frontiers | Frontiers in Environmental Science

TYPE Original Research
PUBLISHED 10 January 2023
DOI 10.3389/fenvs.2022.1075943

OPEN ACCESS

EDITED BY
Rita Yi Man Li,
Hong Kong Shue Yan University, Hong Kong SAR, China

REVIEWED BY
Elchin Suleymanov,
Baku Enginering University, Azerbaijan
Mukesh Kumar Awasthi,
Northwest A&F University, China

*CORRESPONDENCE
Zhongxu Pi,
119076906@qq.com

SPECIALTY SECTION
This article was submitted to Environmental Citizen Science, a section of the journal Frontiers in Environmental Science

RECEIVED 21 October 2022
ACCEPTED 22 December 2022
PUBLISHED 10 January 2023

CITATION
Feng H and Pi Z (2023), Research on the path to improve the efficiency of government social governance based on data mining technology under the background of carbon neutrality.
Front. Environ. Sci. 10:1075943.
doi: 10.3389/fenvs.2022.1075943

Research on the path to improve the efficiency of government social governance based on data mining technology under the background of carbon neutrality

Hao Feng[1] and Zhongxu Pi[2]*

[1]School of Information Engineering, Wuhan College, Wuhan, China, [2]Department of Public Security Administration, Hubei University of Police, Wuhan, China

Using data mining to improve the efficiency of government governance in the context of carbon neutrality is an important way to achieve the modernization of the national governance system. This study starts with the logic of carbon neutral issues, analyzes the factors and indicators that affect the effectiveness of social governance, and constructs the evaluation index system of government social governance efficiency based on data mining application under the background of carbon neutral, including *per capita* GDP, *per capita* domestic power consumption of residents, *per capita* CO_2 emissions, *per capita* green area, industrial waste gas treatment rate, industrial wastewater discharge compliance rate and other indicators, which includes 4 first-class indicators, 19 second-class indicators and 38 third-class indicators. Then, the CV-CRITICAL (coefficient of variation critical) index weight determination algorithm is used to determine the index weight. The Pearson correlation coefficient method is used to evaluate the correlation between the two vectors, and then the rationality of the government social governance efficiency evaluation index system based on data mining applications is evaluated. The evaluation results show that the level of social governance effectiveness of the Chinese government is on the rise from 2016 to 2021. This study promotes the application of improving the efficiency of government social governance in the context of carbon neutrality, and provides tools for relevant assessment through data mining technology. This research can not only deepen the theoretical connotation of government governance effectiveness, but also help promote the application of big data in government governance practice.

KEYWORDS

carbon neutrality, data mining, governance efficiency, index system, weight allocation, CV-CRITICAL, Pearson correlation coefficient

1 Introduction

Carbon neutrality refers to the total amount of greenhouse gas emissions directly or indirectly generated by enterprises, groups or individuals within a certain period of time, offset by afforestation, energy conservation and emission reduction, and achieve "zero emission" of carbon dioxide. As China's economy enters the stage of high-quality development, carbon peak and carbon neutrality have been mentioned repeatedly in important conferences. Since then, the construction of beautiful China has really had a time node and a stage goal. However, as a profound economic and social systemic change, the realization of carbon peak and carbon neutrality is not smooth sailing.

