

Article

Not peer-reviewed version

General evaluation model for complex environment system

Mei Zi Yang , Yong Xing Wang , [Juan Wu](#) ^{*} , Jun Jie Song , Yu Gao , [Zhong Gen Wang](#) , Shi Hao Lan , Fu Bao Gan

Posted Date: 5 July 2023

doi: 10.20944/preprints202307.0299.v1

Keywords: Optimized improved entropy evaluation; Fuzzy comprehensive evaluation; Practical environmental system assessment; Linear programming.



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

General Evaluation Model for Complex Environment System

Mei-Zi Yang [†], Yong-Xing Wang [†], Juan Wu ^{*}, Jun-Jie Song, Yu Gao, Zhong-Gen Wang, Shi-Hao Lan and Fu-Bao Gan

College of Electrical and Information Engineering, Anhui University of Science and Technology, Huainan 232001, China

* Correspondence: yj165@mail.ustc.edu.cn

[†] These authors contributed equally to this work.

Abstract: Establishing an effective evaluation model to analyze the actual environmental level of the multi-indicator systems is an urgent challenge. However, due to the complexity of the environmental system, the factors that determine the practical system are usually interrelated. To solve this problem, many methods have been proposed and verified. However, the disadvantage of unreasonable weight determination and only single-indicator assessment limits the practical application of these analysis methods. Here, we established an optimized entropy weight model and integrate it with the fuzzy comprehensive evaluation (FCE) method to quantify the complex environmental system. By introducing the linear programming method into the entropy weight solution, we can obtain high accuracy weight values and strong resistance to extreme data. Theoretical and simulation results demonstrate that our method can enhance the precision of weight calculation and evaluation results for multi-factor systems. Our work develops an effective method to quantify the complex environment system comprehensively and is significant to real applications of evaluation method.

Keywords: optimized improved entropy evaluation; fuzzy comprehensive evaluation; practical environmental system assessment; linear programming

1. Introduction

As an important evaluation method, the comprehensive evaluation model has been widely applied in various fields, such as ecological civilization development level [1], environmental evaluation [2], ecosystem stability assessment [3] and water quality pollution evaluation [4]. This evaluation model can help researchers to quantify the environment system comprehensively. However, the practical environmental system is complex and its determining factors are often interrelated. Therefore, a single factor evaluation is insufficient to assess the status of practical system. However, most comprehensive evaluation models are usually used to assess single-factor systems, which fail to accurately capture the complexity of practical environmental systems [5]. Thus, the developing of effective evaluation model for multi-factors system is a crucial challenge.

In order to solve these issues and improve the performance of assessment, various methods have been proposed [6–11]. The traditional entropy weight (TEW) method has been proposed to obtain an objective weight based on the uncertainty of information. However, in situations where the differences in entropy are small, this method may yield excessively high weight values [12]. These results indicated that the TEW method has poor resistance to extreme values (data with a small difference in entropy). To effectively reduce the weight differences and enhance the resistance to extreme data, the improved entropy weight (IEW) method has been proven to optimize the entropy evaluation formula [13–18]. However, this method resulted in overcorrection of normal data weight, causing it to deviated from the law of consistency in entropy weight variation. Besides, this method is typically used in conjunction with single-factor evaluation (SFE) method, which is particularly suitable for systems with fewer or independent factors. But, for a complex systems with multiple factors, the SFE method may lack rationality due to the interrelationship between these factors. The fuzzy comprehensive evaluation

(FCE) method utilizes the membership degree theory of fuzzy mathematical model to transform qualitative assessments into quantitative ones, can effectively evaluate complex environmental systems that are influenced by multiple predetermined factors [19–25]. However, literature [26,27] have pointed that the FCE method is greater subjectivity in weight determination, which results in lower accuracy of evaluation results.

In this paper, we establish an optimized improved entropy weight (OIEW) model and integrate it with the FCE method to achieve multi-factors system assessment. To improve the accuracy of weight determination and enhance the anti-interference ability in handling extreme data, we develop an objective function based on the law of consistency in entropy weight variation using the mathematical programming concepts. This single objective model can be effectively solved by the linear programming algorithm, which prevents normal data from being suppressed. Furthermore, our model can effectively realize the comprehensive evaluation of multi-factors system by integrating the FCE method. The theoretical and simulation results demonstrate that our model can enhance the precision of weight determination and evaluation outcomes. Our works provide theoretical guidance for the practical application in the complex system evaluation field.

2. Disadvantages of entropy evaluation method

The traditional entropy evaluation (TEW) method is an objective weighting method that employs the calculation of information entropy value (IEV) to determine the weight of each factor. In the calculation process, the evaluation results follow the law of consistency in entropy weight variation, whereby an increase in IEV corresponds with decrease the weight, and vice versa [28] (The detail calculation process see Appendix A). However, this method has poor resistance to extreme data and is greatly affected by outliers. To overcome this limitation, several studies [15–18] have proposed by incorporating a correction factor to enhance its robustness against extreme data. Although these corrections can enhance their robustness to extreme data, the correction factor will overcorrect the weight of normal data. To further describe the drawbacks of the entropy evaluation method, we utilize Table 1 to delineate the issues in weight determination within both TEW and IEW methods.

Table 1. Weights of TEW method and IEW method. H_i represents the i -th information entropy value (IEV). W_{TEW} and W_{IEW1} are the weight calculated by the TEW method and IEW1 method, respectively.

Condition	(a) Extreme data						(b) Normal data					
H_i	0.9999	0.9998	0.9997	0.9996	0.9995	0.9994	0.989	0.975	0.946	0.928	0.905	0.875
W_{TEW}	0.0476	0.0952	0.1429	0.1905	0.2381	0.2857	0.0288	0.0654	0.1414	0.1885	0.2487	0.3272
W_{IEW1}	0.1666	0.1666	0.1667	0.1667	0.1667	0.1667	0.1579	0.1602	0.1651	0.1681	0.1719	0.1769

The yellow section in Table 1(a) shows the weights of extreme data (where minimal differences are observed among the $IEV(H_i)$). It can be observed that the weights calculated by the TEW method are 0.0476 when the H_i is 0.9999. However, a decreases of 0.005 in the IEV results in an approximate sixfold increase in the weights. The IEV exhibits minimal variation, whereas the TEW undergoes significant weight fluctuations. The weight formula [12] of the TEW method is displayed as follows

$$\omega_i = \frac{1 - H_i}{\sum_{i=1}^n (1 - H_i)}$$

(1)

where w_i refers the i -th weight calculated by TEW method. n is the number of evaluation factors. H_i represents the IEV of the i -th factor.

