Available online www.jsaer.com

***Journal of Scientific and Engineering Research*, 2024, 11(2):151-159**

Research Article

ISSN: 2394-2630
CODEN(USA): JSERBR

Forecast and analysis of the number of fatalities in major coal mine accidents

Guoxun Jing[1,2] , Xiaoyu Ma[1*] , Fang Jiang[1]

[1]School of Safety Science and Engineering, Henan Polytechnic University, Jiaozuo Henan 454002, China;
[2]Anyang Institute of Technology, Anyang Henan 454002, China
Email: 1023996147@qq.com

Abstract In order to improve the prediction accuracy of coal mine accidents and master the law of coal mine accidents, the number of deaths in coal mine major accidents nationwide from 2012 to 2022 is taken as the sample data, and the gray correlation analysis method is used to study the influence of the number of deaths in coal mines on coal mine safety. The results show that the largest correlation is found between three violations and the number of deaths in coal mine accidents, followed by chaotic safety management, an incomplete supervision system, and inadequate education and training.Making predictions about the number of fatalities from major coal mine accidents between 2023 and 2025 using both the improved and conventional gray models. The improved grey model has higher prediction accuracy and reduced relative errors, according to the results.

Keywords accident statistics; gray theoretical models; variable weight buffer operator; gray correlation analysis

1. Introduction

The suddenness, catastrophic nature, and destructiveness of coal mine accidents pose a major threat to the lives and safety of those working underground. A crucial component of coal mine safety management is analyzing and projecting the future safety situation of the mines. Projecting and analyzing the number of fatalities in coal mine accidents allows for a complete assessment of the disaster's impact, allowing for preparation for the prevention of coal mine accidents[1].

Many scholars at home and abroad have carried out research related to coal mine accident prediction.In 1982, Professor Deng Julong[2] founded the gray system theory.In 2003, Yang Zhong[3] et al. firstly applied the gray prediction to the analysis of coal mine safety accidents, which provided theoretical references to the quantitative research of coal mine safety accidents.In 2019, Yousuo Joseph Zou[4] et al. established a linear regression model to make a fatal number prediction, but linear regression can only deal with the problem of linear relationship and is sensitive to outliers.In 2020, Lan Zequan[5] et al. established a gray system theoretical system to predict the number of special major accidents in the national coal industry, but the gray prediction is only suitable for short- and medium-term prediction. Huang Jiguang[6] et al. predicted China's million-ton mortality rate from 2009 to 2018 through Brown's linear trend model. Bai Yanlong[7] et al. introduced the idea of BP neural network (GA-BP) and wavelet neural network based on the improvement of genetic algorithm to predict the number of coal mine accidents in 2019 and 2020.2021 Zhang Peisen[8] et al. set up a quadratic exponential smoothing system to predict and analyze the million-ton fatality rate of coal mines in 2022 and 2023, which has a certain degree of accuracy, but the method lacks the data turning points,However, the method lacks the ability to identify the turning points of the data. Due to the shortcomings of the aforementioned prediction methods and the frequency of accidents, it is evident that the system in question is a typical gray system. However, during the medium- and long-term prediction process, the gray system is susceptible to shock perturbations, which causes the prediction results to diverge from the actual development trend. Accordingly, this paper uses gray correlation to examine the degree of correlation of the influencing factors based on the available data and creates a variable-weighted buffer-based The variable weight buffer operator-based gray model was developed to forecast the amount of fatalities in major coal mine accidents.

2. Coal mine major accident fatality prediction model

According to the State Mine Safety Supervision Bureau and the Coal Mine Safety Production Network, Figure 1 shows the statistics of the number of major accidents and deaths in coal mines in China from 2012 to 2022. Major coal mine accidents in China from 2012 to 2022 resulted in 91 instances and 1,668 deaths, with an overall declining trend. At 21.05% of all incidents, 2013 had the largest number of major accidents. In 2022, there were 3 big accidents that resulted in 38 deaths.With localized rises in 2015 and 2021, the number of fatalities in significant coal mining accidents decreased gradually from 254 in 2013 to 38 in 2022.