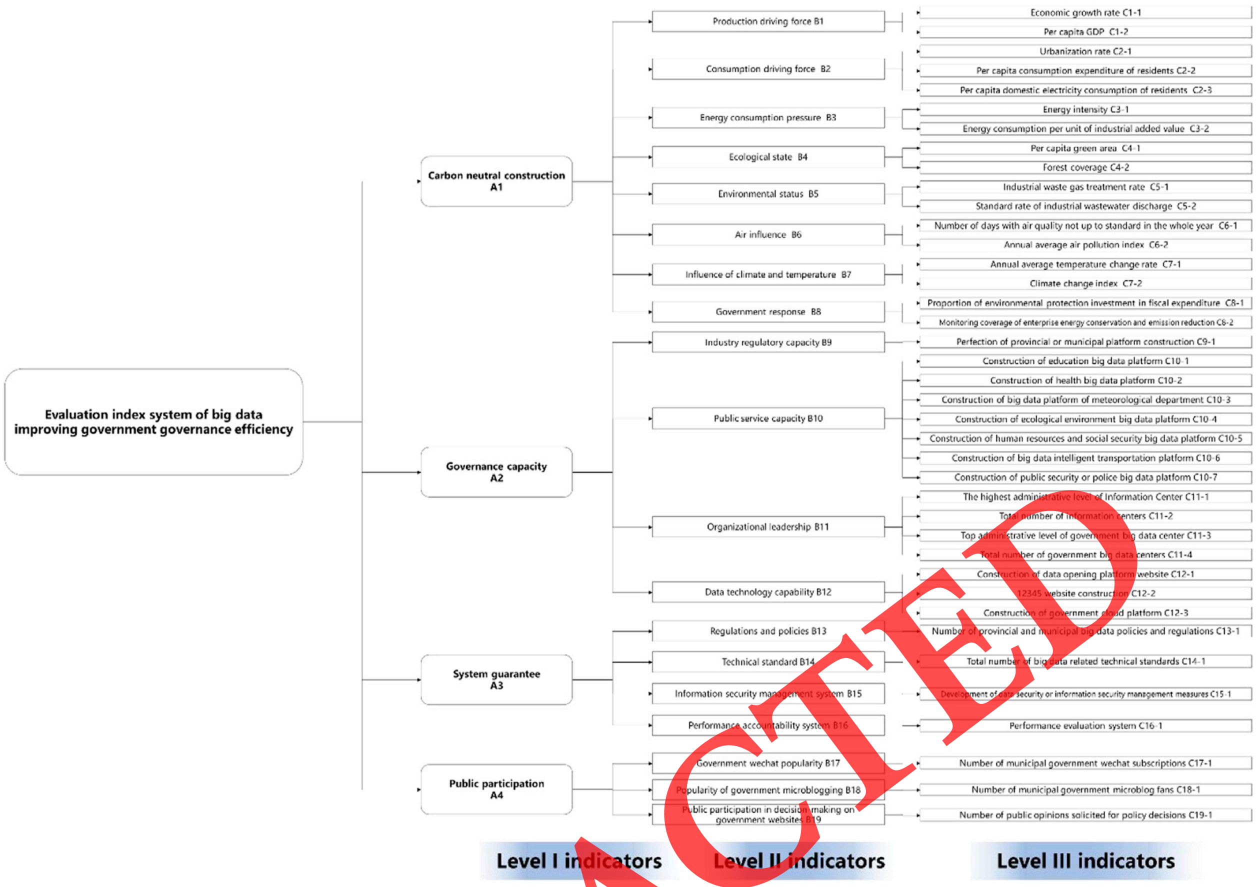

FIGURE 1
Evaluation index system of government social governance efficiency based on data mining application.

Research shows that technological progress has increased the average distance traveled by cars using gasoline per gallon in the United States by 30% since 1980, but has not reduced the total amount of gasoline used (Polimeni et al., 2008). Juliet believes that in the past 35 years, the energy consumption per unit of GDP has been halved, but the demand for energy has increased by 40% instead of decreasing (Foster et al., 2010). Therefore, to resolve the "jevans paradox" and achieve carbon peak and carbon neutrality, we need to change the perspective, that is, from the perspective of government social governance efficiency.

In recent years, with the rapid development of data mining and data mining technology, the cost of data collection, development and utilization has been continuously reduced, and the application field has become increasingly wide. It is regarded as a strategic resource in the new era by many countries. The United States launched the "data mining research and development initiative" in 2012 (Jian and Rui, 2022) and formulated the "federal data mining research and development plan" in 2016 (The Networking And Information Technology Research And Development Program and Group B, 2016); The EU launched the "Data-Driven economy" strategy in 2014 (Cavanillas et al., 2016). In addition, Britain, Japan, and Australia and other countries have also issued a series of policies to promote the application of data mining and industrial development. In 2015, the State Council of China issued the action plan for promoting the development of data mining (Liu and Wang, 2017), and in 2016, the Ministry of industry and information technology formulated the data mining industry development plan (2016-2020) (Wang, 2020). In 2018, the Guangdong Provincial Government of China issued the master plan for the construction of "digital government" in Guangdong Province (2018-2020), which proposed to build a "digital government" with high standards.

In the meantime, China makes much account of the improvement of government governance efficiency. From the 1980s to 2018, China carried out seven large-scale institutional reforms to control the size of the government. In April 2016, Xi'an city launched the "government administrative efficiency revolution". In 2017, Shanghai issued the Trial Measures of the Shanghai municipal government on the management of efficiency construction, which requires that "efficiency construction units should establish a database of basic information in relevant management fields. Efficiency construction units should share management data according to law, cancel or simplify duplicate materials or forms that need to be provided by relevant management personnel, and reduce administrative costs." In 2018, Chinese Premier Li Keqiang proposed to improve the government's efficiency in an all-round way, calling for "optimizing the government's institutional setup and functional allocation, deepening institutional reform, forming a government governance system with clear responsibilities and administering according to law, and enhancing the government's credibility and executive power".

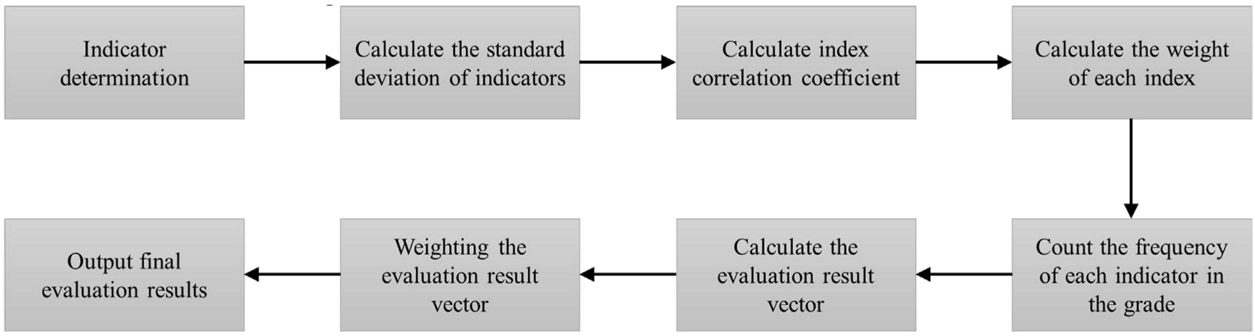

FIGURE 2
Flow chart of weighting based on critic method.

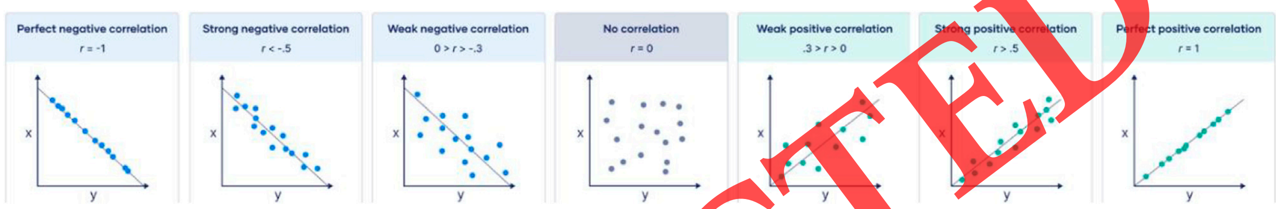

FIGURE 3
Visualization of Pearson correlation coefficient.

The rapid development of data mining technology provides a new technical means to improve the efficiency of government governance. At present, Chinese government departments have made some progress in using big data to improve governance.