According to Equation 1, the weights of TEW method are determined by $1 - H_i$, and this value retains the significant decimal places when the difference in IEV is small. To provide further clarification on the issue of inadequate resistance to extreme data by the TEW method, the weight ratio Q [12] is introduced in ours work.

$$Q = \frac{X_i}{X_d}$$

(2)

Where X_i is the i -th weight of evaluation method. For extreme data, H_i and H_d exhibits nearly equal values ($H_i/H_d \approx 1$), thus using 1 as the baseline is a suitable metric to assess the degree of weight variation. As the difference in the decimal part decreases, the deviation between Q and 1 will increase. The greater the deviation from 1, the more inconsistent the variation in entropy weight. Table 2(a) shows the calculated Q value of TEW method for extreme data. We use 0.9996 as a fixed comparison. It can be observed that the weight ratio Q exhibits obviously deviation from 1 when the H_d is 0.9999, and the Q value is notably lower than 1 when H_d is 0.9994. The difference in IEV is negligible, whereas the deviation of Q value from 1 is significant. These results further substantiate the inadequacy of the TEW method in weight calculation for extreme data.

Table 2. Q value of different EW methods for extreme data. The correct factor of IEW1 is approximately 0.99965.

Method	(a) TEW method					(b) IEW1 method				
H_i	0.9996					0.9996				
H_d	0.9999	0.9998	0.9997	0.9995	0.9994	0.9999	0.9998	0.9997	0.9995	0.9994
Q	4.002	2.001	1.333	0.8001	0.6668	1.0003	1.0002	1.0001	0.9999	0.9998

To effectively solve this problem, the correct factor is used in TEW method [12,14,18]. Using the IEW1 method, as mentioned in the literature [18], we demonstrate the limitations of IEW method. The green section in Table 1(b) shows the influence of correction factor for normal data (where larger differences are observed among the IEV). It can be observed that the corresponding weights are 0.1579 and 0.1769 when H_i is at 0.9890 and 0.8750, respectively. There is a significant disparity in H_i , but the change in weight remains inconspicuous. The weight changes are smooth when the formula incorporating correction factor. To elucidate this phenomenon, we give the formula [18] of correction factor, as shown in follows

$$\omega_i' = \frac{1 - H_i + \bar{H}}{\sum_{i=1}^n (1 - H_i + \bar{H})} \quad (3)$$

where w_i' refers the i -th weight calculated by IEW1 method. The average value \bar{H} is the correct factor. It can be observed that the increase of the \bar{H} results in an increment of the numerator. For the extreme data, this increment in the numerator will increase the integer part of the weight value while diminishing the influence of the decimal portion. Ignoring the effect of the fractional part will mitigate the deviation of Q from 1. Therefore, this correction approach can reduce the differences between weights and improves the resistance to extreme data. The green section in Table 2(b) shows the calculated Q value of IEW1 method for extreme data. It can be seen that a small fluctuation in the IEV results in a correspondingly minor deviation of the Q value from 1. This phenomenon demonstrates that the addition of correction factors can improve the resistance to extreme data.

However, the integer part also weakens the weight difference of normal data and leads to a larger relative error between the changes in weight values and IEV, as shown in Table 3. We examine the impact of adding or omitting correction factors on Q values within the same dataset. The TEW method exhibits a clear decreasing trend in Q value as H_d decreases. However, the introduction of the correction factor \bar{H} only slightly decrease the Q value for IEW1 method. Therefore, the correction factor will impact the weight of normal data.

Table 3. Q value of different EW methods for normal data. The correct factor of IEW1 is 0.9305.

Method	(a) TEW method					(b) IEW1 method				
H_i	0.928					0.928				
H_d	0.989	0.975	0.946	0.905	0.875	0.989	0.975	0.946	0.905	0.875
Q	6.545	2.882	1.333	0.758	0.576	1.065	1.049	1.018	0.978	0.950

It should be emphasized that we use 1 as the benchmark to measure the degree of deviation between Q and 1 to assess the accuracy of the evaluation method for extreme data, but this benchmark is not applicable to normal data. This is because that there exists a significant disparity between the H_i and H_d for normal data ($H_i/H_d \gg 1$), thus rendering 1 as a benchmark is inadequate to assess the weight variation. To assess the entropy weight variation of normal data, the G value is introduced as an evaluative tool for assessing consistency. Equation 4 is formula of G [29].

$$G = \sqrt{\frac{1}{n-1} \sum_{i=1}^n \left[\frac{|X_i - X_{i+1}| - |H_i - H_{i+1}|}{|X_i - X_{i+1}|} \right]^2} \tag{4}$$

The G value is the standard deviation of the relative error between the differences in adjacent entropy values and those in weights. It denotes the consistency level of the entropy weight variation. A smaller value of G indicates a higher level of consistency, whereas a larger value of G signifies that the weights fail to meet the requirement for consistency. According to the calculation formula of standard deviation, G is a constant value. Table 4 shows the G value of different method for normal data and extreme data. It can be observed that the G value of IEW method is greatly higher than that of the TEW method. These results prove that the consistency of the TEW method is superior for normal data. But, the correction factor used in the IEW method overcompensates for normal data weights, leading to a violation of consistency principle.

Table 4. G value of normal data. IEW1, IEW2 and IEW3 are the improved entropy weight method obtained by the literature [12,14,18].

Method	TEW	IEW1	IEW2	IEW3
G	0.6909	5.5902	5.6087	3.0777

Therefore, building an appropriate evaluation model is a pressing issue, which should not only optimize the correction factors to normal data but also enhance the anti-interference ability to extreme data. Moreover, the evaluation process must adhere to the principle of consistency in entropy weight variation.

3. Disadvantage of SFE method

Two crucial aspects must be considered during assessing the complex environment systems [30]. Firstly, assigning weights to each factor individually becomes more difficult when the system comprises many factors. Secondly, when these factors exhibit categorical or hierarchical relationships, determining weights at different levels will add complexity to the evaluation process. If all factors are measured on the same level, it will weaken the internal correlation of each attribute and result in meaningless evaluation outcomes. Therefore, the SFE method is insufficient for evaluating multi-factor systems.

4. Theoretical model construction

To enhance the robustness to extreme data and mitigate excessive correction in IEW method, we propose an optimized improved entropy weight (OIEW) method to improve the weight determination process. By utilizing linear programming methodology and adhering to the principle of consistency in entropy weight variation, a programming model is constructed to calculate the weight. Furthermore, the FCE method is introduced to the OIEW method to overcome the simplistic of SFE method and provides a comprehensive evaluation for multi-factors systems.

4.1. Construction of evaluation datasets and reference sets, as well as the determination of membership functions

Appropriate sets of evaluation criteria, datasets, and membership functions can effectively reflect the true condition of the evaluated objects and enhance the accuracy of the evaluation results. The detailed construction process see Appendix B and Appendix C.

