Control Light Pollution by Genetic Algorithms

Hangrui Fan [a, #], Shiwen He [b, #]

Southern University of Science and Technology, Shenzhen 518055, China

[a] 3336978717@qq.com, [b] FarmerKN2019FK@outlook.com

Hangrui Fan, Shiwen He are Co-first author

Abstract. Based on Light pollution, the team first investigated the important factors that may lead to light pollution, and is divided into three factors, respectively, they are cultural factors, geographical factors and development factors, each and there are several secondary factors, they are GDP, air pollution, river coverage, vegetation coverage, residential coverage, industrial coverage, population quality and other factors. After obtaining these key factors affecting light pollution, our group conducted quantitative modeling based on the two-level AHP hierarchical analysis and derived the weights of these key factors. Noted that because the AHP model largely depends on expert opinion, has a lot of subjectivity, so our group further introduced the grey correlation analysis to more accurately and our assessment results and actual situation is very well, such as protection areas, rural areas for almost no light pollution or light pollution, and urban areas or urban areas are severe or moderate light pollution. Based on the above work, our group adopted the multi-target gene optimization method to obtain the optimal set of strategies, and greatly reduced the light pollution level under the premise that it did not affect the development.

Keywords: Light Pollution; AHP; Grey Correlation Analysis; Genetic Algorithm; Statistics Analysis.

1. Introduction

Light pollution is used to describe any excessive or poor use of artificial light. With the increasing use of artificial light, light pollution has become an environmental problem that needs urgent attention. Light pollution is widespread in many places, has negative impacts on both environment and organism. The deluge of light can also confuse human circadian rhythms, even leading to traffic accidents. Therefore, we should formulate reasonable evaluation criteria and propose effective intervention measures.

However, some lighting is necessary, such as city road lighting. Too little light at night leads to a rise in crime. And the suburbs due to the sparse population, the demand for light is greatly reduced. Instead, the natural areas need to be protected from artificial light.

So, a broadly applicable metric of light pollution should consider various factors such as the location's level of development, population, biodiversity, geography, and climate. Similarly, the intervention strategies also must be tailored to a specific location.

2. Models

2.1 Light Pollution Risk Level Evaluation Model Via AHP

2.2 AHP Model

AHP is the core model based on quantitative importance evaluation. In order to obtain the importance of the relationship between massive factors, we introduced the AHP model, integrated several different correlation indicators, obtained the final risk level through AHP.

2.3 Investigation of Influencing Factors

Through extensive investigation, we give the following indicators that will affect light pollution:
Development factors (C_1): GDP **(P_1),** The degree of air pollution **(P_2)**

Humanistic factors (C_2): Residential area coverage **(P_3),** Industrial area coverage **(P_4),** Population quality **(P_5)**

Geographical factors(C_3): Vegetation coverage rate **(P_6),** River coverage **(P_7)**

From the perspective of the geographical environment, the vegetation coverage rate can greatly increase the diffuse reflection and reduce the light pollution. And rivers will produce light pollution due to mirror reflection. And the air pollution will increase the scattering of light, aggravating the light pollution. From the perspective of humanity, the increase of the coverage rate of residential areas leads to light pollution with environmental deterioration, while the increase of GDP will drive the green industrial chain to reduce environmental pollution and thus reduce light pollution. The increase of population quality will make people consciously protect the environment and reduce light pollution, but the increase of the coverage rate of industrial areas will aggravate air pollution and increase light pollution.

2.4 Two-level Model to Metric the Light Pollution Risk Level

We now use the metrics obtained from the above survey to build a quantitative two-level model to evaluate the light pollution level, as shown below. Each indicator in the figure corresponds to 3.2 indicators respectively.

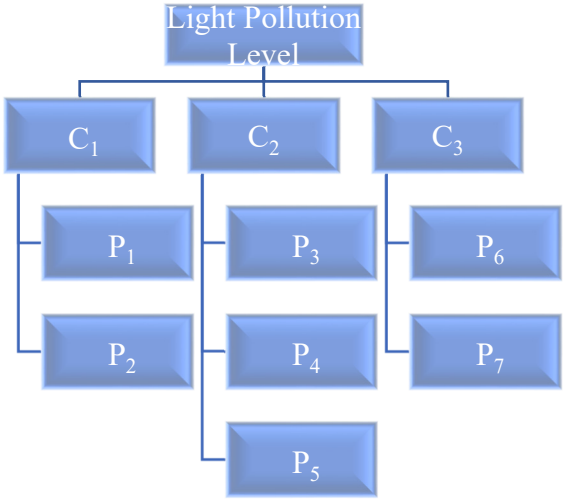


Fig 1. Two-level AHP model

Each secondary level can not only determine its parent project, but also act on different primary levels. The specific relationship is as follows. The secondary level is briefly described in Pi:

