Method for lake eutrophication levels evaluation: TOPSIS-MCS

Song-Shun Lin a, Shui-Long Shen b,c,d,1, Ning Zhang b,2,* , Annan Zhou c,3

a Department of Civil Engineering, School of Naval Architecture, Ocean, and Civil Engineering, Shanghai jiao Tong University, 800 Dong Chuan Road, Shanghai, Minhang District 200240, China
b MOE Key Laboratory of Intelligent Manufacturing Technology, College of Engineering, Shantou University, Shantou, Guangdong 515063, China
c Discipline of Civil and Infrastructure, School of Engineering, Royal Melbourne Institute of Technology (RMIT), Victoria 3001, Australia
d Shanghai Key Laboratory for Digital Maintenance of Buildings and Infrastructure, School of Naval Architecture, Ocean, and Civil Engineering, Shanghai jiao Tong University, Shanghai 200240, China

ABSTRACT

Monte Carlo simulation (MCS) is applied in the engineering with great fuzziness and uncertainty. Technique for order preference by similarity to an ideal solution (TOPSIS) method is used to deal with multi-criteria decision-making issue. Membership function is used to determine the membership degree of evaluated index. This paper presents the method for lake eutrophication level evaluation. The developed approach merges MCS method, TOPSIS method and membership function. The evaluated results are consistent with real eutrophication level in Lake Erhai, China. Global sensitivity analysis (GSA) is conducted. Results show that potassium permanganate index (CODMn) displays the highest negative correlation with the evaluated results and Secchi disc (SD) performs the highest positive correlation under different errors in measured data. The novelty of this work are: (1) the application of TOPSIS considers surface water environmental quality standards and measured data. Besides, the Monte Carlo simulation method is applied to generate a normal distributed dataset to overcome the errors caused by human and equipment in data collection. The approach is utilized in the article, titled “Approach based on TOPSIS and Monte Carlo simulation methods to evaluate lake eutrophication levels” (Lin et al., 2020) [1].

- Developed approach merges TOPSIS and MCS method.
- It can increase the reliability of evaluated result.

© 2021 The Author(s). Published by Elsevier B.V.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

Article info
Method name: Integrated method based on TOPSIS and Monte Carlo simulation.
DOI of original article: 10.1016/j.watres.2020.116437
* Corresponding authors at: College of Engineering, Shantou University, Shantou, Guangdong 515063, China.
E-mail addresses: shenssl@stu.edu.cn (S.-L. Shen), zhangning@stu.edu.cn (N. Zhang).
1 ORCID: 0000-0002-5610-7988
2 ORCID: 0000-0001-5302-600X
3 ORCID: 0000-0001-5209-5169

https://doi.org/10.1016/j.mex.2021.101311
2215-0161/© 2021 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)
Method name: Integrated method based on TOPSIS and Monte Carlo simulation.

Keywords: Monte Carlo simulation, TOPSIS method, Membership function, Lake eutrophication evaluation

Article history: Received 8 October 2020; Accepted 12 March 2021; Available online 18 March 2021

Specification table

| Subject area: | Environmental Engineering |
|---------------|---------------------------|
| More specific subject area: | Lake eutrophication evaluation |
| Method: | Integrated method based on TOPSIS and Monte Carlo simulation. |
| Name and reference of original method: | Hwang, C.L., Yoon, K. (1981). Multiple Attributes Decision Making Methods and Applications. Springer, Berlin Heidelberg. https://doi.org/10.1007/978-3-642-48318-9 [2]. |
| Resource availability: | |

Method details

Monte Carlo simulation (MCS)

In the process of MCS, determining the range and relative probability in simulations utilizes probability density functions (PDFs) [2,3]. The distribution of common PDFs is normal in this research when the exact distribution of influential factor is not known. The normal distribution is given in Eq. (1):

\[
f(x_{j}^{MCS}, \mu_j, \sigma_j^2) = \frac{1}{\sigma_j \sqrt{2\pi}} e^{-\frac{(x_{j}^{MCS} - \mu_j)^2}{2\sigma_j^2}}
\]  