4.2. OIEW method

The TEW method employs a single formula for weight calculation, which results in insufficient robustness against extreme data. Although correction factors can be applied to each factor, they still rely on a static weight calculation formula. As entropy variation is a dynamic process, it is necessary to use a dynamic solution method that adjusts weight calculations accordingly. Therefore, we transform the mathematical planning model into linear programming model by simplex algorithm.

The weights calculated using the correction factor have a significant impact on normal data. Therefore, we introduce the concept of linear programming and optimize the computation process by imposing specific constraints on each weight. As G value can effectively characterize the consistency of entropy weight variation, we select it as the objective function in our linear programming model. According to the Equation 4, it can be observed that the smaller the value of $||H_i - H_{i+1}| - |X_i - X_{i+1}||$ results in a lower G value, thereby increasing the consistency. Therefore, we construct the objective function Z , which is defined as

$$Z = \min ||H_i - H_{i+1}| - |X_i - X_{i+1}|| \quad (5)$$

X_i and H_i represent the generalized weight value and IEV, respectively. When the constraints are satisfied, a smaller value of Z indicates more precise weights. The constraints of the objective function are specified as follows: (1) The total weight of all factors is 1; (2) The changes in weights are inversely proportional to the changes in IEV; (3) The weights X_i should range from 0 to 1.

To facilitate comparison with the TEW and IEW methods, we replace X_i with w''_i to denote the weights calculation in the OIEW method. Therefore, the formula of the OIEW model can be derived by synthesizing the aforementioned concepts.

$$Z = \min ||H_i - H_{i+1}| - |w''_i - w''_{i+1}||$$

$$s.t. \begin{cases} \frac{w''_i - w''_{i+1}}{H_i - H_{i+1}} < 0 \\ \sum_{i=1}^n w''_i = 1 \\ 0 \leq w''_i \leq 1 \end{cases} \quad (6)$$

By utilizing Equation 6 to calculate the weights, maximum consistency with the trend of information entropy changes can be ensured. To achieve stable data acquisition and simplify the calculation process, the single objective function problem is transformed into a linear programming model to optimize the weight calculation process (see Appendix D for the detailed description). The linear programming model that undergoes transformation is formulated as

$$\min \sum_{i=1}^n (u_i + v_i)$$

$$\begin{cases} \frac{\zeta_i - \epsilon_i}{H_i - H_{i+1}} < 0 \\ \sum_{i=1}^n w''_i = 1 \\ \zeta_i + \epsilon_i + u_i - v_i = |H_i - H_{i+1}| \\ u_i, v_i, \zeta_i, \epsilon_i \geq 0 \quad 0 \leq w''_i \leq 1 \end{cases} \quad (7)$$

Where u_i, v_i represent two variables that characterize $||H_i - H_{i+1}| - |w''_i - w''_{i+1}||$. Where ζ_i and ϵ_i represent two variables that characterize $|w''_i - w''_{i+1}|$.

This linear programming model can be solved by the simplex method. Therefore, to achieve the optimal solution, we must first convert the inequality in formula 7 into the standard form of constraints, resulting in the following equation.

$$\frac{\zeta_i - \epsilon_i}{H_i - H_{i+1}} + \Gamma_i = 0 \quad (8)$$

Where Γ_i is relaxation variables. By introducing this relaxation variables, the inequality constraints of the linear programming problem can be transformed into equality constraints. Then, the optimal solution can be obtained through the simplex method. And, the weights calculated by the dynamic solution process of the simplex method can greatly comply with the principle of consistency in entropy weight variation.

4.3. OIEFC method

The evaluation of multi-factors system by SFE method lacks comprehensiveness. Therefore, we incorporate the FCE method into the OIEW method to establish a multi-factors evaluation model (we call it the OIEFC evaluation model). In this method, the multi-factor set is divided into multiple sub-factor sets, and the first-level FCE is conducted for each subset. Subsequently, the obtained first-level evaluation vectors are further utilized as a new factor set to conduct second-level FCE. By repeating the aforementioned process, it is feasible to carry out comprehensive evaluations at third, fourth, and even higher levels in order to obtain the final evaluation results. This evaluation method fully considers the interrelationships among factors, and the hierarchical processing enhances the significance of weights to yield more accurate evaluation results. In order to further enhance the comprehension of the evaluation process, we present the two-level FCE models an illustrative example. The two-level model include first and second evaluation level.

(1) Establish the second-level FCE matrix.

After obtaining the weight judgment matrix and the weight vector, the FCE matrix can be constructed (The weight judgment matrix obtained is presented in Appendix C). The comprehensive result at the current level are determined by fuzzy algorithm. The final evaluation results can be obtained through step-by-step calculations. The detailed method is as follows:

The FCE matrix is calculated based on fuzzy operations. According to weights formula and simplex algorithm, the FCE matrix of the second-level B_p can be expressed as

$$\begin{aligned} B_p &= \vec{w}_p'' \cdot R_p \\ &= [\vec{w}_{p1}'', \vec{w}_{p2}'', \vec{w}_{p3}'', \dots, \vec{w}_{pn}''] \times \begin{pmatrix} r_{p11} & \dots & r_{p1j} \\ \dots & \dots & \dots \\ r_{pi1} & \dots & r_{pij} \end{pmatrix} \quad \text{Where } \begin{matrix} (p = 1, 2, \dots, t, \\ k = 1, 2, \dots, n) \end{matrix} \\ &= [b_{p1}, b_{p2}, b_{p3}, \dots, b_{pj}] \end{aligned} \quad (9)$$

Where p is p -th first level indicator, k is k -th second level indicator. \vec{w}_p'' , and R_p are the weight vector, and weight judgment matrix of the second-level factors corresponding to the p -th first-level indicator, respectively. \vec{w}_{pk}'' is the weight of the k -th secondary indicator corresponding to the p -th first-level indicator. $b_{p\alpha}$ (where $\alpha = 1, 2, \dots, j$) represents the overall membership of the α level corresponding to the p -th first-level indicator.

(2) Establish the first-level FCE matrix.