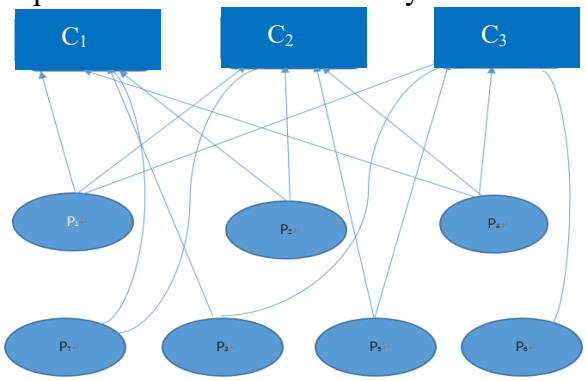


Fig 2. Two-level relations

Then we Establish a judgment matrix:

By reviewing the survey, we obtain the following weights of factors,

C1:(5) C2:(3) C3:(1) P1:(6) P2:(2) P3:(2) P4:(4)

P5:(3) P6:(3) P7:(2).

Pairwise comparisons as to various factors in the hierarchy were made, and the judgment matrix was established based on the index factors and judgment matrix elements of Q 1

The following is the judgment matrix given from this table:

1) Hierarchical single ranking and consistency inspection

For the judgment matrix, solutions are obtained using the MATLAB matrix:

G-C:

$$W=\begin{pmatrix}0.5556\\0.3333\\0.1111\end{pmatrix} \quad \lambda_{max}=3.00 \quad CI=0 \quad RI=0.58 \quad CR=0<0.1 \tag{1}$$

C1-P:

$$W=\begin{pmatrix}0.3750\\0.1250\\0.1250\\0.2500\\0.0000\\0.0000\\0.1250\end{pmatrix} \quad \lambda_{max}=5.00 \quad CI=0 \quad RI=1.12 \quad CR=0<0.1 \tag{2}$$

C2-P:

$$W=\begin{pmatrix}0.3158\\0.1053\\0.1053\\0.2105\\0.1579\\0.0000\\0.1053\end{pmatrix} \quad \lambda_{max}=6.00 \quad CI=0 \quad RI=1.12 \quad CR=0<0.1 \tag{3}$$

C3-P:

$$W=\begin{pmatrix}0.3010\\0.0000\\0.1003\\0.1919\\0.1505\\0.1505\\0.1057\end{pmatrix} \quad \lambda_{max}=6.0092 \quad CI=0.00184 \quad RI=1.24 \quad CR=0.00148<0.1 \tag{4}$$

2) Hierarchical overall ranking and consistency inspection

For computational consistency, the correction factor RI is introduced, where the order is the order of the judgment matrix.

Table 1. CR correlations

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	0	0.58	0.9	1.12	1.24	1.32	1.41	1.45	1.49	1.52	1.54	1.56	1.58	1.59

The resulting eigenvectors are placed in the table in column form and the product of Cj and Pi and taken as weight factors. The table is as below:

Weight factor consistency test:

$$CI=\sum_{j=1}^{3}a_j\,(CI)_j=0.000204424 \tag{5}$$

$$RI=\sum_{j=3}^{3}a_j(RI)_j=1.133332 \tag{6}$$

$$CR=\frac{CI}{RI}=0.0001803743299<0.1 \tag{7}$$

Table 2. Factor weights

Hierarchy P	C1	C2	C3	Overall ranking ρ
	0.5556	0.3333	0.1111	
P1	0.3750	0.3158	0.3010	0.3470
P2	0.1250	0.1053	0.0000	0.1045
P3	0.1250	0.1053	0.1003	0.1157
P4	0.2500	0.2105	0.1919	0.2304
P5	0.0000	0.1579	0.1505	0.0693
P6	0.0000	0.0000	0.1505	0.0167
P7	0.1250	0.1053	0.1057	0.1162

The degree of inconsistency is within the tolerance range.

At this point, the AHP hierarchical analysis model was initially established. From the AHP model, we get seven influencing factors, which will be used in the score evaluation model:

Table 3. Influencing factors

Pi	P1	P2	P3	P4	P5	P6	P7
ρi	0.0167	0.1162	0.1157	0.2304	0.0693	0.1045	0.347

3. The Evaluations of Four Areas Example

Since the weight determined by the above AHP model is still subjective to a certain extent, in order that my own subjective judgment has a great influence on the results, the gray association analysis method is introduced to calculate the more objective weight.