(1)

\[
\mu_j = \frac{1}{N} \sum_{i=1}^{N} x_{i,j}^{MCS}
\]  

(2)

\[
\sigma_j^2 = \frac{1}{N-1} \sum_{i=1}^{N} (x_{i,j}^{MCS} - \mu_j)^2
\]  

(3)

where \( f(.) \) is the function of the normal distribution. \( x_{i,j}^{MCS} \) is the \( i \)th sample data of the \( j \)th indicator, and \( \mu_j \) and \( \sigma_j \) are the mean and standard deviations of the distribution about \( j \)th indicator. Eqs. (2) and (3) are the mean and standard deviations of these sample data sets regarding on \( j \)th water quality indicators. Five influential factors on lake eutrophication are determined. Besides, MCS is used to generate sample datasets in normal distribution based on measured data.

TOPSIS method

The detail of the TOPSIS method implementation is given [1,4–6]:

**Step 1:** Construct object matrix and standard matrix based on the observed data and water quality level: there are \( m \) water quality indicators. Levels of water quality parameters are classified into five categories, with each level denoted as I, II, III, IV, and V. There are \( m \) water quality indicators in the standard matrix. The standard matrix is constructed in Eq. (4) [1]:

\[
\begin{align*}
&x_{ij}^S = \begin{cases} 
1 & \text{if level } i \\
0 & \text{otherwise}
\end{cases} \\
&x_{ij}^O = \begin{cases} 
1 & \text{if level } j \\
0 & \text{otherwise}
\end{cases}
\end{align*}
\]
\[
B = [b_{c,j}]_{5 \times m} = \begin{bmatrix}
C_1 & C_2 & C_3 & \cdots & C_m \\
I & b_{1,1} & b_{1,2} & b_{1,3} & \cdots & b_{1,m} \\
II & b_{2,1} & b_{2,2} & b_{2,3} & \cdots & b_{2,m} \\
III & b_{3,1} & b_{3,2} & b_{3,3} & \cdots & b_{3,m} \\
IV & b_{4,1} & b_{4,2} & b_{4,3} & \cdots & b_{4,m} \\
V & b_{5,1} & b_{5,2} & b_{5,3} & \cdots & b_{5,m}
\end{bmatrix}
\]

where \(c = 1, 2, \ldots, 5\) and \(j = 1, 2, \ldots, m\), \(b_{c,j}\) is the \(c_{th}\) standard concentration of the \(j_{th}\) water quality indicator. With regard to the object matrix construction, there are \(m\) water quality indicators, which is consistent with object matrix. Object matrix is constructed based on sample datasets and standard matrix as in the Eq. (3) of the companion paper [1].

**Step 2:** Normalise the data in object matrix: Data in object matrix are normalised since the water quality indicators have different scales and units. Moreover, it facilitates implementation of the TOPSIS method, in which the water quality indicators must be classified into benefit and cost indicators. For example, larger values of SD indicate better water quality. Thus, SD is a benefit indicator. By contrast, larger values of TN concentration indicate the worse water quality. Thus, TN is a cost indicator. Eq. (5) is used to normalise the data in the object matrix [6].

\[
x_{i,j(norm)} = \begin{cases} 
\frac{x_{i,j} - b_{1,j}}{b_{5,j} - b_{1,j}} & \text{for cost criteria} \\
\frac{x_{i,j} - b_{1,j}}{b_{1,j} - x_{i,j}} & \text{for benefit criteria} 
\end{cases}
\]

Note that \(i f x_{i,j(norm)} < 0, x_{i,j(norm)} = 0; i f x_{i,j(norm)} > 1, x_{i,j(norm)} = 1\).

