the correlation between two variables under the condition of linear correlation.It is generated by combining three indicators: the covariance of two variables (X, Y), the standard deviation of the X variable and the standard deviation of the Y variable, and its basic formula is as follows

$$r(x,y)=\frac{\sigma^2(x,y)}{\sigma(x)*\sigma(y)} \tag{10}$$

Where, r(x, y) denotes the correlation coefficient, σ2(X, Y) denotes the covariance of the two variables, σ(X) denotes the standard deviation of the variable X,σ(Y) denotes the standarddeviation of the variable Y

The slope variance is calculated as follows

$$\sigma^2(x,y)=\frac{\sum(X-\bar{X})(Y-\bar{Y})}{n}$$

(11)

Where $\bar{X}=\frac{\sum x}{n}$ is the arithmetic mean of the variable X, $\bar{Y}=\frac{\sum Y}{n}$, is the arithmetic mean of Y Establish the Pearson correlation coefficient formula based on the above equation

$$r(x,y)=\frac{\sum_{i=1}^{n}\left(X_i-\bar{X}\right)\left(Y_i-\bar{Y}\right)}{\sqrt{\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2}\sqrt{\sum_{i=1}^{n}\left(Y_i-\bar{Y}\right)^2}} \tag{12}$$

Where the linear correlation coefficient r indicates the closeness and direction of the linear correlation between the two variables, the larger |r| is larger, the stronger the correlation betweenX and Y, and vice versa, the weaker

Then the Kendall correlation coefficient formula is established:

$$k_{t-b}=\frac{\sum\sum_{i<j}\xi\left(X_i,X_j,Y_i,Y_j\right)}{\sqrt{\frac{1}{2}n(n-1)-T_{x^*}}\sqrt{\frac{1}{2}n(n-1)-T_y}} \tag{13}$$

### 4.2 Model Testing

In order to validate the above model, four different types of regions are set as follows, and four specific locations within the same city of Beijing with significant differences in relevant indicators and clear data are selected for the study. Under the control variables method, the fourselected regions have clear geographical boundary distinctions, the same human background, and the same climatic conditions.

Select some appropriate evaluation indicators to describe the degree of light pollution, such as: light level, night brightness, etc. Here, we choose 9 indicators for evaluation: citing excel file related data as shown in Fig 2

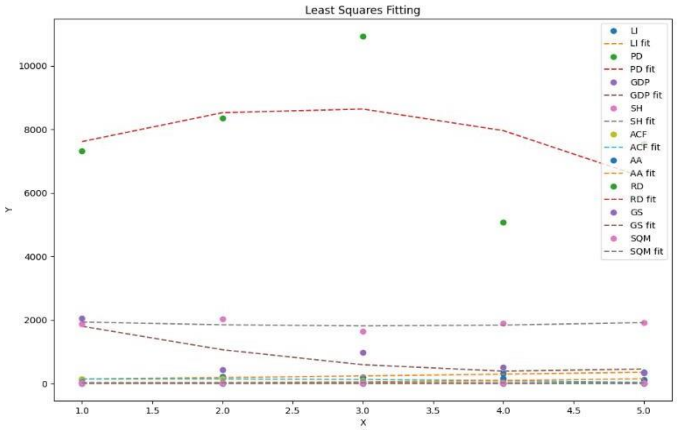



**Fig 2.** Fitting images using least squares.

### 4.3 Weight distribution

The importance of different indicators in the evaluation may be different, so a weight needs to be assigned to each indicator. Here we adopt the data assignment method, and based on the collected data, the indicators LI, LN, PD, GDP, CC, TA, AA, and GS are derived in order, respectively.

The ranking of model results is unchanged. The new weights are obtained as follows Table 4

**Table 4.** New weights.

| *Indicators* | *Weights (%)* |
|---|---|
| **LI** | 20.802 |

| | |
|---|---|
| **LN** | 16.762 |
| **PD** | 17.533 |
| **GDP** | 9.383 |
| **CC** | 9.183 |
| **TA** | 9.98 |
| **AA** | 4.917 |
| **GS** | 11.441 |