Gray association analysis (GRA) is a multi-factor statistical analysis method developed based on gray theory. It is a method of quantitative analysis and comparison of the development trend of dynamic process of a system. It on the sample data of each factor describes the degree of correlation of each factor. This method has the characteristics of simple principle, easy to master, simple calculation, and no special requirements for data distribution types and related types among various variables. If the sample data reflects that the relative changes of the two factors in the development process of the system, the correlation between the two factors is large: on the contrary, the correlation degree is small. The application of gray correlation has been greatly applied in the social and economic fields, such as the regional economic advantage analysis, industrial structure adjustment, project comprehensive evaluation and other aspects have achieved good results.

The calculation formula of this method is as follows:

There are *m* solutions to solve the problem and *n* evaluation indicators for each scheme, so the decision matrix is as follows:

$$X = \begin{bmatrix} lx_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix} \tag{8}$$

There are m solutions to solve the problem and n evaluation indicators for each scheme, so the decision matrix is as follows:

$$b_{ij} = \frac{x_{ij} - \min(x_j)}{\max(x_j) - \min(x_j)} \tag{9}$$

$$b_{ij} = \frac{\max(x_j) - x_{ij}}{\max(x_j) - \min(x_j)} \tag{10}$$

The standardized matrix is weighted by the index, and the gray correlation coefficient is calculated:

$$\xi_i(k) = \frac{\min\limits_i \min\limits_k |c_0(k) - x_i(k)| + \rho \max\limits_i \max\limits_j |c_0(k) - c_i(k)|}{|c_0(k) - x_i(k)| + \rho \max\limits_i \max\limits_j |c_0(k) - c_i(k)|} \tag{11}$$

To obtain the gray correlation matrix:

$$\xi = \begin{bmatrix} l\xi_{11} & \xi_{12} & \cdots & \xi_{1n} \\ \xi_{21} & \xi_{22} & \cdots & \xi_{2n} \\ \vdots & \vdots & & \vdots \\ \xi_{m1} & \xi_{m2} & \cdots & \xi_{mn} \end{bmatrix} \tag{12}$$

The thermal map of the correlation coefficient matrix is obtained as:



Fig 3. Heat map

The abscissa represents seven areas, the ordinate indicates seven indicators, and the color depth is the correlation degree.

By entropy method, the weight of each factor is below:

Table 4. The weight of each factors

	WEIGHT
P1	0.226415094
P2	0.041509434
P3	0.358490566
P4	0.030188680
P5	0.256603774
P6	0.049056604
P7	0.037735850

Table 5. Factor weights

Factors	**P1**	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**
Protected Land	0.42	4.55	27.85	70.69	27569.49	26.15	25.83
Rural Community	0.44	8.51	12.58	75.69	117392.36	31.79	52.36
Suburban community	0.49	0.19	0.54	80.25	41045.63	35.19	15.09
Urban community	0.35	15.14	5.67	64.57	3385.07	10.25	2.01

$$Y = \sum_{i=1}^{7} Y_i^* * \rho_i = 1.258 \tag{13}$$

Table 6. Light pollution risk level

Y score	0-2	2-4	4-7	7-10
Light Pollution Level	Almost no light pollution	Mild light pollution	Moderate light pollution	Heavy light pollution

From the scores and the ratings in the table above, we can simply calculate the light pollution risks of four areas:

Table 7. Light pollution results

Protected land	Almost no light pollution (1.258)
Rural community	Mild light pollution (2.764)
Suburban community	Moderate light pollution (6.325)
Urban community	Heavy light pollution (8.377)

4. Optimal Intervention Strategy Selection Via Genetic Algorithm

4.1 Genetic Algorithm

The parameters of the solved genetic algorithm are set as follows:

Population size: M=20; maximum number of generations: G=100; crossover rate: =1; variation rate: =1.0, and the details are showed as follows:

Table 8. Detailed parameters

The concept of biological genetics	The role in the genetic algorithm	The corresponding concept in this topic
individual	solution	A set of operational parameters for the intervention strategy versus the actual impact parameters
chromosome	The code of solution	Serial number for each measurement of the intervention strategy
gene	Characteristics of each component in the solution	An operational variable produced by the intervention strategy
adaptability	The fitness function values	Neural network prediction model
population	A set of solutions selected from the value of the fitness function	An optimized set of operational parameters for the intervention strategy
copulation	The process of generating a set of new solutions through the mating principle	
variation	A process by which a certain component of the encoding changes	

Specifically, we choose two location which are **Rural community** and **Urban community.** We controlled the degree of light pollution through an intervention strategy. Combined with what was said before, when the strategy is adopted, although the degree of light pollution is reduced, it also needs to sacrifice and develop and other conditions, the purpose of introducing the genetic algorithm is to optimize and balance these two conditions. When we find that it severely affects the development during the selection of the stem prediction strategy, we will randomly change some initial variables of the GA optimization parameters and re-optimize them until the resulting optimization parameters satisfy the optimal state. The following is the convergence curve of our method.