Then, the FCE matrix of second-level factors are taken as the weight judgment matrix of the first-level factors. And the membership of the corresponding first-level factor is determined by the

comprehensive membership degree of the secondary factors. Therefore, the weight judgment matrix of the first-level factors U can be written as

$$U = \begin{pmatrix} b_{11} & \dots & b_{1j} \\ \dots & \dots & \dots \\ b_{p1} & b_{p\alpha} & b_{pj} \\ \dots & \dots & \dots \\ b_{t1} & \dots & b_{tj} \end{pmatrix} = \begin{bmatrix} B_1 \\ \dots \\ B_p \\ \dots \\ B_t \end{bmatrix} \quad (10)$$

Where U is the weight judgment matrix of the first-level indicators. $b_{p\alpha}$ is the α -level membership of the p -th first-level indicator. By utilizing the weights of the second-level factors as raw data and applying the OIEW method, we can derive the weights of the first-level factors. Then, the FCE matrix of the first-level factors can be constructed. Ultimately, the evaluation of these indicators can be assessed by the principle of maximum membership degree. The evaluation model can be expressed as

$$\begin{aligned} L &= \vec{S} \cdot U \\ &= [S_1, \dots, S_p, \dots, S_t] \times \begin{bmatrix} B_1 \\ \dots \\ B_p \\ \dots \\ B_t \end{bmatrix} = [L_1, \dots, L_p, \dots, L_t] \end{aligned} \quad (11)$$

Where L is the first-level FCE matrix. \vec{S} is the weights matrix of first level indicator, S_p is the weights of p -th first level indicator. L_p is overall membership of p -th first level. If $L_\alpha = \max [L_1, \dots, L_p, \dots, L_t]$, the level results of the environmental system are level α .

5. Results and Analysis

5.1. The accuracy of OIEW Method

Figure 1 shows the weight comparisons obtained by different methods for extreme data (a) and normal data (b). As shown in Figure 1(a), when extreme data conditions cause a decrease in IEV at equal intervals of 0.0001, the TEW method exhibits a linear increasing trend and produces weight values that differ significantly. This phenomenon demonstrates that the TEW method violates the principle of consistency of entropy weight variation. But, the slight decrease in extreme entropy calculated by the OIEW method does not result in a significant change in weight, and the overall trend remains stable. This finding is consistent with that of the IEW1 method, indicating that the OIEW method is equally robust to extreme data. Figure 1(b) presents a comparison for the normal data. When the IEV randomly decreases from 0.989 to 0.875, the weight value calculated by TEW method exhibits a obvious upward trend, indicating that the TEW method adheres to the principle of consistency of entropy weight variation for normal data. But the weight values calculated by the IEW1 method show a relatively stable trend, with only minor fluctuations ranging from 0.1579 to 0.1769. This result demonstrates that the IEW1 method overcorrects normal data and violates the principle of consistency. In contrast, the weight of the OIEW method doubles from 0.1140 to 0.2280 as IEV increase, thereby satisfying the principle of consistency of entropy weight variation.

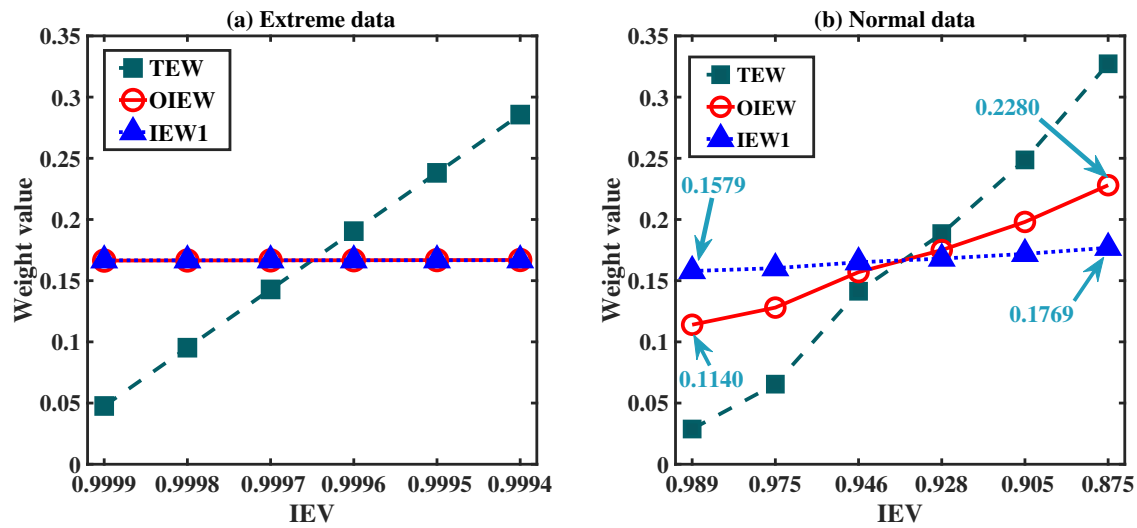


Figure 1. Weight comparison among different methods for extreme data (a) and normal data (b). IEV is the information entropy value. TEW, IEW1 and OIEW denote the traditional entropy weight method, the improved entropy weight method and the optimized improved entropy weight method, respectively. The blue numbers are the weight values corresponding to the information entropy.

Figure 2 compares the Q -value distribution of extreme data. We adopt $H_i = 0.9996$ as a fixed benchmark to compare the weight ratio of other entropy values. It can be seen that the Q value of TEW method exhibits a deviation from the baseline and deviates significantly from 1 when there is only a small difference between H_i and H_d . This phenomenon indicates a substantial deviation in weight obtained by the TEW method, demonstrating its poor resistance to extreme data. However, the Q value of IEW1 and OIEW methods consistently remains at the baseline. These results demonstrate that IEW and OIEW methods are robust against extreme data.

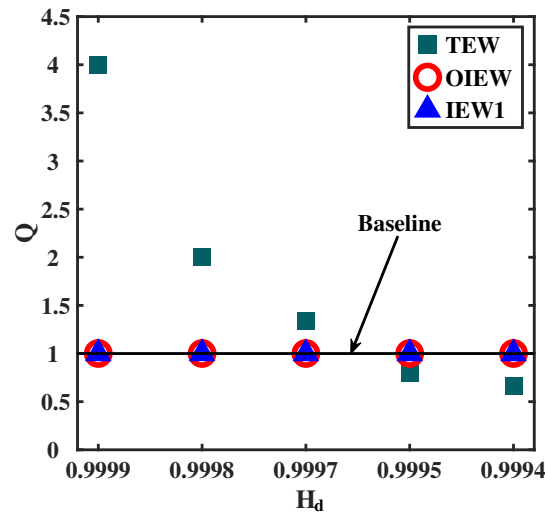


Figure 2. Q values of extreme data. H_d is d -th information entropy value. The black baseline line is 1.

To prove the consistency of entropy weight variation, we compare the G value of different method as shown in Figure 3. IEW1 [18], IEW2 [12] and IEW3 [14] method respond the entropy evaluation method with correction factor. It can be apparently seen that the G value of OIEW method is 1.9825×10^{-16} , which is greatly lower than the other four method, indicating its adherence to the principle of consistency in entropy weight variation. Furthermore, we can also observe that the G

values with the correction factor for normal data exceed that of TEW method, demonstrating an overcorrection by the correction factor and consequent violation of consistency principle.

