RESEARCH ARTICLE

Ecological economic analysis based on force-oriented edge binding of segmented skeletons improved clustering data

Ning Yan[1, *], Aili Ma[2]

[1]Hebei Youth Administrative Cadres College, Shijiazhuang, Hebei, China. [2]Huaxin College, Hebei GEO University, Shijiazhuang, Hebei, China.

Received: October 18, 2023; **accepted:** November 28, 2023.

Visualization techniques can aid in the comprehension and appreciation of the inherent connections between ecological and economic data. However, traditional visualization techniques still suffer from the problems of information overload and missing information. To address this problem, the study proposed a technique to improve the clustered data based on force-directed edge binding of segmented skeletons and constructed a visual analysis model for ecological and economic data based on the improved technique. The study used the parallel coordinates technique, the technique of presenting complex data relationships in parallel coordinates, as the control experimental group in the technical performance comparison test. The results showed that the data relevance indexes of the improved technique and all the control experimental groups were 0.145 and 0.138, respectively. The maximum positive value of the accuracy error of the number of clusters in the improved model was 0.012, while the minimum negative value was -0.003. The control experimental groups were 0.022, -0.016 and 0.022, -0.018, respectively. The study proposed a superior visual analysis model for ecological and economic research. The model utilized a force-directed edge-bound clustering technique based on segmented skeleton data, providing accurate and comprehensive insights. The model could provide more accurate and comprehensive data analysis and decision-making basis to promote the development and sustainability of ecological economy.

Keywords: force-oriented edge binding; segmented skeletons; cluster binding; ecological and economic data; high dimensional data; visualization.

[*]**Corresponding author:** Ning Yan, Hebei Youth Administrative Cadres College, Shijiazhuang 050021, Hebei, China. Email: 15081116164@163.com.

Introduction

In recent years, with economic globalization and the rapid development of information technology, ecological and economic analysis has gradually become a hot field of research [1, 2]. The ecological economic analysis mainly focuses on the impact of economic activities on the environment and the interaction between ecosystems and economic systems [3, 4]. Ecological and economic analysis is a thorough research method with the goal of assessing and analyzing the interaction between economic activities and ecosystems. Regarding the application of ecological economic analysis, numerous scholars have carried out a great deal of research. Wu constructed a spatial econometric model by combining ecological constraint intensity, green innovation efficiency, and a series of high-quality economic development indicators, and empirically analyzed the impact of environmental constraint intensity on green innovation efficiency using Chinese provincial panel data [5]. Cuimei *et al*. proposed a quantitative method based on the energy value theory to evaluate the ecological and economic benefits of using recycled water for irrigation. The method could calculate the cost of different irrigation water sources and the additional benefits associated with using saved tap water [6]. Song *et al*. used the city panel data to establish a comprehensive model to evaluate the competitiveness of the cities in the Huaihe River ecological and economic belt. They applied a dynamic factor approach to analyze and discuss the results. The study's findings revealed that the competitiveness of cities in the Huaihe River ecological economic belt lacked balance in terms of static scores, while the distribution of comprehensive scores among these cities varied significantly. The results also showed that the eastern coastal region was stronger than the central and western regions [7]. Yang *et al*. constructed a comprehensive model using the carbon pressure index, the carbon occupancy index, and the ecological and economic coordination index of Caijiaopo Town for 2012-2017. They used the decoupling index to describe the relationship between capital adequacy, economic growth, and per capita disposable income. The results showed that taking carbon emissions as a constraint could effectively improve the energy utilization rate and optimize the consumption structure of residents [8].

Ecological economy is a subject area combining ecology and economics, which seeks to balance economic growth with natural resource conservation to promote sustainable development by seeking a sustainable development path through the exploration of the interaction between ecosystems and economic systems. The intricate and diverse ecological economic data can pose challenges in interpreting and comprehending. By integrating data on ecological resources, economic development, and energy consumption, as well as combining methods such as data mining and visual analysis, it is possible to study in depth the intrinsic links between macroeconomic policies, residents' daily lifestyles, and the consumption and protection of ecological resources [9, 10], which will provide decision-making assistance for the formulation of economic policies for sustainable development. In the field of ecological economy, data visualization is an important tool to help analysts better understand and interpret patterns and relationships in data. The ecological data can be analyzed in depth using the visual analysis model of ecological and economic data to reveal the hidden knowledge such as the data patterns and intrinsic connections behind [11]. However, there are still some problems with the application of traditional visualization in analysis such as information overload and information loss [12, 13]. Traditional methods of data analysis are often limited to statistical indicators and tables that lack intuitiveness and interactivity [14]. In order to better understand and communicate ecological economic data, researchers have gradually adopted visualization techniques for data presentation and analysis. In the work of ecological economy visualization, it is first necessary to collect the raw data of the ecological economy to form the ecological economy dataset. Then, preprocessing is conducted to generate a dataset suitable for visualization purposes. Subsequently, the data is visually encoded in various dimensions based on the requirements of color, size, and other attributes to complete the conversion of data to view. Finally, the converted view will be transmitted to the user to enable them to comprehend and evaluate the ecological and economic data at hand. However, the traditional visualization technology for complex, high-dimensional data conversion is not ideal. The Cluster Analysis is a technique for statistical data analysis, which is a method of studying individuals based on the characteristics of the things themselves with the purpose of categorizing similar things [15]. It is based on the principle that individuals in the same category have greater similarity and individuals in different classes are very different, and it is widely used in many fields.