In order to better guide the government to use data mining technology to improve the government's social governance efficiency, this study intends to use value focus thinking (VFT) to build an evaluation index system for data mining to improve the government's social governance efficiency, and then use CV-CRITICAL index weight determination algorithm to determine the index weight, using data mining technology to build a scientific and reasonable government's social governance efficiency, So as to promote the application of government social governance efficiency improvement in the context of carbon neutrality.

2 Related works

Governance is a concept with evolving connotation, which has evolved from the meaning of "rule" and "control" in the early days to the widely accepted meaning of "multi-agent cooperative management of public affairs" (Sun). The purpose of governance is to use power to guide, control and regulate various activities of citizens in various institutional relationships so as to maximize public interests (Chen, 2016). In 1997, in its report entitled "decentralized governance: strengthening people centred development capacity", the United Nations summarized 15 core concepts of good governance pursued by the contemporary governance movement, including participation, openness, transparency, response, fairness, responsibility, and legitimacy (Spanhove and Verhoest, 2007).

Most of the existing studies on efficiency building are carried out from the perspective of corporate and organizational governance, while there are few studies on public domain or government governance performance. In 1993, the national performance review committee of the United States established a performance evaluation system for the staff of the government and its functional departments (Ofpresident, 1993); In 2002, the British Government Audit Commission issued a performance evaluation framework for local governments, including three parts: resource utilization, service evaluation and municipal authority evaluation (The Audit Commission and UK, 2019); Charles et al. Proposed governance performance indicators of coastal management in terms of management objectives and different management stages (initiation, planning, and adoption), including authority, leadership, vision, institutional capacity, human resource development, empowerment, financial resource capacity, planning capacity, information management capacity, public participation, formalization, and support (Ehler, 2003); Ryzin et al. (2004) Evaluated the public's satisfaction with the New York government from the perspective of customer satisfaction. The public governance, performance and Accountability Act published by the Australian government in 2013 stipulates that the annual performance statements of federal entities should include statements, results and analysis (Australia, 2019). In 2016, the world bank proposed the local governance performance index (lgpi), including education, health, physical safety and dispute resolution, social assistance and welfare,

10.3389/fenvs.2022.1075943

TABLE 1 Weight calculation results of indicator system

Target	Level I indicators	Weight	Level II indicators	Weight within the primary indicator	Final weight	Level III indicators	Weight in this secondary indicator	Final weight of level III indicators
Evaluation index system of big data improving government governance efficiency	Carbon neutral construction A1	0.4729	Production driving force B1	0.1223	0.0578	Economic growth rate C1-1	0.50	0.0289
						Per capita GDP C1-2	0.35	0.0202
			Consumption driving force B2	0.0900	0.0425	Urbanization rate C2-1	0.15	0.0087
						Per capita consumption expenditure of residents C2-2	0.55	0.0234
						Per capita domestic electricity consumption of residents C2-3	0.45	0.0191
			Energy consumption pressure B3	0.0981	0.0464	Energy intensity C3-1	1.00	0.0464
						Energy consumption per unit of industrial added value C3-2	0.05	0.0039
			Ecological state B4	0.1649	0.0780	Per capita green area C4-1	0.40	0.0312
						Forest coverage C4-2	0.15	0.0117
			Environmental status B5	0,1755	0.0830	Industrial waste gas treatment rate C5-1	0.20	0.0166
						Standard rate of industrial wastewater discharge C5-2	0.35	0.0290
			Air influence B6	0.1599	0.0756	Number of days with air quality not up to standard in the whole year C6-1	0.45	0.0373
						Annual average air pollution index C6-2	1.00	0.0756
			Influence of climate and temperature B7	0.0567	0.0268	Annual average temperature change rate C7-1	1.00	0.0268
						Climate change index C7-2	0.40	0.0312
			Government response B8	0.1327	0.0627	Proportion of environmental protection investment in fiscal expenditure C8-1	1.00	0.0627
						Monitoring coverage of enterprise energy conservation and emission reduction C8-2	0.35	0.0290
	Governance capacity A2	0.2487	Industry regulatory capacity B9	0.1442	0.0358	Perfection of provincial or municipal platform construction C9-1	1.00	0.0358
			Public service capacity B10	0.4748	0.1181	Construction of education big data platform C10-1	0.12	0.0147
						Construction of health big data platform C10-2	0.17	0.0196
						Construction of big data platform of meteorological department C10-3	0.12	0.0147
						Construction of ecological environment big data platform C10-4	0.12	0.0147
						Construction of human resources and social security big data platform C10-5	0.12	0.0147
						Construction of big data intelligent transportation platform C10-6	0.17	0.0196
						Construction of public security or police big data platform C10-7	0.17	0.0198
			Organizational leadership B11	0.1197	0.0298	The highest administrative level of Information Center C11-1	0.17	0.0050

(Continued on following page)

10.3389/fenvs.2022.1075943

TABLE 1 (*Continued*) **Weight calculation results of indicator system**

Target	Level I indicators	Weight	Level II indicators	Weight within the primary indicator	Final weight	Level III indicators	Weight in this secondary indicator	Final weight of level III indicators
						Total number of information centers C11-2	0.25	0.0074
						Top administrative level of government big data center C11-3	0.25	0.0074
						Total number of government big data centers C11-4	0.33	0.0099
			Data technology capability B12	0.2614	0.0650	Construction of data opening platform website C12-1	0.40	0.0260
						12345 website construction C12-2	0.40	0.0260
						Construction of government cloud platform C12-3	0.20	0.0130
	System guarantee A3	0.1698	Regulations and policies B13	0.3624	0.0615	Number of provincial and municipal big data policies and regulations C13-1	1.00	0.0615
			Technical standard B14	0.1952	0.0332	Total number of big data related technical standards C14-1	1.00	0.0332
			Information security management system B15	0.1729	0.0294	Development of data security or information security management measures C15-1	1.00	0.0294
			Performance accountability system B16	0.2694	0.0458	Performance evaluation system C16-1	1.00	0.0458
	Public participation A4	0.1086	Government wechat popularity B17	0.2222	0.0241	Number of municipal governments wechat subscriptions C17-1	1.00	0.0241
			Popularity of government microblogging B18	0.1111	0.0121	Number of municipal government microblog fans C18-1	1.00	0.0121
			Public participation in decision-making on government websites B19	0.6667	0.0724	Number of public opinions solicited for policy decisions C19-1	1.00	0.0724