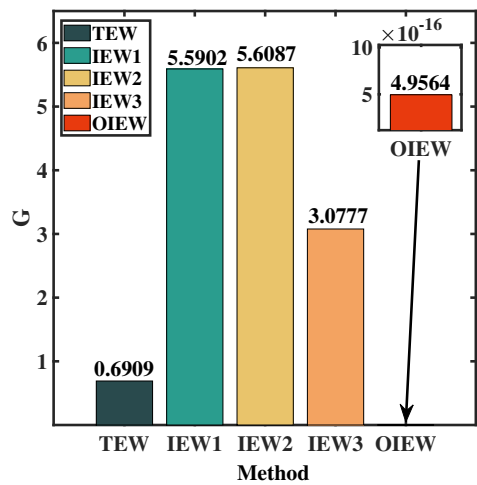


Figure 3. G values of normal data calculated by different EW methods.

Therefore, compared with the TEW and IEW1 methods, the weight changes of the OIEW method remain a relatively stable when IEV changes slightly, but show a corresponding trend when the IEV decreases randomly. The Q value of the OIEW method exhibits stability with the baseline for extreme data, whereas the G value is significantly small for normal data. These results obviously demonstrate that the OIEW method not only effectively alleviates the excessive correction of normal data and improves the resistance to extreme data, but also adheres the theory of consistency in entropy weight variation with high precision.

5.2. The rationality of the OIEFC model

To evaluate the validity of the OIEFC method, a comprehensive evaluation is performed on the soil dataset from Shandong Province of Chinese Soil Database (HWSD)(was collected from 2009/01/01 to 2009/12/31, and updated from 2020-12-31) [31,32]. This dataset provides actual measured physical indicators and chemical indicators collected from 154 different sampling points (The dataset is provided by National Cryosphere Desert Data Center. (<http://www.ncdc.ac.cn>)). Referring to the current national standards of ‘GB/T33469-2016: Quality Grade of Cultivated Land’ and ‘Agricultural Soil Nutrient Grading Standard’, two first-level factors are selected, which are further subdivided into six second-level factors. The first-level factors are the physical and chemical properties of the soil. The second-level factors are the soil pH(PH), organic carbon content (OC), soil electrical conductivity (ECE), soil cation exchange rate (CEC), soil bulk density (BULK), and soil texture (TEX). These factors constitute the assessment factor system for soil quality.

Table 5 displays the IEV and weight values obtained from different evaluation methods. As can be observed from the table, when the difference between the IEVs of the two physical factors is small, there is a significant discrepancy in weight values calculated by the TEW method for BULK and TEX factors, which are 0.101 and 0.899, respectively. In contrast, the OIEFC method produces nearly consistent weight values for both indicators, demonstrating its robustness to extreme data when evaluating the complex systems. For chemical factors, it can be seen that the OIEW method demonstrates the corresponding change in response to changes in entropy values as the IEV decrease from 0.9837 to 0.6733, while the IEW1 method exhibits a relatively stable trend. These results fully prove the high accuracy of the OIEFC method in determining weights for factors in practical systems.

Table 5. Evaluation values for the chemical and physical indicators. The chemical indicators include organic carbon (OC), potential of hydrogen (PH), cation exchange capacity (CEC), and electrical conductivity (ECE). The physical indicators include soil texture (TEX) and soil bulk density (BULK). The weight values for each indicator should be discussed separately based on the second-level indicator classification.

Indicator/Method		IEV	TEW	IEW1	OIEW
Chemical Indicator	OC	0.9837	0.0447	0.2313	0.1752
	PH	0.9979	0.0058	0.2278	0.161
	CEC	0.9807	0.053	0.2321	0.1782
	ECE	0.6733	0.8695	0.3089	0.4856
Physical Indicator	TEX	0.8932	0.899	0.5327	0.5474
	BULK	0.988	0.101	0.4763	0.4526

The evaluation results of soil quality and the proportion of soil utilization, which are obtained by the OIEFC method containing 154 soil data, are presented in Figure 4. The red area denotes the evaluation results of soil quality. From these bar chart, it can be observed that the soil quality level in Shandong Province is predominantly classified as Grade I (excellent), accounting for 66.22% of the total soil dataset. This is followed by Grade II (good) with a proportion of 18.92%. Grades III (moderate) and IV (poor) each account for 12.84%, while the smallest proportion belongs to Grade V (very poor) at only 0.68%. These analysis results indicate that the overall soil quality in Shandong Province meets an excellent standard. Compared to the 2020 soil distribution announced by Shandong Province (as shown in the blue area), the proportion of farmland (F) is 65.58%, while woodland (W), grassland (G), residential areas (R) and abandoned land (A) are 6.30%, 8.40%, and 18.60% and 1.12%, respectively. The predominant proportion of farmland further attests to the superior overall soil quality in Shandong Province, while the minimal proportion of abandoned land serves as evidence of poor soil quality. This findings fully validate the rationality and effectiveness of OIEFC method in practical system evaluation.

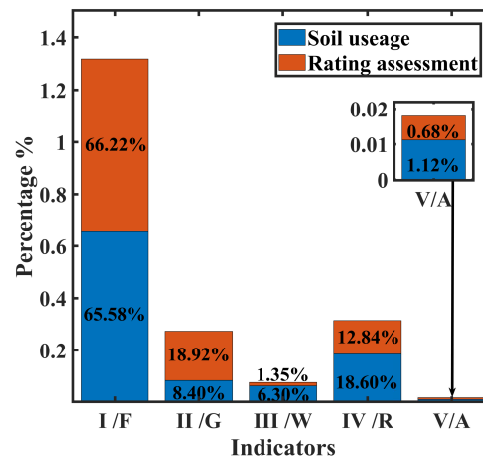


Figure 4. Soil evaluation results and soil utilization percentage. The orange bars are the evaluation grades. The blue bars represent land use types obtained from Shandong Province of Chinese Soil Database, including farmland (F), grassland (G), woodland (W), residential areas (R), and abandoned land (A). The soil evaluation results are divided into five levels: I (excellent), II (good), III (moderate), IV (poor), and V (extremely poor).

To further demonstrate the precision of the OIEFC method in multi-factor systems, we conducted a comparison with the SFE method adopted in [33]. Only the initial 30 samples from the Shandong soil dataset in the HWSO are selected to facilitate the comparison. Figure 5 shows the comparison results. Most of the line charts exhibit overlap, indicating that the evaluation results for both SFE

and OIEFC method on these samples are congruent. However, there exist notable variances in the evaluation results for ID15 and 24-27. To further elucidate the disparities among these data points, we conduct an analysis of the physical and chemical characteristics of ID15 and ID24 samples, as presented in Table 6. The yellow section in Table 6 are actual measurement data obtained by HWSO dataset. This dataset provides raw physical and chemical indicators data. The soil’s physical and chemical properties of ID26 and ID27 are same as ID24. From the Table 6, we can find that the physical and chemical properties of ID15 and ID24 are nearly identical, indicating that their soil grades are equivalent. The actual measurement data show that most of the indicator values for samples ID15 and ID24 are at level IV or V, indicating a relatively poor soil condition(The indicator level is evaluated according to the current national standards of ‘GB/T33469-2016: Quality Grade of Cultivated Land’ and ‘Agricultural Soil Nutrient Grading Standard’). However, the calculated result obtained by the SFE method is rated as level II (good) as shown in green section in Table 6, which did not align with the actual situation. In contrast, the OIEFC method’s evaluation result is rated as level IV (poor), demonstrating greater accuracy. Therefore, with almost identical indicator weights, the OIEFC method outperforms the SFE method in terms of comprehensiveness and precision in soil evaluation.