This study proposed a method to improve the clustered data based on force-directed edge binding of segmented skeletons. Additionally, a visual analysis model for ecological and economic data was constructed using this improved method. It was expected that the introduction of force-directed edge binding of segmented skeleton could better capture the correlation between data and provide more accurate clustering results. This research could provide a new perspective and methodology for processing and analyzing complex eco-economic data through the innovative utilization of a force-directed edge-binding technique for segmented skeletons in the field of ecological and economic analysis and was valuable in improving the accuracy of data visualization and reducing information loss. As a result, policy makers and researchers could better comprehend and evaluate the impact of economic activities on ecosystems. The application of this technique would promote the development of ecological and economic research methodologies and provide more accurate data to support the formulation of sustainable development policies. In addition, the results of this study would have a profound impact on promoting the integration between interdisciplinary research such as environmental science, economics, and data science. Ultimately, by optimizing the analysis of eco-economic data, this study was expected to make a positive and instructive contribution to global ecological conservation and sustainable development policymaking. Additionally, it emphasized the need and urgency of eco-economic research in the context of current global environmental changes and economic challenges.

Materials and methods

Cluster-binding-based visual analysis of ecological and economic data

This study used k-means clustering algorithm in cluster analysis for visual analysis model design of ecological and economic data [16]. The k-means algorithm belongs to one of the basic and the most widely used classification methods in cluster analysis methods, which is a method to discover clusters and cluster centers in classless labeled data during the classification process. The k-means algorithm constructed a similarity matrix based on the data similarity in the classification process, which was expressed as:

$$A=\begin{pmatrix} x_{11} & x_{12} & \dots & x_{1j} \\ \dots & \dots & \dots & \dots \\ x_{i1} & x_{i2} & \dots & x_{ij} \end{pmatrix} \tag{1}$$

where *i* was the data objects. *j* was the features of each data object. The degree of similarity between different objects could be reflected by the distance between the objects. The smaller the distance indicated that the objects were more similar. On the other hand, it indicated that the objects were more different. There were three commonly used distance calculation formula in clustering with the mathematical expressions as shown in Equations (2), (3), and (4) respectively.

$$D_1(a,b)=\sqrt{\sum_{i=1}^{n}\left(a_{ij}-b_{ij}\right)^2} \tag{2}$$

where D_1 was the Euclidean distance between the objects *a* and *b*. a_{ij} was the coordinates of the object *a*, b_{ij} was the coordinates of the object *b*. *n* was the dimension of the object point.

$$D_2(a,b)=\sum_{i=1}^{n}\left|a_{ij}-b_{ij}\right| \tag{3}$$

where D_2 was the city block distance between objects *a* and *b*.

$$D_3(a,b)=\max_{1\le i\le n}\left|a_{ij}-b_{ij}\right| \tag{4}$$

where D_3 was the Chebyshev distance between objects *a* and *b*. In the clustering algorithm, not only the distance function could be used to represent the similarity between the objects, but also the similarity coefficient could be used to determine the degree of similarity between the objects, which was calculated below.

$$D^1(a,b)=\frac{\sum_{i=1,j=1}^{n} a_{ij}b_{ij}}{\sqrt{\sum_{i=1,j=1}^{n} a_{ij}^2 \sum_{i=1,j=1}^{n} b_{ij}^2}} \tag{5}$$

where $D^1(a,b)$ was the similarity between objects a and b. In the use of clustering, the intersection of the mean line and the virtual main axis was the point of action of the clustering gravity as shown in Figure 1, where X_i was the coordinate axes, P_i was the data line composed of sampled points in the clusters, V_i was the virtual principal axis, C_i was the intersection of the clustered mean line with the virtual principal axis, Q_i was the intersection of the sampled data line with the principal axis, which served as points that would move to C_i under the effect of cluster gravity to the position of Q_i. The moving distance was calculated as:

$$Q_i^{'} = Q_i + \beta(C_i - Q_i) \tag{6}$$

where β was the adjustment factor of the gravitational field in the clustering with the value range of [0, 1].

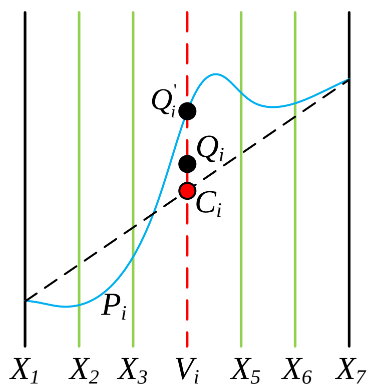

Figure 1. Cluster-center gravitational analysis.