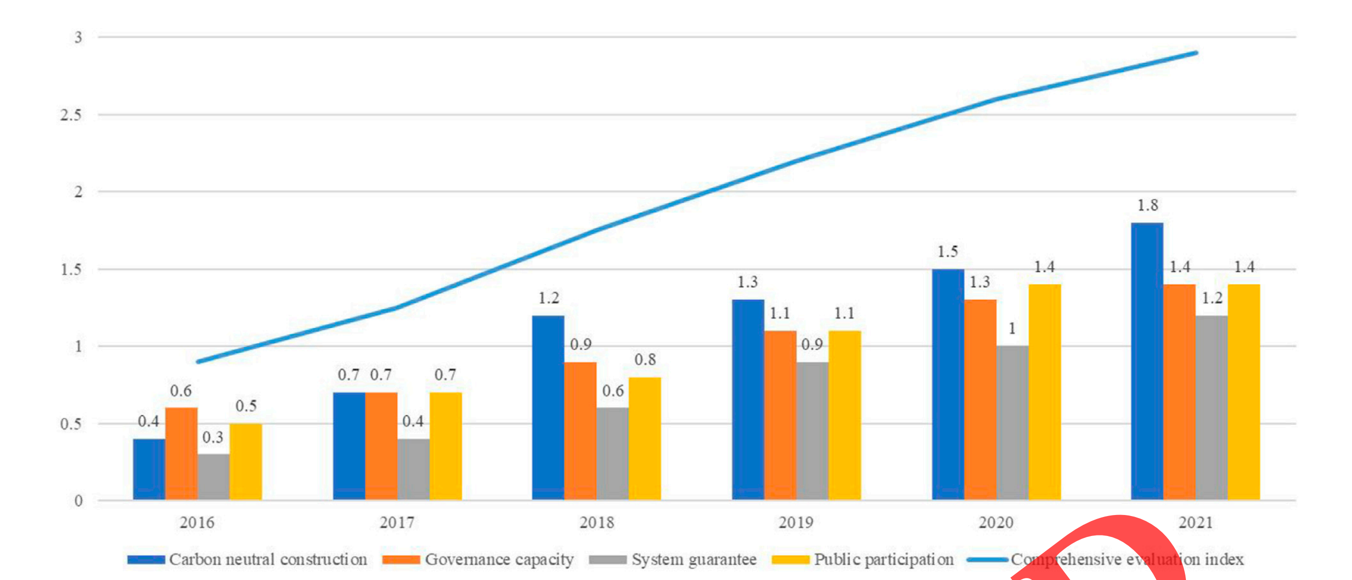

FIGURE 4
Change trend of four primary assessment indicators and comprehensive assessment indicators from 2016 to 2021.

citizen state ties and corruption, social composition, and culture (Bank, 2016).

Edwin based on the data of 66 cities (including 50 in England, 5 in Wales, 6 in Scotland, and 5 in Northern Ireland) published by the British government, used the indicator of the proportion of clean energy in primary energy to evaluate the construction level of low-carbon cities in 66 cities. The larger the proportion, the higher the level of low-carbon cities (Chan et al., 2013). Rory assessed the construction level of 10 low-carbon project cities (Geneva, Abu Dhabi, Stockholm, etc.) of the United Nations by using eight indicators, including harmless garbage disposal rate, carbon productivity, carbon emissions per unit building area, and forest cutting and storage proportion, according to the world bank data in 2010. The results show that Stockholm has the highest construction level of low-carbon cities (Sullivan et al., 2013).

With the advent of the era of data mining, the issue of using data mining to improve the efficiency of government social governance has received the attention of some scholars. Zhang believes that using data mining technology can improve the scientificity of government governance decisions, enhance the effectiveness of government governance, and improve the performance evaluation and application of government governance (Zhang et al., 2016); Chen believes that data mining can help government departments master the needs and attitude preferences of Internet users, judge the effect of early administration, adjust and optimize public policies, reengineer management service processes, and improve the government's awareness, response and Governance (Chen and He, 2020); Pannu proposed that in the era of data mining, the knowledge-based government has the advantages of adaptability, operability and intelligence (Pannu et al., 2016).

The above research discriminates the connotation and structural dimension of government governance and its efficiency, and discusses the role, realization way and influence mode of data mining to improve the efficiency of government and social governance, laying a theoretical foundation for the evaluation research of data mining to improve the efficiency of government and social governance.

3 Construction of evaluation system of government's social governance efficiency

3.1 Evaluation system construction method

When designing the evaluation indicators, this study used the gold standard of value Focused Thinking (VFT) (Parnell, 2007), extracted the fundamental value goal of data mining to improve the efficiency of government social governance from the national policy documents, and converts it into evaluation indicators.

Through policy analysis, literature research and case studies, this study extracted the indicators of government social governance efficiency that can be improved by data mining in the context of carbon neutrality, such as energy consumption, economic development, population structure, market supervision, public satisfaction, etc. In terms of policy analysis, this study mainly selected data mining development strategic plans and important speeches of leaders at the national level of China, the United Kingdom and the United States as the analysis objects. China's policy texts include the outline of action for data mining development issued by the State Council in 2015 (Zhengce, 2015), the data mining industry development plan (2016-2020) (MIIT, 2017) formulated by the Ministry of industry and information technology in 2016, and the speeches of state leaders on data mining, Internet + and government efficiency on December 9, 2017 and April 2018; The policy text of the United States is mainly the data mining research and development strategic plan issued by the federal government in 2012 (Bemelmans, 1979); The UK's policy text is mainly "seizing the opportunities brought by data: UK data capability strategy" published in 2014 (Abbasi et al., 2016). Analyze the content of the above-mentioned policies and speech texts, and extract the strategic objectives of the government to use data mining to improve governance efficiency.