Table 6. Difference of evaluation results between ID15 and ID24. Actual measurement data obtained by HWSO dataset which provides raw physical and chemical indicators data for diverse soil indicators. Soil grade calculation results obtained by SFE and OIEFC method.

ID			ID15	ID24
Actual measurement data	Physical indicator	TEX	12	12
		BULK	1.58	1.6
	Chemical indicator	OC	0.3	0.63
		PH	7.1	6.5
		CEC	7	6
		ECE	0.1	0.1
Calculation evaluation results		SFE	II	II
		OIEFC	IV	IV

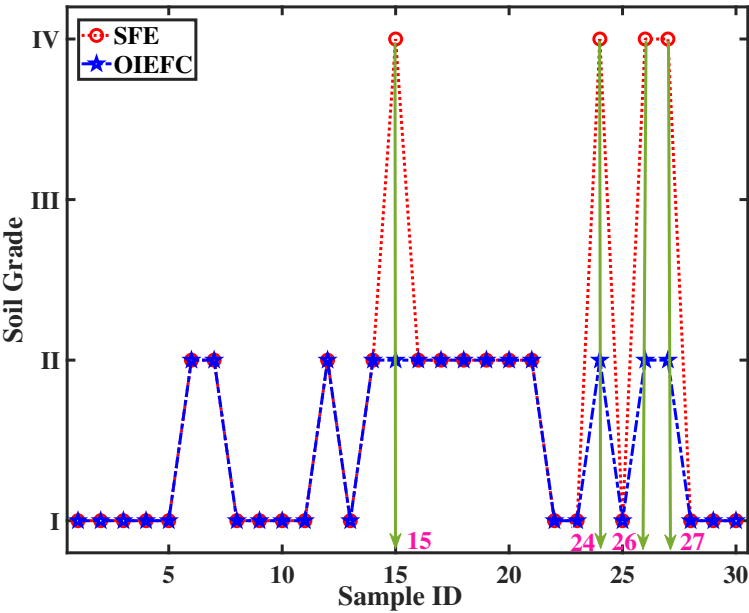


Figure 5. Comparison of evaluation results between SFE and OIEFC method. Sample ID is the serial number of the soil data. The blue dotted line and the red dashed line are the evaluation results calculated by the SFE method and OIEFC method, respectively. The magenta numbers indicate the serial numbers of the differential evaluation results.

6. Conclusion

In this paper, we conduct an optimized and improved entropy evaluation method to obtain more accurate weights. And, we also integrate our method with the FCE method to conduct a comprehensive evaluation for the multi-factor system. Theoretical and simulation results demonstrate that our method can enhance resistance to extreme data and improve the over-correction of normal data. Compared to the traditional method, our method is more adaptable to diverse information entropy conditions. Besides, it also provides a more comprehensive and accurate assessment results compared to the SFE method. Our works is instructive to quality evaluation for complex environment system, and provide guidance for the practical application in the complex system evaluation field.

Author Contributions: Conceptualization, M.Z.Y. and Y.X.W.; methodology, M.Z.Y.; software, Y.X.W.; validation, M.Z.Y., Y.X.W. and J.W.; formal analysis, M.Z.Y.; investigation, Y.G.; resources, J.J.S.; data curation, Z.G.W.; writing—original draft preparation, F.B.G.; writing—review and editing, M.Z.Y., Y.X.W. and J.W.; visualization, M.Z.Y.; supervision, J.W.; project administration, J.W.; funding acquisition, S.H.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Natural Science Foundation of China (62105004); University Natural Science Research Project of Anhui Province (KJ2020A0308); Natural Science Foundation of Anhui Province (2108085MF200); University-level general projects of Anhui University of science and technology (QNYB2021-05, QNYB2021-10).

Data Availability Statement: Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A: Calculation process of TEW method

The weight calculation using TEW method is outlined as follows:

(1) Construct the raw data matrix Y Suppose there are n evaluation indicators and μ evaluation factors. The raw data matrix Y can be expressed as

$$Y = (y_{ij})_{n \times \mu} \quad (A1)$$

Where y_{ij} is the j -th raw data of i -th evaluate indicator.

(2) Normalization of raw data To mitigate the impact of the original data dimension, we normalized the raw data. The normalization process is as follows

$$P_{ij} = y_{ij} / \sum_{j=1}^{\mu} y_{ij} \quad (A2)$$

Where P_{ij} is normalized factor values.

(3) Calculate the entropy value of each factor The information entropy value (IEV) of each indicators can be defined as follows

$$H_i = - \frac{\sum_{j=1}^{\mu} P_{ij} \log P_{ij}}{\ln \mu}, i = 1, 2, \dots, n \quad (A3)$$

Where H_i is information entropy of i -th indicator. Considering the possibility of zero values in some original data, errors may occur during entropy calculation. To solve this problem, a small increment is added to the indicator value. (e.g. 10^{-4}). This small value can prevent errors while preserving the integrity of the original data.

(4) Calculation of indicator weights

The weights assigned to each indicator can be determined upon completion of the IEV calculation.

$$w_i = \frac{1 - H_i}{\sum_{i=1}^n (1 - H_i)} \quad (A4)$$

Where w_i represents the weights of i -th indicator, and $0 \leq w_i \leq 1, \sum_{i=1}^n w_i = 1$

Appendix B: Evaluation Data Set and evaluation criteria set Construction

The evaluation data set is a raw data matrix. The evaluation criteria for each indicator are included in the evaluation set, which is consistently divided into five levels. However, different from the conventional FCE method, each indicator adheres to a strict ranking standard. Thus, this paper utilizes national standards and ranks each evaluation factor in the form of intervals

$$\begin{cases} F = \{f_{k1}, f_{k2}, \dots, f_{k5}\} \\ f_{ka} = [a_{k\alpha}, b_{k\alpha}] \quad (k = 1, 2, \dots, n; \alpha = 1, 2, \dots, j) \end{cases} \quad (\text{B.1})$$

Where F is evaluation set, $f_{k1} \sim f_{k5}$ represent five levels: I (excellent), II (good), III (moderate), IV (poor), and V (extremely poor). $a_{k\alpha}, b_{k\alpha}$ represent the α -th rank and the upper and lower limits of the interval corresponding to the k -th indicator, respectively.