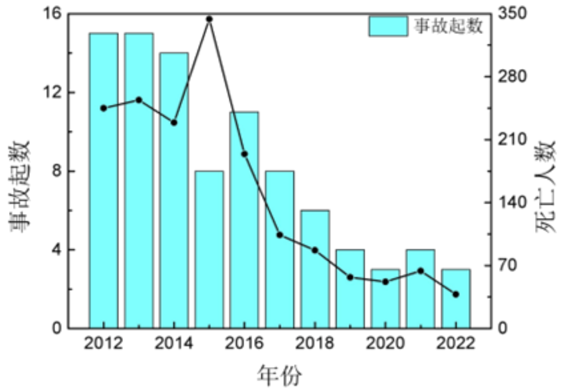

Figure 1: Statistics on the number of major accidents and fatalities in coal mines, 2012-2022

Table 1: Number of coal mine accidents and deaths in China, 2012-2022

vintages	Number of general accidents	Number of fatalities in general accidents	Number of larger accidents	Number of fatalities in larger accidents	Number of major accidents	Deaths in major accidents
2012	36	24	13	60	15	245
2013	32	22	15	71	15	254
2014	6	2	20	92	14	229
2015	22	17	8	37	8	334
2016	32	28	8	39	11	194
2017	62	59	12	50	8	104
2018	117	107	22	95	6	87
2019	76	75	26	130	4	57
2020	70	75	17	79	3	52
2021	39	42	14	67	4	64
2022	28	30	5	21	3	38

The figure shows that, of all the accidents that have occurred in the past ten years, general coal mine accidents have occurred the most (67% of total accidents), followed by larger incidents (21% of total accidents). From 2012 to 2022, the overall number of coal mine accidents decreased, increased, and then decreased again, as seen in the table. In 2018, there were 117 accidents and 107 fatalities, which was the greatest number to date.2019 was the highest number of major coal mine accidents, with 26 incidents and 130 fatalities. There were 91 major incidents overall, with 1,668 fatalities. Although the number of major accidents is low, the number of deaths is the highest and the harm caused is more severe. Therefore, it is important to attach great importance to the pattern of development of major accidents and to minimize the property damage and casualties caused when accidents occur.

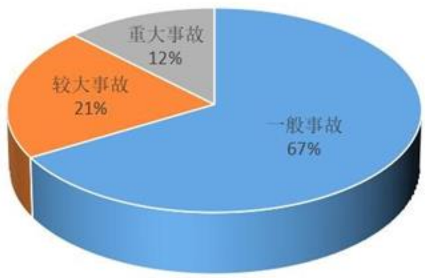

Figure.2 Pie chart of the number of major accidents in China's coal mines, 2012-2022

2.1 Gray modeling of variable-weight buffer operator

Gray correlation is in the process of system development, according to the degree of similarity or dissimilarity between the development trend of the factors to identify the degree of closeness of their relationship, so as to find out its main influencing factors and remove redundant factors. Gray system prediction is from the idea of gray system modeling, correlation and residual identification, through a series of data generation methods, a set of raw data sequences that are not at all regular, disorganized or less regular become obviously regular.However, studies have shown that the prediction of fast-growing time series using slow-growing time series tends to have a large deviation[10],in order to improve the accuracy of the model prediction, the historical slow-growing energy consumption data can be preprocessed using variable weight buffer operator, so as to make the model prediction results more close to the reality[11] .

2.2 Gray correlation analysis

1) Determination of reference and comparison sequences

   Reference sequences refer to sequences that reflect the behavioral characteristics of the research subject, and comparison sequences refer to sequences that influence the behavior of the research subject.

2) dimensionless processing

   The mean value method was used for dimensionlessness and the resolution was taken as 0.5.

3) Calculating Gray Correlation

   The gray correlation coefficient is the value of the degree of correlation between the reference sequence and the comparison sequence at each moment in time.

2.1.2 Definition of the buffer operator

Let the system behavior data sequence be

$X = (x(1), x(2), \cdots, x(n))$

1) If $\forall k \in 2,3,\cdots,n$, $x(k) - x(k-1) > 0$, then X is said to be a monotonically growing sequence.
2) If $\forall k \in 2,3,\cdots,n$, $x(k) - x(k-1) < 0$, then X is said to be a monotonically decreasing sequence.
3) If there exists $k1,k2 \in \{2,3,\cdots,n\}$,there is x(k1)-x(k1 -1)>0,x(k2)-x(k2 -1)<0, then X is said to be an oscillatory sequence, and also when

$M = \max\{x(k) | k = 1,2,\cdots,n\}$ $m = \min\{x(k) | k = 1,2,\cdots,n\}$

Call M-m the amplitude of the oscillatory sequence X.

Let X be a sequence of system behavior data, D be an operator acting on X. The resulting data sequence after the action of operator D is denoted as

$$XD = (x(1)d, x(2)d, \cdots, x(n)d) \quad (1)$$

Then D is called a sequence operator and XD is a sequence of first-order operator actions[12] . In practice, the first-order buffer operator tends to cause the effect of buffering action to be too strong or too weak, so the variable-weight buffer operator is introduced.

2.1.3 Gray model theory

Gray model, abbreviated as GM model, builds a gray differential prediction model to give a fuzzy long-term description of the law of things through a small amount of incomplete information[13] .