For the high-dimensional data in the visualization, it was still possible to use the k-means algorithm to rationalize the location of the clustering center point between each level, which could reflect the hierarchical relationship in high-dimensional data (Figure 2). In the process of high-dimensional data clustering, C_1 and C_2 were the two clusters clustered in the upper layer, while the lower layer was decomposed into two and three sub-clusters as C_{11}, C_{12}, C_{21}, C_{22}, and C_{23}, respectively. The center value positions of the two sub-clusters C_{23} and C_{11} were interleaved with each other, which destroyed the overall distribution relationship of the upper clusters (Figure 2a). Thus, the process of view transformation in data visualization would produce the problem of data lines blocking and interlacing each other and increased the center value position of lower clusters on the virtual spindle through clustering to rationalize the hierarchical relationship of clustering. The sub-clusters were adjusted the position on the virtual spindle and could be calculated as:

$$C_{i2} = (1-\beta_1)C_{2i} + \beta Q_i + \beta_1 C_i \tag{7}$$

where β_1 was the adjustment factor of the lower level clustering center of C_1. C_{i2} was the position of the sub-cluster C_{2i} after moving. The new binding locus position of the lower level cluster center value after moving was calculated below.

$$Q^{'} = (1-\beta_2)Q_i + \beta_2 C_{2i} \tag{8}$$

where Q was the location of the new binding positioning point of the lower-level cluster center value. β_2 was the lower-level cluster center adjustment factor of C_2. After the lower level clustering center value was adjusted, the distribution of each clustering mean line on the virtual spindle corresponded to the hierarchical topology relationship, which reduced the mutual occlusion of data lines and improved the view accuracy.

Design of visual analysis model for ecological and economic data

The energy-ecology-economy model is an analytical method that integrates energy resource utilization, ecological environmental protection, and economic development [17, 18]. Through this model, the ecological-economic model of sustainable development can be analyzed in depth (Figure 3). Among them, the energy module mainly selected data dimensions from fossil energy and electricity. The ecological module mainly selected data dimensions from the aspects of ecological footprint and biological carrying capacity. The economic module mainly

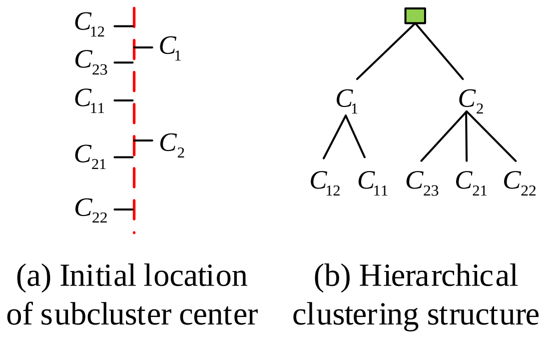

Figure 2. Location adjustment of hierarchical clustering.

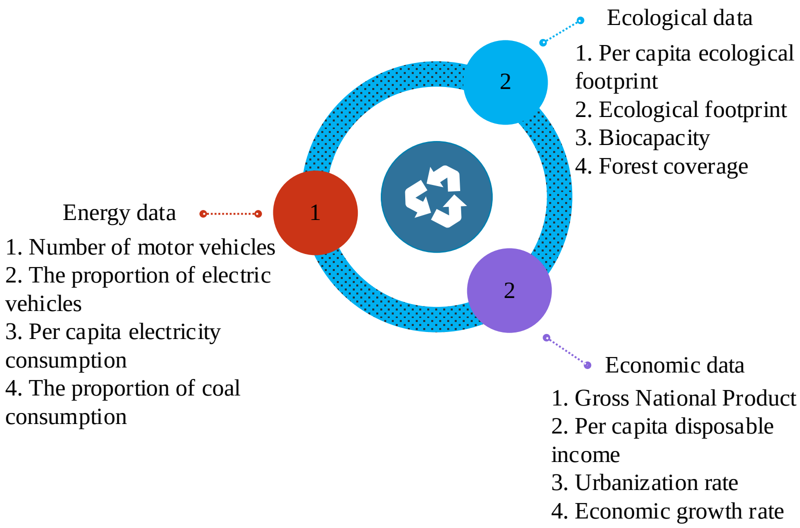

Figure 3. Structural diagram of the energy-ecological-economic model.

selected data dimensions from the aspects of economic development and residents' life. For each statistical index of those multi-attributes, a database was formed for visualization and analysis. The force-oriented edge binding of segmented skeleton (FDSBB) is a technique based on a force-directed algorithm that creates associations between bones and vertices of the model mesh by creating edges on the model mesh, thus enabling the influence of bones on the mesh to improve the visual analysis model based on cluster analysis. In constructing the model mesh, the skeleton reached the equilibrium position under the action of gravitational and repulsive forces. The minimum distance between each clustered skeleton was defined as the equilibrium distance, so as to obtain the gravitational and repulsive forces of the model, which was calculated in Equations (9) and (10).

$$\begin{cases} K = t \times \dfrac{H}{N} \\ F_1(P,Q) = K^2 / D_{P,Q} \end{cases} \tag{9}$$

where K was the equilibrium distance. H was the height of the coordinate axis. N was the number of clusters. t was the adjustment factor. $F_1(P,Q)$ was the repulsion between the skeletons of the clusters P and Q. $D_{P,Q}$ was the distance between the centers of the skeletons of the clusters P and Q.