Appendix C: Membership determination

The membership can be determined by the membership function, which is dependent upon the evaluation level value of the evaluation set. The evaluation set is expressed in the form of intervals. Therefore, the membership function is established based on the correlation between factor values and various levels. To facilitate the representation of the membership function, it is advisable to arrange all rank intervals in descending order when considering the irregular arrangement of each rank interval. Ultimately, reverting to the original order after calculating the membership. Membership function are expressed as follow (1) $a_{k\alpha} \leq \rho \leq b_{k\alpha}$

$$r_{pk\alpha}(\rho) = \begin{cases} \frac{\max(b_{k\alpha}) - \rho}{\max(b_{k\alpha}) - b_{k(\alpha-h)}} & a_{(\alpha-h)k} \leq \rho \leq b_{(\alpha-h)k} \\ 1 & a_{k\alpha} \leq \rho \leq b_{k\alpha} \\ \frac{\rho}{a_{k(\alpha+g)}} & a_{(\alpha+g)k} \leq \rho \leq b_{(\alpha+g)k} \end{cases} \quad (\text{C.1})$$

($h = 1, 2, \dots, \alpha - 1; g = 1, 2, \dots, j - \alpha$)

$$(2) \rho < \min(a_{k\alpha})$$

$$r_{pk\alpha}(\rho) = \frac{\rho}{a_{k\alpha}} \quad (\alpha = 1, 2, \dots, j) \quad (\text{C.2})$$

$$(3) \rho > \max(b_{k\alpha})$$

$$r_{pk\alpha} = \begin{cases} \frac{\max(b_{k\alpha}) - \rho}{\max(b_{k\alpha}) - b_{k(\alpha-h)}} & \alpha = 1, 2, \dots, j - 1 \\ 1, & \alpha = j \end{cases} \quad (\text{C.3})$$

Where ρ is variable. $r_{pk\alpha}$ is the membership of the p -th first indicator corresponding to the k -th secondary indicator at level α . $a_{k(\alpha-h)}$ and $b_{k(\alpha-h)}$ are the upper and lower limits of the h levels preceding level α for the k -th secondary indicator, respectively. $a_{k(\alpha+g)}$ and $b_{k(\alpha+g)}$ are the upper and lower limits of the g levels following level level α for the k -th secondary indicator, respectively.

Appendix D: Linearization of nonlinear programming problems

For certain nonlinear programming problems, it can be reformulated as linear programming problems to facilitate the solution process and ensure stable results. For instance, the following equation shows a nonlinear process

$$\begin{aligned} \min & |\delta_1| + |\delta_2| + \dots + |\delta_n| \\ \text{s.t.} & A\delta \leq b \end{aligned} \quad (\text{D.1})$$

where $\delta = [\delta_1, \delta_2, \dots, \delta_n]^T$ is the matrices of the corresponding dimensions. A and b are vectors.

To convert the aforementioned issue into a linear programming problem, it is simply necessary to take note of the following fact when u_i and v_i equal and bigger than 0.

$$\begin{aligned}\delta_i &= u_i - v_i \\ |\delta_i| &= u_i + v_i\end{aligned}\quad (\text{D.2})$$

Therefore, we can get $u_i = \frac{\delta_i + |\delta_i|}{2}$, $v_i = \frac{|\delta_i| - \delta_i}{2}$. By writing $u = [u_1, u_2, \dots, u_n]^T$, $v = [v_1, v_2, \dots, v_n]^T$, the aforementioned issue can be transformed into an equivalent form

$$\begin{aligned}\min & \sum_{i=1}^n (u_i + v_i) \\ \text{s.t.} & \begin{cases} A(u - v) \leq b \\ u, v \geq 0 \end{cases}\end{aligned}\quad (\text{D.3})$$

Here $u \geq 0$ indicates that each component of the vector u is greater than or equal to 0.

Combining the above linearization conversion methods, we can substitute the subsequent variables for Equation 7.

$$\begin{aligned}||H_i - H_{i+1}| - |\omega''_i - \omega''_{i+1}|| &= \mu_i + v_i \\ |H_i - H_{i+1}| - |\omega''_i - \omega''_{i+1}| &= \mu_i - v_i\end{aligned}\quad (\text{D.4})$$

where H_i, H_{i+1} are the entropy values of the i -th known metric and its neighbors. w''_i and w''_{i+1} are the corresponding weight.

Introducing two new variables ζ_i, ϵ_i converts $|\omega''_i - \omega''_{i+1}|$ and $w''_i - w''_{i+1}$ equivalently, we can obtain

$$\begin{aligned}|\omega''_i - \omega''_{i+1}| &= \zeta_i + \epsilon_i \\ \omega''_i - \omega''_{i+1} &= \zeta_i - \epsilon_i\end{aligned}\quad (\text{D.5})$$

Therefore, by utilizing the Equation D.4 and Equation D.5, we can linearize the weight difference between adjacent indicators to derive the relationship of new various, that is $\zeta_i + \epsilon_i + u_i - v_i = |H_i - H_{i+1}|$. However, since the constraints are imposed on each weight rather than the difference of the weights, and introducing new variables to represent separate weights would increase complexity, w''_i remains a variable in our model.

Bring the aforementioned equation into Equation 7, the linear programming model after variable substitution can be obtained

$$\begin{aligned}\min & \sum_{i=1}^n (u_i + v_i) \\ & \begin{cases} \frac{\zeta_i - \epsilon_i}{H_i - H_{i+1}} < 0 \\ \sum_{i=1}^n w''_i = 1 \\ \zeta_i + \epsilon_i + u_i - v_i = |H_i - H_{i+1}| \\ u_i, v_i, \zeta_i, \epsilon_i \geq 0 \quad 0 \leq w''_i \leq 1 \end{cases}\end{aligned}\quad (\text{D.6})$$

References

1. Hu, P. Evaluation algorithm of coastal city ecological civilization development level based on improved BP neural network. *Journal of Environmental Management* **2022**, 321, 116039.
2. Lee, K.I.; Lin, S.C. A fuzzy quantified SWOT procedure for environmental evaluation of an international distribution center. *Information Sciences* **2008**, 178, 531–549.
3. Zhang, R.; Zhang, X.; Yang, J.; Yuan, H. Wetland ecosystem stability evaluation by using Analytical Hierarchy Process (AHP) approach in Yinchuan Plain, China. *Mathematical and Computer Modelling* **2013**, 57, 366–374.
4. Li, R.; Zou, Z.; An, Y. Water quality assessment in Qu River based on fuzzy water pollution index method. *Journal of environmental sciences* **2016**, 50, 87–92.
5. Chen, J.F.; Hsieh, H.N.; Do, Q.H. Evaluating teaching performance based on fuzzy AHP and comprehensive evaluation approach. *Applied Soft Computing* **2015**, 28, 100–108.