By comparing and combining the above analysis results, we conclude that the expected main objectives of the government's use of data mining to improve governance efficiency are energy conservation and emission reduction, economic development, social governance, government services, market supervision, improvement of people's livelihood, scientific research and innovation, talent training, scientific decision-making, data sharing, law and system construction.

In terms of literature research, the current evaluation research related to government social governance efficiency mainly includes four categories: environmental governance evaluation, social governance evaluation, government performance evaluation and government efficiency evaluation. The worldwide governance indicators (WGI) divide governance into seven dimensions: climate change response, voice and accountability, political stability and violence avoidance, government effectiveness, control quality, legal system, and corruption prevention (Kaufmann et al., 2011). The China social governance index (CSGI) proposed by Farooque includes seven secondary indicators: environmental pollution control, human development, social equity, public services, social security, public security and social participation, as well as 35 tertiary indicators (Farooque et al., 2022). Governments or scholars in different countries have studied the government performance evaluation index system. In 1993, the national performance review committee of the United States established a performance evaluation index system to evaluate the staff of the government and its functional departments, including six indicators: input indicators, energy indicators, output indicators, outcome indicators, efficiency and cost-effectiveness indicators, and productivity indicators (Ofpresident, 1993).

In the case study, on the basis of the above analysis, this study collects typical cases of government governance using data mining in the context of carbon neutrality, and extracts evaluation indicators through case analysis. The cases include "carbon neutral construction + government governance" cases such as the governance of the Qinghai Tibet Plateau based on carbon neutral assessment, the carbon neutral transformation governance of Henan Province, and the ecological environment governance of China's counties. We extracted 21 indicators, including *per capita* GDP, *per capita* domestic power consumption of residents, *per capita* CO_2 emissions, *per capita* green area, industrial waste gas treatment rate, industrial wastewater discharge compliance rate, enterprise energy conservation and emission reduction monitoring coverage The proportion of households using solar, geothermal energy or biogas energy, the annual average air pollution index, and the proportion of environmental protection investment in fiscal expenditure. As well as data mining + government governance cases such as microblogs and online government platforms in Beijing, Shenzhen, Guangdong and Zhengzhou. We extracted 54 indexes, including crime rate, information release efficiency, administrative approval process and transparency, public transport, reemployment time and financial expenditure, number of media platforms of new government media, thematic update frequency, feedback time, interaction frequency User subscriptions, etc.

Based on the theory of government efficiency and government governance, this study analyzes the basic objectives of government governance based on data mining applications by using the value focus thinking method, as shown in Figure 1.

3.2 Determination of evaluation index weight based on CV-CRITIC algorithm

The index weight is an important basis for the application of the evaluation index system, which reflects the value goal of evaluation and has a guiding role. In this study, the CV-CRITICAL algorithm is used to solve the problem of the index weight of the government's social governance efficiency evaluation index system, so that the index weight can be dynamically adjusted to adapt to the change of evaluation needs.

CRITICAL algorithm (Markowitz et al., 1993) is suitable for solving decision problems with multiple objectives. This algorithm belongs to a qualitative and quantitative method to determine the index weight. The weight assignment flow chart based on CRITIC method is shown in Figure 2.

The main calculation formula of CRITICAL algorithm is shown in (1). Where R represents the correlation coefficient matrix of each index. R_{ij} is the correlation coefficient between index i and index j, m represents the data size of the evaluation object, and n represents the number of evaluation indexes. S_{ij} is used to quantify the conflict between index i and index j. Q_i is the standard deviation of index j, which is used to quantify the impact of the information contained in the index on the index weight. K_i represents the amount of information contained in the index i.

$$K_i = Q_i * S_{ij}, \quad S_{ij} = \sum_{j=1}^{n} \left(1 - R_{ij}\right) \tag{1}$$

According to the basic idea of the algorithm, first calculate the correlation coefficient matrix R of each index, and then calculate the standard deviation vector $Q = (Q_1, Q_2, \ldots, Q_n)$ of the index. Then, according to Eq. 1, the information contained in the indicators is obtained, and the weight value of each indicator is directly determined through discrimination and conflict. The greater the amount of information contained in the indicator ?, the greater the value of K_j, and the greater the relative importance of the indicator. The objective weight vector W_j is obtained by calculating the objective weight W of the index according to Eq. 2 based on the information value CRITICAL of each index.

$$W_j = K_j \Big/ \sum\nolimits_{j=1}^{n} K_j \tag{2}$$

In order to reflect the importance difference between DDV and civ in the process of indicator weight determination, CV-CRITICAL algorithm in the process of indicator weight allocation is as follows: first, the difference coefficient method is introduced to analyze the value of indicator conflict coefficient, and then its important parameter value in the process of indicator determination is calculated, and then the final weight vector based on objective data is calculated. Assuming that the two vectors ? and vector ? have been calculated, the difference coefficient method is used to fully consider the internal data distribution difference of the vector to calculate the undetermined parameter ?, so as to solve the possible error impact caused by traditional methods. The main calculation is divided into the following steps:

The first step is to reorder the components of the ? vector from small to large to obtain the ordered vector ?, where $P = (P_1, P_1, \ldots, P_1)$.

$$Gg = \left(\frac{2}{n}\right)*(1*P_1 + 2*P_2 + \ldots + n*P_n) - \left(n + \frac{1}{n}\right) \quad (3)$$

The second step is to use Eq. 3 to calculate Gg, where Gg is the difference coefficient of each component in the vector ?, and n is the number of indicators.