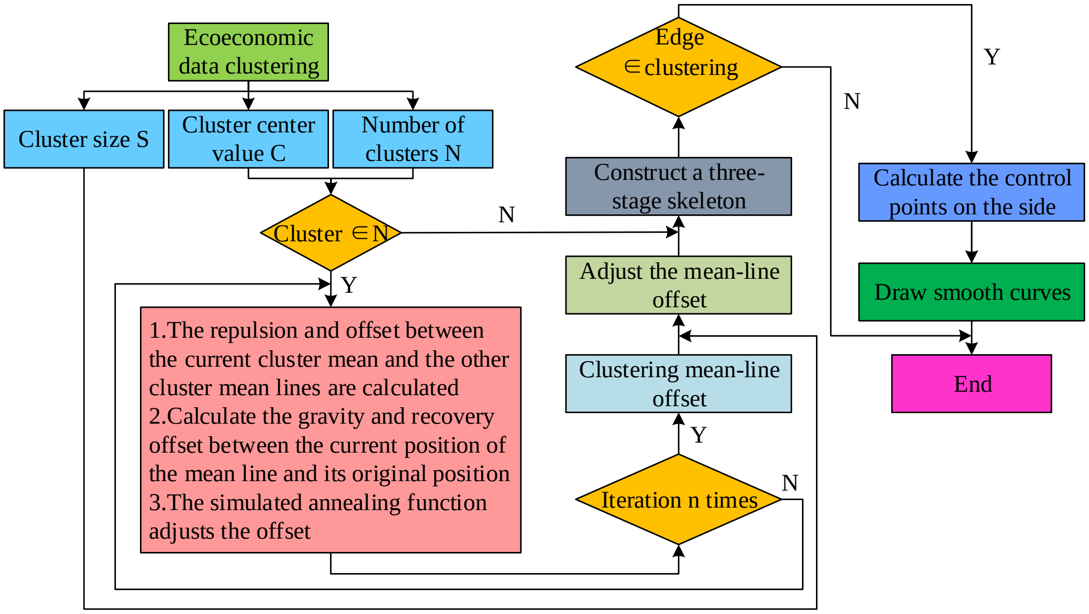

Figure 4. Visual analysis model of ecological and economic data based on the improved algorithm.

$$\begin{cases} F_2(P) = \Delta P^2 / K \\ F_2(Q) = \Delta Q^2 / K \end{cases} \tag{10}$$

where $F_2(P)$ and $F_2(Q)$ were the constraint gravitational force on the skeletons of the clusters P and Q, respectively, which were subject to their own initial positions. ΔP and ΔQ were the offset of the skeletons of the clusters P and Q, respectively, which were subject to the force to move and were calculated as Equations (11) and (12).

$$\begin{cases} \Delta P = (K - D_{P,Q}) \times d_P \\ d_P = n_Q / (n_P + n_Q) \end{cases} \tag{11}$$

where d_P was the conditioning operator. n_P and n_O were the number of samples included in the clusters P and Q, respectively.

$$\begin{cases} \Delta Q = (K - D_{P,Q}) \times d_Q \\ d_Q = n_P / (n_P + n_Q) \end{cases} \tag{12}$$

where d_Q was the adjustment operator. According to the offset of the clustering skeleton movement, the position of the clustering center after the movement could be calculated as Equations (13) and (14).

$$c_1.disp = c.disp + (\delta / dist(c) * F_1(c)) \tag{13}$$

where c_1 was the position of the cluster center after moving under the repulsive force. c was the position of the cluster center before the repulsive force. δ was the integrated offset.

$$c_2.disp = c.disp + (\delta / dist(c) * F_2(c)) \tag{14}$$

where c_2 was the position of cluster center after moving under gravitational force. c was the position of cluster center before gravitational force. Based on the reference point of the skeleton, the gravitational force calculation model within the cluster was then constructed, which allowed for determining the position of the data line's control point within the cluster.

$$E = \alpha \cdot E_1 + (1 - \alpha) E_2 \tag{15}$$

where E was the total energy value of the data curve in the gravitational factory within the cluster. E_1 was the energy value when the curve was in its own initial position. E_2 was the energy of the curve subject to the gravitational field of the clustering center. α was the intra-cluster adjustment factor, and the smaller the α value, the more centralized the binding of the wire bundle.

The flowchart of the visual analysis model of ecological and economic data with improved clustering data based on the force-directed edge binding technique of segmented skeleton was demonstrated in Figure 4, where clustering was performed on the multi-attribute data using the k-means algorithm to obtain the number of clusters, cluster centers, and cluster sizes. Subsequently, the offset was calculated, and the annealing algorithm was simulated to iteratively calculate the equilibrium position for each cluster center under repulsive and gravitational forces to regulate the offset. Subsequently, a three-stage skeleton calculation was performed to control the intra-cluster curves, and smooth curves were plotted to complete the data visualization transformation.

Test of the developed model

To test the performance of the improved technique, Parallel coordinates plot (PCP) and A Data Scalable Approach for Identifying Relationships in Parallel Coordinates (DSPCP) as the control group were employed in the technique performance comparison test. The test used part of the data from United Nations Auntie Mo's database as the experimental dataset. The data relevance index (Cr), data dispersion index (Id), location distribution index (Is), cluster center index (Ic) were applied as evaluation indexes in the comparison test. To verify the practical application of the proposed visual analysis model of ecological and economic data, 20 volunteers with 10 males and 10 females from the age of 20-25 years old and the same level of education were recruited to take the test. All participants had no color vision impairment and no prior experience in data visualization and analysis.