6. Yue-Ju, X.; Shu-Guang, L.; Yue-Ming, H.; Jing-Feng, Y. Soil quality assessment using weighted fuzzy association rules. *Pedosphere* **2010**, *20*, 334–341.
7. Xie, L.; Zhong, J.; Chen, F.; Cao, F.; Li, J.; Wu, L. Evaluation of soil fertility in the succession of karst rocky desertification using principal component analysis. *Solid Earth* **2015**, *6*, 515–524.
8. Zhang, X.; Wang, C.; Li, E.; Xu, C. Assessment model of ecoenvironmental vulnerability based on improved entropy weight method. *The Scientific World Journal* **2014**, *2014*.
9. Liu, J.; Liu, C.; Han, W. Efficiently evaluating heavy metal urban soil pollution using an improved entropy-method-based TOPSIS model. *Archives of environmental contamination and toxicology* **2016**, *71*, 377–382.
10. Kucuker, D.M.; Giraldo, D.C. Assessment of soil erosion risk using an integrated approach of GIS and Analytic Hierarchy Process (AHP) in Erzurum, Türkiye. *Ecological Informatics* **2022**, *71*, 101788.
11. Heil, J.; Häring, V.; Marschner, B.; Stumpe, B. Advantages of fuzzy k-means over k-means clustering in the classification of diffuse reflectance soil spectra: A case study with West African soils. *Geoderma* **2019**, *337*, 11–21.
12. Deng, X.; Chen, C. A novel portfolio selection with prospect value constraint and distance measure of IFSs based on the improved entropy-weighted method. *Journal of Intelligent & Fuzzy Systems* **2020**, *39*, 3519–3543.
13. He, Y.; Jiao, Z.; Yang, J. Comprehensive evaluation of global clean energy development index based on the improved entropy method. *Ecological Indicators* **2018**, *88*, 305–321.
14. Wang, S.; Xu, Y.; Wang, D.; Gao, B.; Lu, M.; Wang, Q. Effects of industry structures on water quality in different urbanized regions using an improved entropy-weighted matter-element methodology. *Environmental Science and Pollution Research* **2020**, *27*, 7549–7558.
15. Xu, H.; Ma, C.; Lian, J.; Xu, K.; Chaima, E. Urban flooding risk assessment based on an integrated k-means cluster algorithm and improved entropy weight method in the region of Haikou, China. *Journal of hydrology* **2018**, *563*, 975–986.
16. Deng, X.; Xu, Y.; Han, L.; Yu, Z.; Yang, M.; Pan, G. Assessment of river health based on an improved entropy-based fuzzy matter-element model in the Taihu Plain, China. *Ecological Indicators* **2015**, *57*, 85–95.
17. Zuo, L.; Dong, S.; Zhu, C.; Shu, L.; Han, G. A cloud resource evaluation model based on entropy optimization and ant colony clustering. *The Computer Journal* **2015**, *58*, 1254–1266.
18. Ouyang, S.; Liu, Z.W.; Li, Q.; Shi, Y.L. A new improved entropy method and its application in power quality evaluation. *Advanced Materials Research. Trans Tech Publ*, 2013, Vol. 706, pp. 1726–1733.
19. Torbert, H.; Krueger, E.; Kurtener, D. Soil quality assessment using fuzzy modeling. *International Agrophysics* **2008**, *22*, 365–370.
20. Liu, L.; Zhou, J.; An, X.; Zhang, Y.; Yang, L. Using fuzzy theory and information entropy for water quality assessment in Three Gorges region, China. *Expert Systems with Applications* **2010**, *37*, 2517–2521.
21. Qi, Y.; Wen, F.; Wang, K.; Li, L.; Singh, S. A fuzzy comprehensive evaluation and entropy weight decision-making based method for power network structure assessment. *International Journal of Engineering, Science and Technology* **2010**, *2*, 92–99.
22. Li, Z.w.; Huang, J.q.; Li, Y.y.; Guo, W.; Zhu, J.f. Assessment on soil fertility of Dongting Lake wetland area (China) based on GIS and fuzzy evaluation. *Journal of Central South University* **2011**, *18*, 1465–1472.
23. Khalili-Damghani, K.; Tavana, M.; Santos-Arteaga, F.J. A comprehensive fuzzy DEA model for emerging market assessment and selection decisions. *Applied Soft Computing* **2016**, *38*, 676–702.
24. Bo, L.; Cheng, S.; Li, D. Establishment and application of fuzzy comprehensive evaluation of green building design based on data mining. *Journal of Intelligent & Fuzzy Systems* **2020**, *38*, 6815–6823.
25. Xie, Q.; Ni, J.Q.; Su, Z. Fuzzy comprehensive evaluation of multiple environmental factors for swine building assessment and control. *Journal of Hazardous Materials* **2017**, *340*, 463–471.
26. Zhang, X.; Wang, Y.; Xiang, Q.; Zhang, H.; Jiang, Z. Remanufacturability evaluation method and application for used engineering machinery parts based on fuzzy-EAHP. *Journal of manufacturing systems* **2020**, *57*, 133–147.
27. Alaoui, Y.L.; Tkiouat, M. Assessing the performance of microfinance lending process using AHP-fuzzy comprehensive evaluation method: Moroccan case study. *International Journal of Engineering Business Management* **2017**, *9*, 1847979017736692.
28. Ouyang, L.Y.; Chen, K.S.; Yang, C.M.; Hsu, C.H. Using a QCAC–Entropy–TOPSIS approach to measure quality characteristics and rank improvement priorities for all substandard quality characteristics. *International Journal of Production Research* **2014**, *52*, 3110–3124.

29. Hodson, T.O. Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not. *Geoscientific Model Development* **2022**, *15*, 5481–5487.
30. Liu, J.P.; Ji, Z.G. Evaluation of the single-factor approach to determining wave exposure. *Journal of Coastal Research* **2004**, *20*, 546–552.
31. lin, L.; chao, L. Chinese soil data set based on world soil database (hwsd) (v1.1).
32. Fischer, G.; Nachtergaele, F.; Prieler, S.; van Velthuisen, H.T.; Verelst, L.; Wiberg, D. Global Agro-ecological Zones Assessment for Agriculture (GAEZ 2008). *IIASA, Laxenburg, Austria and FAO, Rome, Italy* **2008**.
33. Wei-Xin, L.; ZHANG, X.X.; Bing, W.; Shi-Lei, S.; Yan-Song, C.; Wen-Yang, P.; Da-Yong, Z.; CHENG, S.P. A comparative analysis of environmental quality assessment methods for heavy metal-contaminated soils. *Pedosphere* **2008**, *18*, 344–352.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.