The third step is to calculate the undetermined parameter PV according to Eq. 4.

$$PV = \left(\frac{n}{n-1}\right)*Gg \quad (4)$$

Finally, the vectors ? and ? are weighted according to Eq. 5 using the calculated ?? parameters to obtain the final index weight vector $W = (W_1, W_2, \ldots, W_n)$.

$$W_i = PV*C_i + (1 - PV)*V_i \quad (5)$$

Through the above analysis, the algorithm analyzes the relationship between the two factors on the index weight through the difference coefficient method instead of simply thinking that the two factors have the same impact on the index. Subsequent experiments also show that this is more reasonable and can effectively improve the accuracy of the weight allocation results.

3.3 Rationality evaluation of index system based on Pearson correlation coefficient

This study uses Pearson correlation coefficient to evaluate the correlation between the two vectors, and then evaluates the rationality of the evaluation index system of government social governance effectiveness based on data mining applications. Pearson correlation coefficient is a linear correlation coefficient, which can be used to reflect the statistics of the linear correlation degree of two variables and describe the degree of linear correlation between two variables (Chen et al., 2022). If its absolute value is larger, it indicates that the correlation between the two vectors is stronger (Ciric et al., 2022).

The Pearson correlation coefficient can be used to measure the closeness of the observation results to the best fit line. As shown in Figure 3, when the slope is negative, the correlation coefficient r is negative. When the slope is positive, the correlation coefficient r is positive. When the correlation coefficient r is 1 or −1, all points fall exactly on the best fit line. When r is greater than 0.5 or less than −0.5, these points are close to the best fit line. When r is between 0 and 0.3 or between 0 and −0.3, these points are far from the best fit line. And when r is 0, the best fit line does not help to describe the relationship between variables.

Suppose there are two sets of data $X = (X_1, \ldots, X_n)$, $Y = (Y_1, \ldots, Y_n)$, then the Pearson correlation coefficient $R = \langle X, Y\rangle / (||X||*||Y||)$ of these two sets of vectors, where $||X||$ represents the length of the vector, $< X, Y >$ represents the inner product of the two vectors.

$$\langle X, Y\rangle = \sum_{i=1}^{n} X_i*Y_i; \quad ||X|| = \sqrt[2]{X_i*X_i} \quad (6)$$

In order to intuitively analyze the influence of the indicator's own information on the indicator weight in the evaluation process, the concept of indicator discrimination is proposed in this study. In order to quantify the discrimination of the index, first construct a data matrix D from the sample data:

$$D = \begin{bmatrix} D_{11} & \cdots & D_{1n} \\ \vdots & \ddots & \vdots \\ D_{m1} & \cdots & D_{mn} \end{bmatrix} \quad (7)$$

Among them, $i = 1, \ldots, m$; $j = 1, \ldots, n$, a total of ? evaluation objects and n evaluation indicators. $D_i = (D_{1i}, \ldots, D_{mi})$ in matrix D represents the data information of index i. Due to the dimensional difference of the index data, this study first uses the proportion method to dimensionless process the index data to obtain the standardized data matrix Z:

$$D = \begin{bmatrix} Z_{11} & \cdots & Z_{1n} \\ \vdots & \ddots & \vdots \\ Z_{m1} & \cdots & Z_{mn} \end{bmatrix} \quad (8)$$

1) Discrimination vector V of index

Calculate the standard deviation of each index according to the data information of each index in the standardized data matrix, and use the following formula to find the discrimination DDV of each index, which is expressed by the discrimination vector, $V = (V_1, \ldots, V_n)$.

$$V_i = Q_i \Big/ \sum_{i=1}^{n} Q_i \quad (9)$$

2) Index conflict coefficient vector ?

In general, each index of the evaluation index system is not independent, but has a certain correlation. Through the analysis and exploration of the index system structure and index connotation, the conflict between each index and other indicators is found. Pearson's correlation coefficient is used to calculate the correlation between indicators, and the conflict coefficient vector is obtained according to Eq. 10, where ? = (?$_1$, . . . , ?$_?$).

$$C_i = T_i \Big/ \sum_{i=1}^{n} T_i, \quad T_i = \sum_{j=1}^{n} \left(1 - R_{ij}\right) \quad (10)$$

This study introduces the difference coefficient method to analyze the value of the indicator conflict coefficient, then calculates the important parameter value PV in the indicator determination process, and then calculates the final weight vector ? based on the objective data. The main calculation steps are as follows:

First, reorder the components of ? vector from small to large to get the ordered vector ?, where $P = (P_1, P_2, \ldots, P_n)$.

Step 2: Calculate G_g with Eq. 11, G_g is the difference coefficient of each component in the vector ?, ? is the number of indicators.

$$G_g = \left(\frac{2}{n}\right)*(1*P_1 + 2*P_2 + \ldots + n*P_n) - \left(n + \frac{1}{n}\right) \quad (11)$$

Step 3: Calculate the undetermined parameter PV according to Eq. 12.

$$PV = \left(\frac{n}{n} - 1\right) * G_g \tag{12}$$

Finally, we use the calculated PV parameter to weight the vectors ? and ? according to Eq. 13 to get the final index weight vector $W = (W_1, W_2, \ldots, W_n)$.

$$W_i = PV * C_i + (1 - PV) * V_i \tag{13}$$

Based on the above analysis, the algorithm analyzes the relationship between the two factors and the index weight through the difference coefficient method, rather than simply thinking that the two factors have the same impact on the index, so as to effectively improve the accuracy of the weight distribution results.

4 Result analysis and discussion

According to the CV-CRITICAL algorithm proposed in this paper, the first step is to calculate the DDV of each indicator to obtain the discrimination vector ?, the second step is to calculate the civ to obtain the conflict vector ?, the third step is to analyze the values of each indicator of the conflict vector ? by using the difference coefficient method, determine the undetermined parameter ?, and finally calculate the weight values of 17 indicators in the evaluation index system to obtain the final weight vector.