Results and discussion

Comparative performance analysis of improved technologies

The Cr indicator reflects the degree of correlation between different variables or features in a data set. The Cr value test results of the improved technology, PCP technology, and DSPCP technology were shown in Figure 5. The correlation coefficients of the data obtained from the three visualization techniques ranged between [-1.0, 1.0], reflecting strong linear correlation. The Cr value of the proposed improved technique was 0.145, while the Cr values of PCP and DSPCP techniques were both 0.138, which indicated that the Cr value of the improved technique was the highest one, and its performance was the best one. Id is the Pearson's correlation coefficient between the clustering data distribution coefficient (Dc), using the coefficient of variation (CV). The CV measures the dispersion of data points in a dataset relative to their mean value, reflecting the variation between data points. When the positive linear correlation between Dc and CV is stronger, the Id value is higher. The higher Id value indicates a better response to the degree of dispersion of the data distribution. The correlation between Dc and CV of the improved technique was obvious and showed a positive and strong linear correlation, i.e., the improved technique had the highest value of Id, which better responded to the degree of discretization of the data distribution and was better than that of the control group (Figure 6). The evaluation results of the range of Ic values for the improved technique showed that the improved technology delivered actual Ic values of 30-60, 60-90, 15-35, and 50-80 in clusters of 1, 2, 4, and 6, respectively. The range of Ic assessed values was within a similar range of 30-55, 60-90, 15-35, and 55-80, respectively, showing a minimal deviation from the actual values. The range of Ic actual values for the PCP technology were 45-80, 15-40, 50-80, and 40-80, respectively, and the estimated range of Ic assessed values were 50-70, 15-40, 60-100, and 45-75, respectively, which showed the deviation between the assessed and actual values of Ic for the PCP technology was small. However, the actual Ic values for the DSPCP

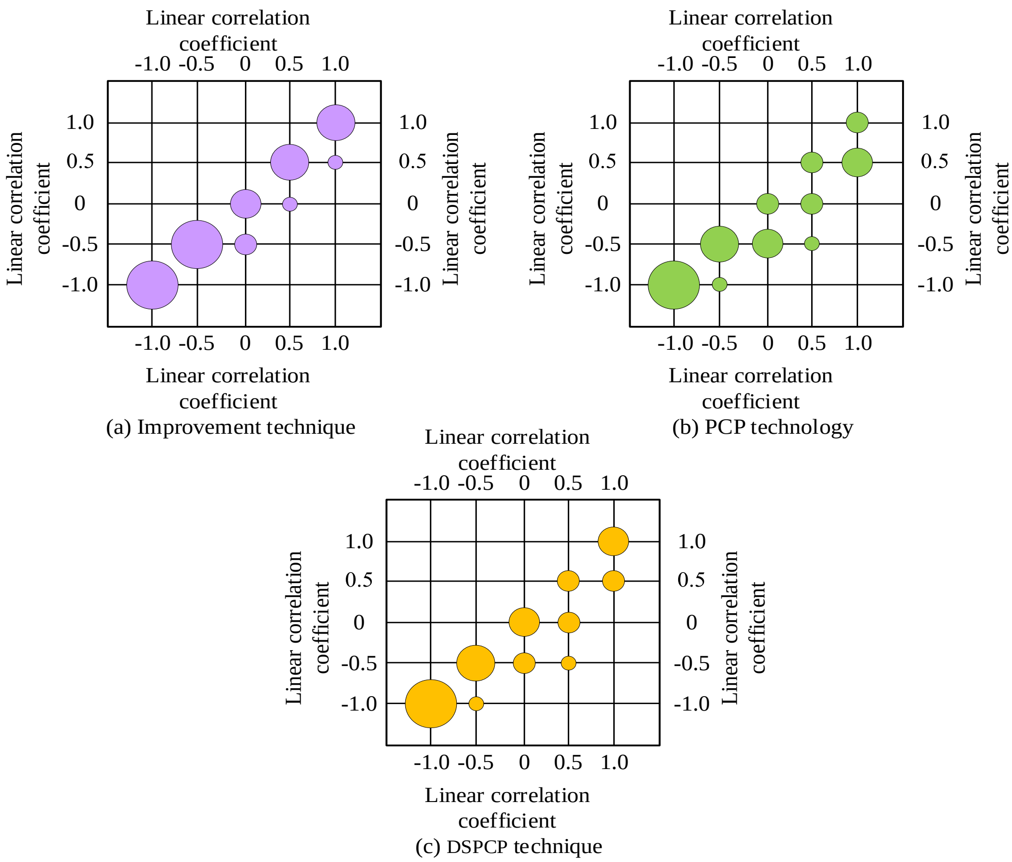

Figure 5. Assessment data of the index Cr.