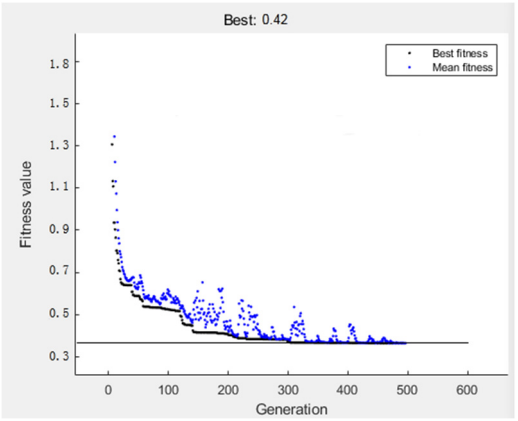


Fig 4. Convergence curve

After 498 generations of inheritance, we obtained the final optimization results, in which light pollution in optimized rural areas decreased from 2.764 to 1.386, while the average decline rate of key indicators such as GDP was only 8%; for urban areas, light pollution decreased from 8.377 to moderate light pollution, and the average decrease rate of key indicators such as GDP was only 13%.

The following figure is the process of the optimization curve of the objective function in the optimization process. It can be seen that the multi-target gene inheritance algorithm we use significantly considers the optimal problem of the two metrics.

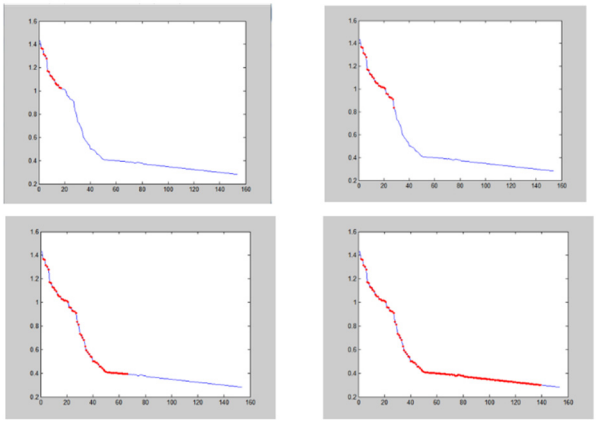


Fig 5. Multi-objective genetic algorithm optimization process

4.2 Statistical Analysis of the Intervention Strategies

As can be seen from the above table, before the optimization, the light pollution intensity in the original area was 2.764-8.377, and the light pollution intensity after the optimization was 1.386-5.436, with an average decrease of 35.11% -49.90%. Therefore, it can be considered that the optimization effect is good.

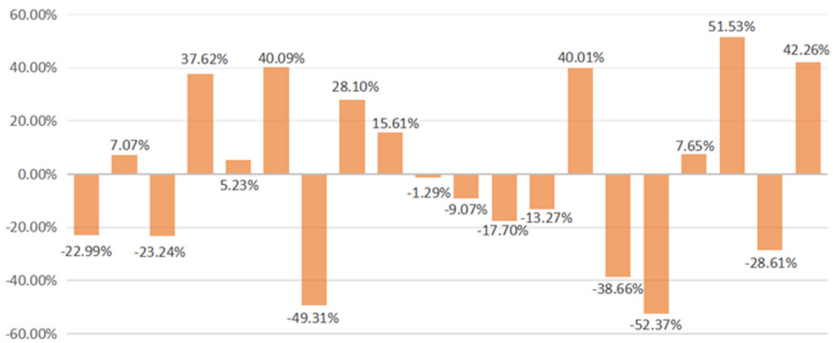


Fig 6. The strategies for different factors

At the same time, we tried to provide more general guidelines on optimal intervention strategies.

We conducted a statistical analysis of some of the main optimized factors and found that factor 2, factor 3, factor 4, factor 5, and factor 7 were larger on average. The rest of the operating group adjustment range is small. In order to control the degree of light pollution, we need to focus on these factors.