In the process of multiple indicator comprehensive evaluation, some indicators are positive evaluation indicators, some indicators are reverse evaluation indicators, and some indicators are indicators whose quantity value is larger and whose performance is better, which are called positive evaluation indicators. Some indicators are indicators whose value is smaller and whose performance is better, which are called moderate evaluation indicators. Different evaluation indicators often have different dimensions and dimension units, so it is inappropriate to directly integrate them. Due to the different units of each indicator of the government's social governance efficiency evaluation factor, there are differences in the dimensions of indicators, which makes it impossible to compare indicators with each other and there are contradictions. The impact can be eliminated through dimensionless, and different types of indicators can be standardized through standardization. The specific implementation methods are as follows:

$$r'_{ij} = \frac{r_{ij} - \min\left(r_{ij}\right)}{\max\left(r_{ij}\right) - \min\left(r_{ij}\right)} \tag{6a}$$

$$r'_{ij} = 1.0 - \frac{\left|\max\left(r_{ij}\right) - r_{ij}\right|}{\max\left(r_{ij}\right)} \tag{7a}$$

In the formula, $\max(r_{ij})$ is the maximum value of data in the evaluation factors of government social governance effectiveness, and $\min(r_{ij})$ is the minimum value of data in the evaluation factors of government social governance effectiveness. Each factor is dimensionless processed by the above formula to prepare for subsequent evaluation.

So far, we have obtained the weights of indicators at all levels of the evaluation index system of data mining to improve the effectiveness of government social governance, as shown in Table 1.

We have established a government social governance efficiency evaluation index system based on data mining applications, including 4 first-class indicators, 19 s-class indicators and 38 third class indicators, and determined the index weight using CV-CRITICAL (coefficient variation CRITICAL) index weight determination algorithm. Among the four first level indicators, governance performance reflects the effect of big data application in government governance, governance capacity and institutional guarantee reflect the ability of the government to use big data, and public participation reflects the characteristics of big data application and the modernization of government governance. The four indicators comprehensively reveal the content of using big data to improve governance effectiveness of the government. When the Pearson correlation coefficient method is used to evaluate the rationality of the evaluation index system of government social governance effectiveness based on data mining applications, the results show that the weight value calculated by CV-CRITICAL algorithm is closer to the real degree of the indicators and is more suitable for determining the weight of the government governance efficiency evaluation indicator system. Because the critical algorithm considers the influence of the index's own information and the correlation between indicators on the index weight, other methods, such as entropy weight method, only consider the influence of the index's own information on the weight. Therefore, the method proposed in this study has strong scientificity and operability.

Then, the specific values of each level I index and comprehensive evaluation index based on the historical data of China from 2016 to 2021 could be calculated. The trend chart is shown in Figure 4. It could be seen that China's carbon neutral construction level is generally on the rise, and the achievements of ecological civilization construction are obvious to all. With the in-depth application of big data technology, the governance ability of the Chinese government has been enhanced year by year, the system guarantee has been gradually improved, and the public participation has been constantly improved.

5 Conclusion

From the new perspective of social governance, this study analyzed the logical way to achieve carbon neutrality. First, we analyzed the evaluation methods for improving the effectiveness of national social governance in the context of carbon neutrality, built an evaluation index system based on data mining technology, combined with theoretical analysis, literature analysis, case studies and other methods, collected typical cases of government governance using data mining in the context of carbon neutrality, and extracted all indicators related to the evaluation objectives, such as *per capita* GDP, *per capita* domestic power consumption of residents, *per capita* CO_2 emissions, *per capita* green area, industrial waste gas treatment rate, industrial wastewater discharge compliance rate, enterprise energy conservation and emission reduction monitoring coverage, proportion of households using solar energy, geothermal energy or biogas energy, annual average air pollution index, proportion of environmental protection investment in financial expenditure. Then, based on the comparison and induction of all relevant indicators, we completed the hierarchical division and system construction of indicators, including 4 first level indicators, 19 s level indicators and 38 third level indicators. Then we use CV-CRITIC (CoefficientVariation-CRITIC) index weight determination algorithm to determine the index weight. Finally, this study uses Pearson correlation coefficient to evaluate the correlation between the two vectors, and then evaluates the rationality of the government social governance efficiency evaluation index system based on data mining applications. The evaluation results showed that the level of social governance effectiveness of the Chinese government is on the rise from 2016 to 2021. This study provides a quantifiable evaluation tool for data mining to improve government social governance, deepens the theoretical research on the effectiveness of government social governance, and helps promote the application of big data in government governance practice. In the future, we will evaluate and study the use of big data by Chinese urban governments to improve government governance efficiency, and dynamically adjust the indicator system and its weight based on different stages of practical development, giving full play to the guiding role of evaluation indicators.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The studies were reviewed and approved by the Hubei University of Police and Wuhan College. The participants provided their written informed consent to participate in this study.

Author contributions

ZP contributed on writing—original draft preparation and data collection, and HF contribution on data preprocessing and design of methodology.

Funding

This work was supported by Hubei Provincial Natural Science Foundation of China: “Research on Customer Satisfaction Mining Based on Big Data”, the project number is ZRMS2019001565.

Acknowledgments

We thank the reviewers whose comments and suggestions helped improve this manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Abbasi, A., Sarker, S., and Chiang, R. H. L. (2016). Big data research in information systems: Toward an inclusive research agenda[J]. *Journal of the association for information systems* 17 (2), 3.

Australia (2019). Public governance, performance and accountability act 2013[EB/OL]. Available at: https://wwwl.egisla-tion.gov.au/Det (Accessed 07 06, 2019).