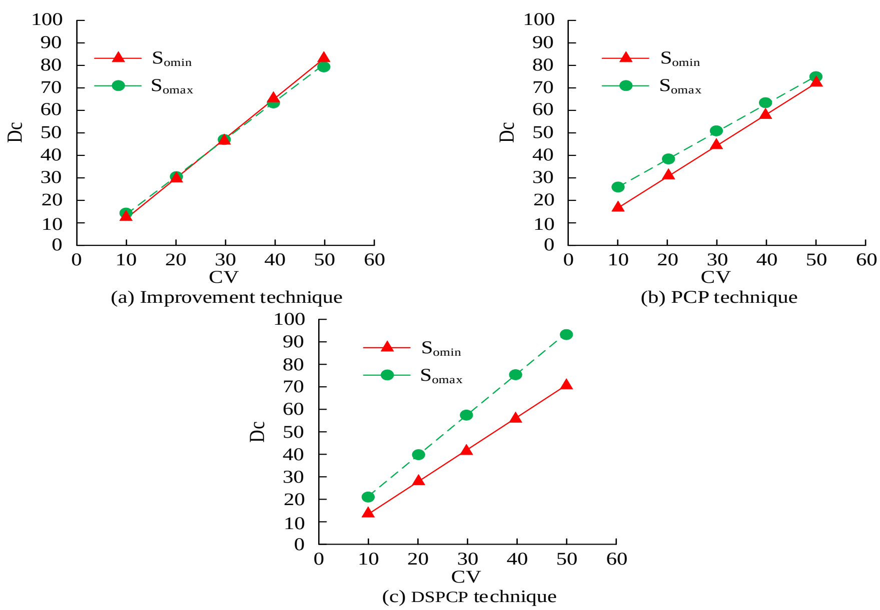

Figure 6. Dc assessed values and CV distribution fitting.

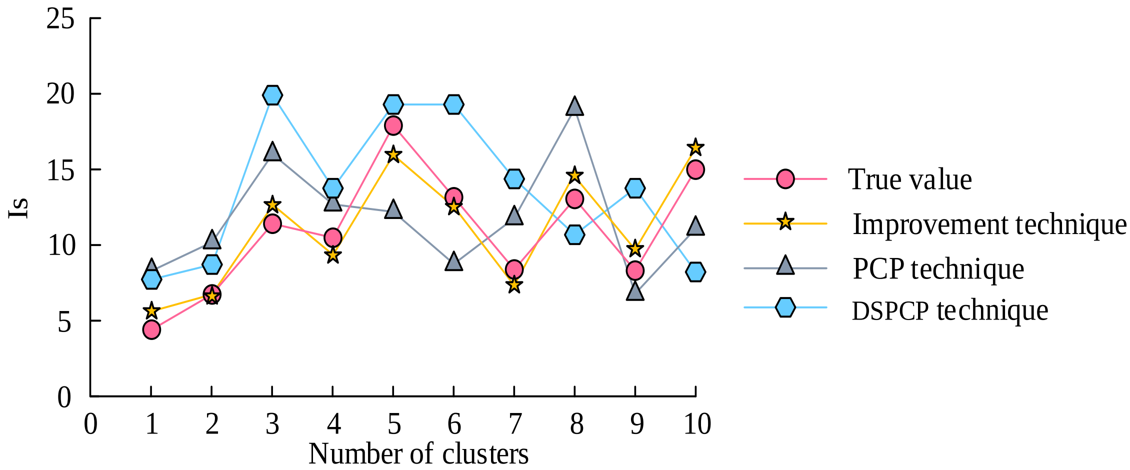

Figure 7. Comparison of the location distribution accuracy of the three technologies.

technology ranged from 25-45, 70-85, 10-30, and 0-30, respectively, while the estimated Ic ranges were 40-70, 30-45, 10-30, and 25-60 demonstrating a larger deviation between the estimated and the actual Ic values of the DSPCP technology. The smaller deviation between the Ic estimated and the actual values indicated that the accuracy for the center values was higher, resulting in a more precise representation of the cluster's central position. Therefore, the results suggested that the performance of the improved technique was better than the other techniques. Is metrics are often used to assess the uniformity of the distribution of data points in each area or space, which is particularly useful when analyzing geographic or spatial data. The results of the Is value test for the improved technique, PCP, and DSPCP techniques demonstrated that, in the case of different number of clustered clusters, the true values of Is were 4.5, 6.8, 12.3, 10.5, 17.5, 13.0, 8.0, 13.0, 8.0, and 15.0, respectively. The values of Is using the improved technique were 5.8, 6.8, 12.8, 8.5, 16.0, 12.5, 7.0, 14.9, 10.5, and 16.0, respectively, which were the closest to the true values (Figure 7). Therefore, the accuracy of the clustered data zones after using the improved technique was the highest one. The results of the evaluation indexes confirmed that the research proposed improved k-means clustering analysis technique based on FDSBB had the best performance and superiority.