**Step 3:** Assign the weight of each indicator: Determination of weights for water quality indicator considers the interaction among indicators. There are two parameters \((\alpha_j\) and \(\beta_j\) consisting of the weights which are presented in Eqs. (6) and (7). The correlation coefficient \((\beta_j)\) was proposed by Diakoulaki et al. [7].

\[
\alpha_j = \frac{1}{n} \sum_{i=1}^{n} \left( x_{i,j(norm)} - \frac{x_{1,j(norm)} + x_{2,j(norm)} + \cdots + x_{n,j(norm)}}{n} \right)^2, \quad j = 1, 2, \ldots, m
\]

\[
\beta_j = \sum_{k=1}^{m} (1 - \delta_{j,k})j = 1, 2, \ldots, m
\]

where:

\[
\delta_{j,k} = \frac{\sum_{i=1}^{n} (x_{i,j(norm)} - \bar{x}_{j(norm)}) (x_{i,k(norm)} - \bar{x}_{k(norm)})}{\sqrt{\sum_{i=1}^{n} (x_{i,j(norm)} - \bar{x}_{j(norm)})^2 \sum_{i=1}^{n} (x_{i,k(norm)} - \bar{x}_{k(norm)})^2}}, \quad j, k = 1, 2, \ldots, m
\]

\(\alpha_j\) is the standard deviations, \(\delta_{j,k}\) is a parameter considers the correlation between the \(j_{th}\) and \(k_{th}\) indicators. Two indicators have higher correlations when the absolute value of \(\delta_{j,k}\) is higher. Eq. (8) shows the weight of \(j_{th}\) water quality indicator:

\[
\omega_j = \frac{\alpha_j \cdot \beta_j}{\sum_{k=1}^{m} \alpha_j \cdot \beta_j
\]

**Step 4:** Identify the NIS and the PIS in the normalised comprehensive object matrix. Before determining the negative ideal solution (NIS) and positive ideal solution (PIS), the normalised comprehensive object matrix should be constructed by multiplying the normalised data and their corresponding weight, as shown in Eq. (9). Next, the PIS and NIS of the comprehensive object matrix
are determined through Eq. (10):

\[
R_{\omega} = \left[ \omega_j \cdot x_{i,j}(\text{norm}) \right]_{n \times m} = \begin{bmatrix}
C_1 & C_2 & \ldots & C_m \\
A_1 \left[ \omega_1 x_{1,1}(\text{norm}) & \omega_2 x_{1,2}(\text{norm}) & \ldots & \omega_m x_{1,m}(\text{norm}) \right] \\
A_2 \left[ \omega_1 x_{2,1}(\text{norm}) & \omega_2 x_{2,2}(\text{norm}) & \ldots & \omega_m x_{2,m}(\text{norm}) \right] \\
\vdots & \vdots & \ddots & \vdots \\
A_m \left[ \omega_1 x_{m,1}(\text{norm}) & \omega_2 x_{m,2}(\text{norm}) & \ldots & \omega_m x_{m,m}(\text{norm}) \right]
\end{bmatrix}
\]

(9)

\[
\begin{cases}
R^+_{i,j} = \max_{1 \leq j \leq n} (\omega_j x_{i,j}(\text{norm})) \\
R^-_{i,j} = \min_{1 \leq j \leq n} (\omega_j x_{i,j}(\text{norm}))
\end{cases}
\]

(10)

where the PIS set \((R^+_{i,j})\) contains all the maximum values of the water quality indicators, while NIS set \((R^-_{i,j})\) contains all the minimum values of the water quality indicators.

**Step 5:** Calculate the revised closeness coefficient of alternatives in normalised comprehensive object matrix: application of the TOPSIS method is to calculate the revised closeness of alternatives with different attributes. The revised closeness of each alternative is obtained by calculating the distance between the alternative and the NIS (PIS). Eqs. (11) are used to obtain the revised closeness coefficient of each alternative in normalised comprehensive object matrix.

\[
\eta_i = \frac{\sqrt{\sum_{j=1}^{m} (\omega_j x_{i,j}(\text{norm}) - R^-_{i,j})^2}}{\sqrt{\sum_{j=1}^{m} (\omega_j x_{i,j}(\text{norm}) - R^+_{i,j})^2 + \sum_{j=1}^{m} (\omega_j x_{i,j}(\text{norm}) - R^-_{i,j})^2}}
\]