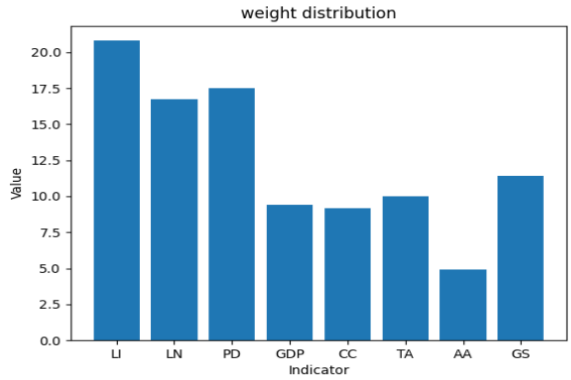



**Fig 3.** Weight Assignment in Python Visual Analytics Histogram Specific Form.

## 5. Conclusions

The light pollution risk index is a comprehensive indicator to measure the impact of light pollution on human and ecological environment. In order to construct the light pollution risk index model, entropy weight method and TOPSIS model can be introduced for improvement. The entropy weight method is a multi-criteria decision making method that can be used to unify different units or magnitudes of different indicators and transform them into dimensionless indicator weights. Specifically, the entropy value of each indicator can be calculated to determine the weight of each indicator, and then the weighted average method is used to aggregate each indicator into a composite indicator. TOPSIS model, on the other hand, is an evaluation method used to determine the relative merit and comprehensive ranking of each indicator. By comparing the positive ideal solution of each indicator with the negative ideal solution, the proximity of each indicator can be calculated, and then the proximity of each indicator is combined to obtain the composite ranking.

Combining the entropy weight method and TOPSIS model can build a light pollution risk index model, which can make a comprehensive assessment of light pollution risk levels in different cities, and then formulate corresponding light pollution control measures.

To promote this model, we need to collect relevant data, including light pollution index, light pollution source type and intensity, population density, green area and other factors of each city, and then pre-process and analyze the data, including data cleaning, data standardization, calculation of each index weight and comprehensive index. Finally, the cities can be ranked according to the comprehensive index to guide the light pollution control work.

## References

[1] G. Admiranto, R. Priyatikanto, S. Maryam, Ellyyani and N. Suryana, “Preliminary Report of Light Pollution in Indonesia Based on Sky Quality Observation”, Journal of Physics: Conference Series, 2018.

[2] Philippe Deverchère, Sébastien Vauclair, Gonzague Bosch, Sylvain Moulherat & Jérémie H. Cornuau, “Towards an absolute light pollution indicator”, Scientific Reports, 2022.

[3] Susan Mander, Fakhrul Alam, Ruggiero Lovreglio, Melanie Ooi, “How to measure light pollution - a systematic review of methods and applications”, Sustainable Cities and Society, 2023.

[4] Ziyan Yan, Minghong Tan, “Changes in light pollution in the Pan-Third Pole's protected areas from 1992 to 2021”, Ecological Informatics, 2023.

[5] Miao Cao, Ting Xu, Daqiang Yin, “Understanding light pollution: Recent advances on its health threats and regulations”, Journal of Environmental Sciences, 2023.

[6] Aguilera Moisés A.; González María Gracia, “Urban Infrastructure expansion and artificial light pollution degrade coastal ecosystems, increasing natural-to-urban structural connectivity”, Landscape and Urban Planning, 2023.

[7] Hadi A. Khorshidi, Uwe Aickelin, “Multicriteria Group Decision-Making Under Uncertainty Using Interval Data and Cloud Models”, Journal of the Operational Research Society, 2020.

[8] Guilherme D. Pelegrina, Leonardo T. Duarte, João M. T. Romano, “Muticriteria decision making based on independent component analysis: A preliminary investigation considering the TOPSIS approach”, Lecture Notes in Computer Science, 2018.

[9] Guilherme Dean Pelegrina, Leonardo Tomazeli Duarte, João Marcos Travassos Romano, “Application of independent component analysis and TOPSIS to deal with dependent criteria in multicriteria decision problems”, Expert Systems with Applications, 2019.

[10] Chun-Ho Chen, “A Novel Multi-Criteria Decision-Making Model for Building Material Supplier Selection Based on Entropy-AHP Weighted TOPSIS”, Entropy, 2020.

[11] Geetanjali Rathee; Sahil Garg; Georges Kaddoum; Bong Jun Choi, “Decision-Making Model for Securing IoT Devices in Smart Industries”, IEEE, 2021.