Effectiveness of the ecological-economic visual analysis model

The ecological and economic visual analysis models were constructed based on the improved technology, the PCP technology, and the DSPCP technology, respectively. 20 volunteers were involved to assess the correctness of the models under the change of Cluster number index (Nc). The results showed that, when the values of Nc were 5, 6, 7, 8, and 9, the improved model had correct volunteer assessment rates of 81%, 85%, 90%, 95%, and 95% in the case of very large values of the coefficients, respectively. In the case of very small values of coefficients, the correct rates of volunteer assessment were 77%, 79%, 86%, 92%, and 93%, respectively (Figure 8a). In addition, the PCP model in the case of very large values of the coefficients demonstrated the volunteers’ assessed correct rate of 75%, 77%, 80%, 83%, and 88%, respectively, while, in the case of very small values of the coefficients, the correct rates of volunteer assessment were 72%, 72%, 75%, 77%, and 82%, respectively (Figure 8b). The DSPCP model in the case of very large values of the coefficients showed the volunteers’ assessed the correct rates of 69%, 72%, 75%, 79%, and 79%, respectively, while, in the case of very small values of the coefficients, the correct rates of volunteers' assessment were 68%, 70%, 72%, 77%, and 77%, respectively (Figure 8c). The improved model yielded the highest rate of

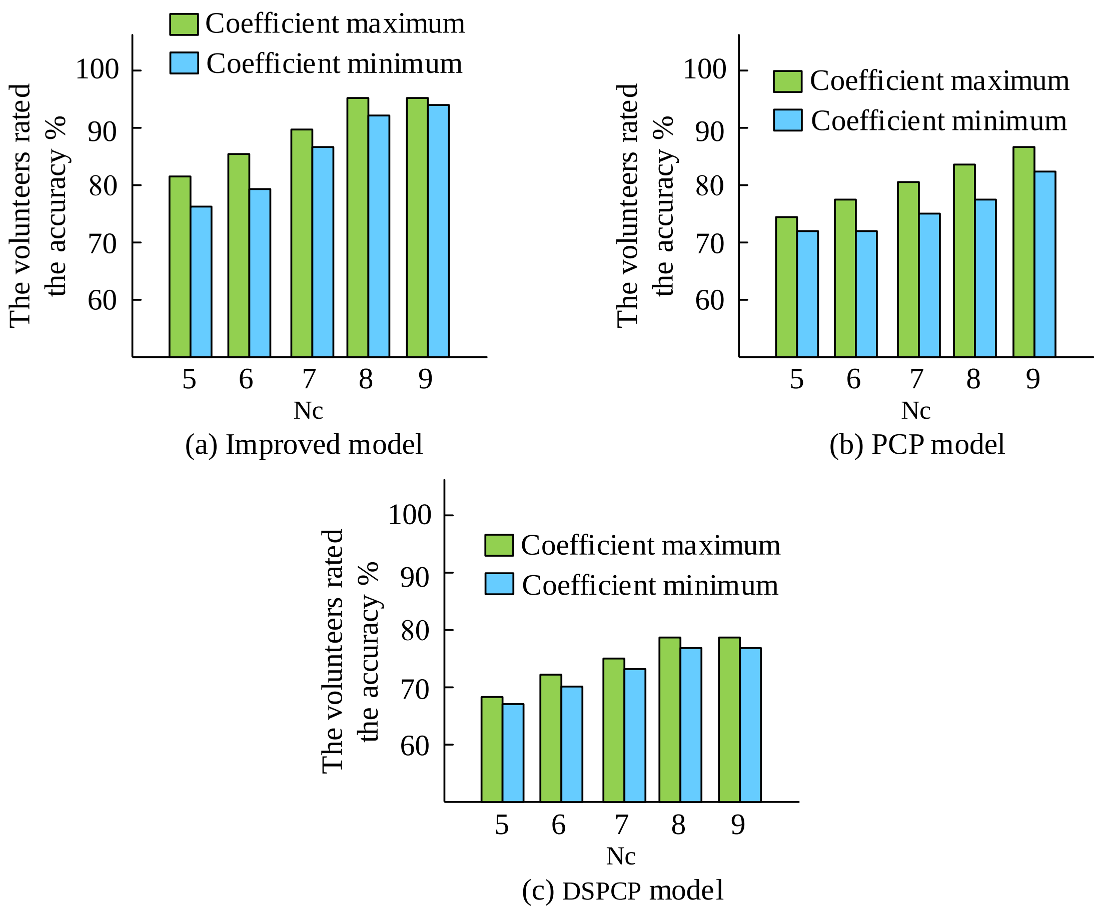

Figure 8. The correct rates of volunteer assessment.

correct volunteer assessments, indicating that it was superior to the other models.

The clustering quantity accuracy error (An) of the improved model, PCP model, and DSPCP model were also tested. The smaller the range of An value, the closer the model to the real data. The results showed that the An value of the improved model fluctuated around 0 with the maximum positive and negative An values of 0.012 and -0.003 among 40 samples, respectively. The An value of PCP model fluctuated around [0.01, 0.02] and [-0.02, -0.01] with the maximum positive and negative An values of 0.022 and -0.016 among 40 samples. The An value of DSPCP model fluctuated in the range of [0.01, 0.02], [-0.02, -0.01] and its maximum positive and negative An values in 40 samples were 0.022 and -0.018, respectively (Figure 9). Therefore, the use of the improved model was better than the other models. In conclusion, the ecological and economic visual analysis model constructed based on the improved k-means cluster analysis technique of FDSBB proposed in the study, is practical and more effective than the comparative model.