(11)

**Membership function and water quality evaluation**

Membership function is a method for the knowledge inference and used to assess the fuzziness of alternative \([8,9]\). Eqs. (12)–(14) are membership functions.

\[
\theta_1(\eta_i) = \begin{cases} 
1, & \eta_i \in [0, \ell_{S_1}] \\
p_x(\eta_i - \ell_{S_1}) \exp\left(-\frac{\eta_i - \ell_{S_1}}{S_2 - S_1}\right), & \eta_i \in [\ell_{S_1}, 1]
\end{cases}
\]

(12)

\[
\theta_c(\eta_i) = \begin{cases} 
\exp\left(-\frac{\eta_i - \ell_{S_c}}{S_c - S_{c-1}}\right), & \eta_i \in [0, \ell_{S_c}] \\
p_x(\ell_{S_c} - \eta_i) \exp\left(-\frac{\ell_{S_c} - \eta_i}{S_{c+1} - S_c}\right), & \eta_i \in [\ell_{S_c}, 1]
\end{cases}
\]

(13)

\[
\theta_5(\eta_i) = \begin{cases} 
\exp\left(-\frac{\eta_i - \ell_{S_5}}{S_5 - S_4}\right), & \eta_i \in [0, \ell_{S_5}] \\
1, & \eta_i \in [\ell_{S_5}, 1]
\end{cases}
\]

(14)

where \(c = 2, 3, 4\). The coefficient \(P\) is an attenuation indicator.

**Lake eutrophication level evaluation**

The membership degree of revised closeness coefficient \((\eta_i)\) is calculated via the membership function. Eq. (4) in the companion paper [1], one type of defuzzification method, is applied to transform the fuzzy membership degree of alternative into a crisp value [10]. Index \(Q_i\) for the \(i\)th alternative is expressed in Eq. (5) in the companion paper [1]. Water quality level is determined based on the distribution the value of index \(Q_i\).
Table 1

Water quality level and its corresponding evaluated index (based on [1]).

| Level | Evaluated index of water quality level (%) |
|-------|---------------------------------------------|
| I     | $0 < Q_i < 20$                             |
| II    | $20 < Q_i < 40$                            |
| III   | $40 < Q_i < 60$                            |
| IV    | $60 < Q_i < 80$                            |
| V     | $80 < Q_i < 100$                           |

Fig. 1. Scatter plots of evaluated index distribution for lake eutrophication level.

**Global sensitive analysis (GSA)**

Sensitivity analysis is used to examine the performance of system in response to a minor change on the input factors [11,12]. Pearson correlation coefficient (PCC) is adopted to evaluate the importance of the input variables. PCC can be calculated using Eq. (6) in the companion paper [1]. The calculation pseudo code for the aforementioned procedure is give in the Appendix.

**Method validation**

Table 1 in companion paper [1] presented the water quality level and corresponding evaluated index. If the evaluated index $Q_i$ mainly distributed in the range 20–40%. The water quality level can be assigned level II. In this study, the standard matrix is constructed based on the division of influential factors concerning on water quality, presented as follows in Eq. (15) [1]:

$$B = [b_{c,j}]_{5 \times m} = \begin{bmatrix} C_{TN} & C_{TP} & C_{COD_{mn}} & C_{chl-a} & C_{SD} \\ I & 0.20 & 0.010 & 2 & 1 & 15 \\ II & 0.50 & 0.025 & 4 & 4 & 4 \\ III & 1.00 & 0.050 & 6 & 10 & 2.5 \\ IV & 1.50 & 0.100 & 10 & 30 & 1.5 \\ V & 2.00 & 0.200 & 15 & 65 & 0.5 \end{bmatrix}$$
Fig. 1 displays scatter plots of the evaluated result via developed approach. The water quality managers can determine the lake eutrophication level via the distribution of evaluated result ($Q_1^i$). In Fig. 1(a), the scatter plots are mainly distributed in the range 50–56% (Level III). It can be concluded that the water quality level is graded level III. In some circumstance, it is difficult to determine the lake eutrophication level. The evaluated result is consistent with real situation. Fig. 2 displays the statistic histogram of lake eutrophication level evaluation. It provides the convenience for water quality managers to determine lake eutrophication level. Fig. 3 presented the scatter plots of five factors via GSA in 2008. In Fig. 3, water quality indicators (COD$_{Mn}$, SD) are the higher sensitivity influential factors. In other words, these two factors displayed higher correlations with evaluated indexes ($Q_1^i$). Apart from that, different errors in measured data poses impacts on GSA result. In Fig. 3, GSA result is presented in four scenarios (errors within 5%, 10%, 15% and 20%). The correlations between each influential factor and evaluated index ($Q_1^i$) in different scenarios is performed similar without obvious fluctuations.

There are some limitations towards this study. There are lots of factors on water quality. In this study, there are only five influential factors considered. Besides, the distribution of measured data should be considered. In this study, the normal distribution is applied in this research. Finally, different countries have their own Surface water environmental quality standards (SWEQS). The water quality evaluation approaches is developed on the basic of SWEQS in China [13,14].
Conclusions

This developed approach can be used to lake eutrophication level (water quality) evaluation. The main contributions of this work are: (1) application of TOPSIS to consider surface water environmental quality standards and measured data; (2) use of the Monte Carlo simulation to overcome the errors caused by human and equipment factors in data collection. The case study validates the performance of developed approach. The conclusions are summarised:

1. The developed approach merges TOPSIS method, MCS and membership function. Monte Carlo simulation method is applied to generate a normal distributed dataset to overcome the errors caused by human and equipment in data collection. TOPSIS method is utilized to obtain the revised closeness.

2. The water quality level is determined based on the distribution of evaluated result. Global sensitivity analysis (GSA) results show that potassium permanganate index (COD$_{MN}$) displays the highest negative correlation with the evaluated results and Secchi disc (SD) performs the highest positive correlation under different errors in measured data.

Future work focuses on followed three aspects: (1) more influential factors are considered in the developed approach; (2) the different distribution of measured data should be considered, such as uniform, Poisson distribution; (3) Surface water environmental quality standards (SWEQS) in different countries should be considered.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relatio The nships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Song-Shun Lin: Investigation, Data curation, Writing - original draft. Shui-Long Shen: Conceptualization, Methodology, Supervision. Ning Zhang: Visualization, Validation, Writing - review & editing. Annan Zhou: Visualization, Writing - review & editing.

Acknowledgments

The research work was funded by “The Pearl River Talent Recruitment Program” in 2019 (Grant No. 2019CX01G338), Guangdong Province and the Research Funding of Shantou University for New Faculty Member (Grant No. NTF19024-2019).

Appendix

Algorithm 1: Pseudocode for the method for lake eutrophication levels evaluation: TOPSIS-MCS.

Required: Determine the influential factors $n$ and divisions of water quality indicators level $n$
Required: Input the measured data of influential factors.
Required: Input the values of water quality indicators level: $b_{c,j}$.
Required: Assign the values of $l_{ij}$ ($i = 1, 2, ..., 5$) and parameter $P$ to determine membership function.
Required: The number of sample data sets generated through MCS: $t$.

Construct the standard matrix $R = [b_{c,j}]_{n 	imes m}$.

Generate the sample data sets based on measured data through MCS.

$h = 1$

while $h \leq t$

Construct the object matrix $R^h = [x_{i,j}]_{(n + 1) \times m}$.

Normalize the data in object matrix: $x_{i,j}^{(norm)} \leftarrow x_{i,j}$

Obtain the weights of influential factors $\omega^h_j$. 