Bemelmans, T. (1979). Strategic planning for research and development. *Long. Range Plan.* 12 (2), 33–44. doi:10.1016/0024-6301(79)90071-2

Cavanillas, J. M., Curry, E., and Wahlster, W. (2016). *New horizons for a data-driven economy: A roadmap for usage and exploitation of big data in europe*. Midtown Manhattan, New York City: Springer International Publishing. doi:10.1007/978-3-319-21569-3

Chan, E., Choy, L., and Yung, E. (2013). Current research on low-carbon cities and institutional responses. *habitat Int.* 37, 1–3. doi:10.1016/j.habitatint.2011.12.007

Chen, K., Chen, L., Xiao, J. Q., Li, J., Hu, Y., and Wen, K. (2022). Speckle reduction in digital holography with non-local means filter based on the Pearson correlation coefficient and Butterworth filter. *Opt. Lett.* 47, 397. doi:10.1364/ol.444769

Chen, L. J. (2016). *Research on soft power in government governance–the definition and performances of function analysis*. South Bridge, United Kingdom: The University of Edinburgh.

Chen, M., and He, J. (2020). Research on the application of big data technology in government management. *IOP Conf. Ser. Mater. Sci. Eng.* 806 (1), 012001. doi:10.1088/1757-899x/806/1/012001

Ciric, D. G., Peric, Z. H., Milenkovic, M., and Vucic, N. J. (2022). Evaluating similarity of spectrogram-like images of DC motor sounds by Pearson correlation coefficient. *Elektron. Ir. Elektrotechnika* 28 (3), 37–44. doi:10.5755/j02.eie.31041

Ehler, C. N. (2003). Indicators to measure governance performance in integrated coastal management. *Ocean Coast. Manag.* 46, 335–345. doi:10.1016/S0964-5691(03)00020-6

Farooque, O. A., Hamid, A., and Sun, L. (2022). National governance index, corruption index and growth rate—international evidence from sub-saharan and MENA countries. *JRFM* 15, 261. doi:10.3390/jrfm15060261

Foster, J., Clark, B., and York, R. (2010). Capitalism and the curse of energy efficiency: The return of the jevons paradox. *Mon. Rev. Indep. Social. Mag.* 62 (6), 1–12. doi:10.14452/mr-062-06-2010-10_1

Jian, L., and Rui, L. (2022). “The revelation of US government "big data research and development initiative" for China’s scientific platform resource share[C],” in China-Japan-Korea joint symposium on optimization of structural and mechanical systems. 0.

Kaufmann, D., Kraay, A., and Mastruzzi, M. (2011). The worldwide governance indicators: Methodology and analytical issues. *Hague J. Rule Law* 3, 220–246. doi:10.1017/S1876404511200046

Liu, C., and Wang, L. (2017). *Textual and quantitative research on China’s action plan for promoting the development of big data from the perspective of policy tools[J]*. Canada: Canadian Social Science.5.

Markowitz, H., Todd, P., Xu, G., and Yamane, Y. (1993). Computation of mean-semivariance efficient sets by the critical line algorithm. *Ann. Operations Res.* 45 (1), 307–317. doi:10.1007/bf02282055

Mdee, A., and Mushi, A. (2021). Untangling blame and responsibility for service delivery and local governance performance: testing a grounded social accountability approach in Tanzania[J]. *Local Government Studies* 47 (6), 993–1013.

MIIT (2017). Notice of the Ministry of industry and information technology on printing and distributing the big data industry development plan (2016-2020). [EB/OL]. (In Chinese). Available at: http://www.miit.gov.cn/n1146295/n1652858/n1652930/n3757016/c5464999/content.html (Accessed 01 17, 2017).

Ofpresident, E. O. (1993). “National performance review,” in *From red tape to results: Creating a government that works better and costs less.*

Pannu, M., Gill, B., Tebb, W., and Yang, K. (2016). “The impact of big data on government processes,” in Information Technology, Electronics & Mobile Communication Conference, Vancouver, BC, 13-15 October 2016 (IEEE).

Parnell, G. S. (2007). *Value-focused thinking*. San Francisco: Academia.

Polimeni, J., Mayumi, K., Giampietro, M., and Alcott, B. (2008). *Jevons’ paradox: The myth of resource efficiency improvements earthscan*. London: Taylor and Francis Group. Research Edition[J].

Ryzin, V. G. G., Muzzio, D., and Immerwahr, S. (2004). Explaining the race gap in satisfaction with urban services. *Urban Aff. Rev.* 39 (5), 613–632. doi:10.1177/1078087404264218

Spanhove, J., and Verhoest, K. (2007). *Corporate governance vs. government governance: Translation or adaptation? [C]*. Brussels: EIASM workshop on corporate governance. Date: 2007/11/15 - 2007/11/16.

Sullivan, R., Gouldson, A., and Webber, P. (2013). Funding low carbon cities: Local perspectives on opportunities and risks. *Clim. Policy* 13 (4), 514–529. doi:10.1080/14693062.2012.745113

Sun, G. A. (2020). "Government governance of smart cities in China[C]," in 2019 International Conference on Management Science and Industrial Economy (MSIE 2019), January 2020.

The Audit Commission, UK (2019). CPA-the harder test 2006[R/OL]. Available at: https://webarchive.nationalarchives.gov.uk/20150423154445/html (Accessed 07 06, 2019).

The Networking And Information Technology Research And Development Program, Group B (2016). *The federal big data research and development strategic plan*. United States: The Networking And Information Technology Research And Development Program.

Wang, Y. (2020). "Adaptability of professional setting and local industrial structure of applied undergraduate colleges in dongguan under the background of big data[M]," in *Advances in intelligent systems and computing* (Berlin, Germany: Springer).

Zhang, J., Luna-Reyes, L. F., and Pardo, T. A. (2016). "Information, policy, and sustainability: The role of information technology in the age of big data and open government[J]," in *Public administration and information technology* (New York: Springer International Publishing), 1–19. 10.1007/978-3-319-25439-5(Chapter 1).

Zhengce (2015). Notice of the State Council on printing and distributing the action outline for promoting the development of big data [EB/OL]. (In Chinese). Available at: http://www.gov.cn/zhengce/content/2015-09/05/content_10137.html (Accessed 09 05, 2015).