We further conducted a sensitivity factor analysis of the intervention operation, which tested the correlation between variables using KMO, whose values were between 0 and 1. The closer the KMO value is to 1, the stronger the correlation between variables, the better the original variables is for factor analysis; When the KMO value is to 0, the weaker the correlation between variables is, the worse the original variable is to the cooperative factor analysis. According to the statistician, factor analysis is not appropriate to perform when KMO values are less than 0.5. The Bartlett's spherical test was used to test whether the individual variable was individually independent. If the variables are independent, the common factor cannot be extracted from them and the factor analysis cannot be applied. In general, the variables are independent when df values are greater than 30 and sig. <0.05 (i. e. p <0.05).

Table 9. Statistic parameters

Elements	Initial eigenvalue			Extract the squared sum and load			Rotary square sum loading		
1	3.003	30.026	30.026	3.003	30.026	30.026	2.964	29.638	29.638
2	2.444	24.438	54.464	2.444	24.438	54.464	2.438	24.384	54.022
3	1.749	17.485	71.949	1.749	17.485	71.949	1.598	15.977	69.999
4	0.839	8.391	80.340	0.839	8.391	80.340	1.034	10.341	80.340
5	0.646	6.463	86.803	-	-	-	-	-	-
6	0.588	5.882	92.685	-	-	-	-	-	-
7	2.739E-005	0.000	100.000	-	-	-	-	-	-

Extraction method: Main component analysis.

The variance interpretation rate table is mainly used to determine how many factors are appropriate to extract, and the variance interpretation rates and cumulative variance interpretation rates for each factor. The larger the variance interpretation rate, the more the factor contains information about the original data, and generally the factor that retains more than 80% of the acceptable variance level. According to the above table, the first four factors are mainly extracted, and the interpretation rate of these four factors is 30.026%, 24.438%, 17.485%, 8.391% respectively., and the interpretation rate of the rotation variance is 29.638%, 24.384%, 15.977%, and the first four factors are greater than 1, the greater the greater the representative factor contribution.

5. Conclusion

We have conducted extensive research and review on the factors that affect Light Pollution in this paper, so as to formulate 3 primary indicators and 7 secondary indicators that affect the pollution risks. Then we obtained the quantitative model AHP according to the indicators via gray association analysis method, so that we can quantitatively analyze the risk level of light pollution in different areas. Besides, We present the optimal strategy screening method based on gene algorithm, which can automatically select the best reduction light pollution strategy for different regional situations, and verify its effectiveness on two regions.

We conducted a statistical analysis of some of the main optimized factors and found that factor 2(The degree of air pollution), factor 3(Residential area coverage), factor 4(Industrial area coverage), factor 5(Population quality), and factor 7(River coverage) were larger on average. The rest of the operating group adjustment range is small. In order to control the degree of light pollution, we need

to focus on these factors. This model describes the assessment of light pollution degree in different regions and the corresponding intervention control.

The model also extensively analyzes the factors affecting light pollution and quantitatively analyzes these factors using mathematical models. After obtaining this model, we can use this model to analyze the degree of light pollution in different situations, such as mild light pollution and severe light pollution. This quantitative model can provide important suggestions for light pollution control.

References

[1] Longcore T, Rich C. Ecological light pollution[J]. Frontiers in Ecology and the Environment, 2004, 2(4): 191-198..

[2] Chepesiuk R. Missing the dark: health effects of light pollution[J]. 2009.

[3] Dominoni D M. The effects of light pollution on biological rhythms of birds: an integrated, mechanistic perspective[J]. Journal of Ornithology, 2015, 156(Suppl 1): 409-418.

[4] Pritchard, Diana J., et al. "Bring the captive closer to the wild: redefining the role of ex situ conservation." Oryx 46.1 (2012): 18-23.

[5] Škvareninová J, Tuhárska M, Škvarenina J, et al. Effects of light pollution on tree phenology in the urban environment [J]. Moravian geographical reports, 2017, 25(4): 282-290.

[6] Al-Harbi, Kamal M. Al-Subhi. "Application of the AHP in project management." International journal of project management 19.1 (2001): 19-27.

[7] Hölker F, Moss T, Griefahn B, et al. The dark side of light: a transdisciplinary research agenda for light pollution policy[J]. Ecology and Society, 2010, 15(4).

[8] Liu S, Yang Y, Forrest J. Grey data analysis[J]. Springer Singapore, Singapore, Doi, 2017, 10(1007): 978-981.

[9] Javed S A, Liu S. Evaluation of outpatient satisfaction and service quality of Pakistani healthcare projects: application of a novel synthetic grey incidence analysis model[J]. Grey Systems: Theory and Application, 2018.

[10] Katoch S, Chauhan S S, Kumar V. A review on genetic algorithm: past, present, and future[J]. Multimedia Tools and Applications, 2021, 80: 8091-8126.

[11] Vose M D. The simple genetic algorithm: foundations and theory[M]. MIT press, 1999.