Construct comprehensive object matrix: \[ R_{h0}^i = [\omega_j \cdot X_{i,j}^{\text{norm}}]_{(n+1) \times m} \]

Identify PIS and NIS: \((R_{i,j}^+, R_{i,j}^-)_h\)

Calculate the revised closeness coefficient of \((n+1)\) alternatives: \(\eta_1^h, \eta_{b_1}^h, \eta_{b_2}^h, \ldots, \eta_{b_n}^h\).

Obtain the evaluated results of water quality: \(Q_1^h, Q_{b_1}^h, Q_{b_2}^h, \ldots, Q_{b_n}^h\).

\[ h \leftarrow h + 1 \]

Determine the water quality level based on the distribution of \(Q_1^h\).

References

[1] S.S. Lin, S.L. Shen, A.N. Zhou, Y.S. Xu, Approach based on TOPSIS and Monte Carlo simulation methods to evaluate lake eutrophication levels, Water Res. 187 (2020) 116437, doi: 10.1016/j.watres.2020.116437.

[2] C.L. Hwang, K. Yoon, Multiple Attributes Decision Making Methods and Applications, Springer, Berlin Heidelberg, 1981, doi: 10.1007/978-3-642-48318-9.

[3] M.P. Allen, Introduction to Monte Carlo Simulation. Observation, Prediction and Simulation of Phase Transitions in Complex Fluids, Springer, Netherland, 1995, doi: 10.1007/978-94-011-0065-6_8.

[4] N. Sadeghi, A.R. Fayek, W. Pedrycz, Fuzzy Monte Carlo simulation and risk assessment in construction, Comput. Aided Civ. Infrastructure Eng. 25 (4) (2010) 238–252, doi: 10.1111/j.1467-8667.2009.00632.x.

[5] G.R. Jahanshashloo, F.H. Lotfi, M. Izadikhah, An algorithmic method to extend TOPSIS for decision-making problems with interval data, Appl. Math. Comput. 175 (2) (2006) 1375–1384, doi: 10.1016/j.amc.2005.08.048.

[6] Z.Y. Li, T. Yang, C.S. Huang, C.Y. Xu, Q.X. Shao, P.F. Shi, X.Y. Wang, T. Cui, An improved approach for water quality evaluation: TOPSIS-based informative weighting and ranking (TIWR) approach, Ecol. Ind. 89 (2018) 356–364, doi: 10.1016/j.ecolind.2018.02.014.

[7] D. Diakoulaki, G. Mavrotas, L. Papayannakis, Determining objective weights in multiple criteria problems: the critic method, Comput. Op. Res. 22 (2) (1995) 763–770, doi: 10.1016/0305-0548(94)00059-H.

[8] T.L. Saaty, A scaling method for priorities in hierarchical structures, J. Math. Psychol. (1977).

[9] T.L. Saaty, Decision making with the analytic hierarchy process, Int. J. Serv. Sci. 1 (1) (2008).

[10] W.V. Leekwijk, E.E. Kerre, Defuzzification: criteria and classification, Fuzzy Sets Syst. 108 (2) (1999) 159–178, doi: 10.1016/S0165-0114(97)00337-0.

[11] F. Pinosi, T. Wagener, A simple and efficient method for global sensitivity analysis based on cumulative distribution functions, Environ. Model. Softw. 67 (2015) 1–11, doi: 10.1016/j.envsoft.2015.01.004.

[12] T. Wagener, F. Pinosi, What has Global Sensitivity Analysis ever done for us? A systematic review to support scientific advancement and to inform policy-making in earth system modelling, Earth Sci. Rev. 194 (2019) 1–18, doi: 10.1016/j.earscirev.2019.04.006.

[13] S.S. Lin, S.L. Shen, H.M. Lyu, A. Zhou, Assessment and management of lake eutrophication: a case study in Lake Erhai, China, Sci. Total Environ. 751 (2021) 141618, doi: 10.1016/j.scitotenv.2020.141618.

[14] H.M. Lyu, S.L. Shen, A. Zhou, The development of IFN-SPA: a new risk assessment method of urban water quality and its application in Shanghai, J. Clean. Prod. 282 (2021) 124542, doi: 10.1016/j.jclepro.2020.124542.