Conclusion

In today's global economic integration and informationization, ecological economic analysis has become a popular topic. Data visualization technology can provide powerful technical support for analysts to better understand and interpret the laws and connections existing in the data. This study focused on improving the k-means cluster analysis technique using FDSBB

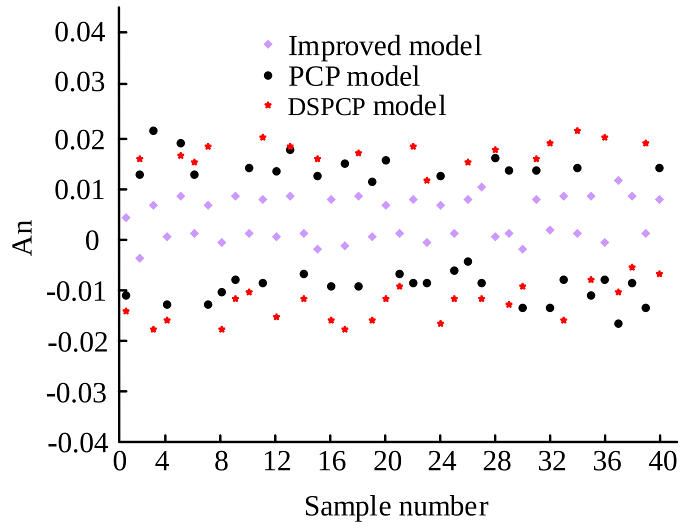

Figure 9. The comparison of the accuracies in Improved, PCP, and DSPCP models.

and constructed a visual analysis model for ecological and economic data based on the improved technique. To assess the effectiveness of the improved method, the study conducted a comparative test and found that the ecological economic visualization and analysis model built upon the upgraded FDSBB k-means cluster analysis technique outperformed alternative models. The findings demonstrated that the model offered more precise and extensive data analysis, thus enabling better-informed decisions regarding the development and sustainability of the ecological economy. However, data visualization is a comprehensive and multi-disciplinary topic, and the extended research on the involved fields will be strengthened subsequently.

References

1. Sharma S, Verma K, Hardaha P. 2023. Implementation of artificial intelligence in agriculture. JCCE. 2(2):155-162.
2. Bataka H. 2023. Economic globalization and public debt in Sub-Saharan Africa. Int J Financ Econ. 28(2):1756-1771.
3. Taesopapong S, Ratanatamskul C, Pornchaiwiseskul P. 2020. Economic analysis of the innovative eco-biofilter/membrane bioreactor (MBR) system for community wastewater recycling. Eng J. 24(5):41-54.
4. Lueddeckens S. 2023. A review on the handling of discounting in eco-efficiency analysis. Clean Technol Environ Policy. 25(1):3-20.
5. Wu M. 2022. Intensity of eco-environmental constraints and green innovation efficiency-based on spatial economic analysis. EEMJ. 21(12):2061-2071.
6. Cuimei L, He Y, Zhang W, Changkuan G, Yang L, Denghua Y. 2021. Quantitative analysis of eco-economic benefits of urban reclaimed water greening based on emergy theory. Water Resour Manage. 35(14):5029-5047.
7. Song M, Xie Q. 2021. Evaluation of urban competitiveness of the Huaihe River eco-economic belt based on dynamic factor analysis. Comput Econ. 58(3):615-639.
8. Yang Y, Meng G. 2020. Analysis of the decoupling effect and eco-economic coordination of the resident energy carbon footprint: a case study of Caijiapo Town, a national key town in western China. Environ Sci Pollut Res. 27(7):6936-6949.
9. Truong ST, Barreto H. 2023. Teaching income inequality with data-driven visualization. AEX. 68(1):140-155.
10. Wu Y, Yang Y, Cao Y. 2023. GPU-based adaptive data reconstruction for large-scale statistical visualization. J Visualization. 26(4):899-915.
11. Novak WRP. 2022. Investigating evolutionary relationships through cluster analysis: A teaching science with big data workshop session. Biochem Mol Biol Educ. 50(5):440-445.
12. Patton AJ, Weller BM. 2023. Testing for unobserved heterogeneity via k-means clustering. J Bus Econ Stat. 41(3):737-751.
13. Xue L. 2023. Cluster analysis-based big data mining method of e-commerce consumer behavior. Int J Web Based Communities. 19(1):53-63.
14. Lethikim N, Vovan T. 2022. Fuzzy cluster analysis for interval data based on the overlap distance. Int J Uncertainty Fuzziness Knowledge Based Syst. 30(4):625-648.

15. Lester JN, Cho Y, Lochmiller CR. 2020. Learning to do qualitative data analysis: A starting point. Hum Resour Dev REV. 19(1):94-106.
16. Lund B, Ma J. 2021. A review of cluster analysis techniques and their uses in library and information science research: k-means and k-medoids clustering. Perform Meas Metrs. 22(3):161-173.
17. Feldman D, Schmidt M, Sohler C. 2022. Turning big data into tiny data: Constant-size coresets for k-means, PCA, and projective clustering. Siam J Comput. 49(3):601-657.
18. Sun Y, Wang N. 2022. Sustainable evaluation of the eco-economic systems in the π-shaped Curve Area of the Yellow River basin of China: a study based on the 3D ecological footprint model. Environ Sci Pollut Res. 29(50):75921-75935.