The core system of grey prediction is the grey model, that is, the original data to do cumulative generation, to generate an approximation of the exponential law and then modeling methods, the prediction process may involve three sequences, whitened differential equations, and a series of tests[14] .

2.2 Modeling Methodology

1) Constructing variable-weight buffer operators

Let X □ （$x(1),x(2),$□□□$,x(n)$) be a nonnegative monotonic growth sequence, a monotonic decay sequence, or an oscillatory sequence such that

□□□xXD(k)□d（□x(1)d,x((x2()kd)),□2□□,x(n)d) (2)

□□ □$x(n)$□(1□□)$x(k)$

where□is a variable weight and 0 □□□1，k □1,2,□□□,n ,then D is said to be a variable weight buffer operator. When□

= 0.2, 0.4, 0.6, 0.8, 1, different columns of data[15] are generated.

Let there be the original data column $x^{(0)}$ □ □$x_1^{(0)},x_2^{(0)}$ □□□$x_n^{(0)}$□ , n is the number of data[16] , and the sequence is the data column just constructed by the variable weight buffer operator.

A single accumulation of this data in order to weaken the volatility and randomness of the random sequence yields the new data sequence.

$x^{(1)}$ □ □$x_1^{(1)}$, $x_2^{(1)}$ □□□ $x_n^{(1)}$□ Among them. $x^{(1)}(t)$ □□$x^{(0)}(k)$,t □1,2,3,...,n

order linear differential equation for x(1) (t) for x(1) (t) ——— □ $ax(1)$ □ b (3)

Eq. (1) where a,b are called development coefficients and gray role quantities, respectively, according to the least squares method $[a,b]^T = (B^T B)\ B^{-1T}\ Y$ (4)

1□ □$x^{(0)}(2)$□

□ □ $^{(0)}(3)$□□ 1□ Y □ □x

□□0.（5 $x^{(1)}(n$ □1)□ $x^{(1)}(n)$) 1□□ □□$x^{(0)}(n)$□□

Bringing a,b into Eq. (3) and solving for it gives:

x□(1)(t □1)□□□x(0)(1)□ b□□e□at □ b (5)

The inverse generating equation can be obtained by discretizing and differencing the functional expression

□(0) □(1) □(1) x (t □1) □ x (t □1)□ x (t) (6)

(3) Posterior difference test for prediction models

Denote the 0th order residual as □$_i^{(0)}$ □ $x_i^{(0)}$ □ x_i (7)

where x is the predicted value obtained by the prediction model，i is the sequence number

mean value of residuals □ □ —□ in□1□i(0) (8)

Residual variance $s2$ □ 1 □ n (□(0) □□) (9)

Raw data means: x □ 1 □ n x(0) (10)

$s22$ □ —1 □ in□1(xi(0) □ x)2 n (11)

c □ s_1 (12)

small error probability P: P □ P□|□ $_1^{(0)}$ □□ $^{(0)}$ □ 0.6745S_2 |□ (13)

Based on the above indicators, the test accuracy level can be checked from the gray model accuracy test level table (Table 2)[17] .

Table 2: Gray model accuracy test level table

Prediction accuracy class	P	c
(of an unmarried couple) be close	P>0.95	c<0.35
eligible (voter etc)	0.95 □P□ 0.8	0.35 □c□ 0.5
barely enough	0.8 □P□ 0.7	0.5 □c□ 0.65
substandard	P□ 0.7	c□ 0.65

3. Gray correlation analysis

According to the gray system theory, the number of deaths in major accidents in coal mines nationwide from 2012 to 2022 as a whole, each influencing factor as a research object, combined with the data from the State Mine Safety Supervision Bureau and the Coal Mine Safety Production Network, respectively, selected the three violations of the law, poor awareness of safety risks, lack of education and training, disorganized safety management, inadequate supervision system, insufficient staffing, and lack of hidden danger investigation and management, The eight indicators of illegal organization of production are taken as influencing factors. The gray correlation was calculated and ranked according to the above steps, and the results are shown in Figure 3 and Table 3.

According to the principle of gray correlation analysis, the value of correlation ranges from 0 to 1. The larger the value, the stronger the correlation between the factor and the "reference value" (parent series); and the higher the correlation, the closer the relationship between the factor and the "reference value" (parent series). [18]The higher the correlation, the stronger the relationship between the factor and the reference value (parent series). As can be seen from the figure and the table, the correlation degree of each influencing factor, in descending order, is the following: three violations (illegal operation, illegal command and violation of labor discipline), chaotic safety management, inadequate supervision system, inadequate education and training, illegal organization of production, inadequate investigation and management of hidden dangers, inadequate staffing, and poor awareness of safety risks. The correlation values of all the influencing factors are higher than 0.6, indicating that the selected factors are closely related to the number of deaths in coal mine accidents.

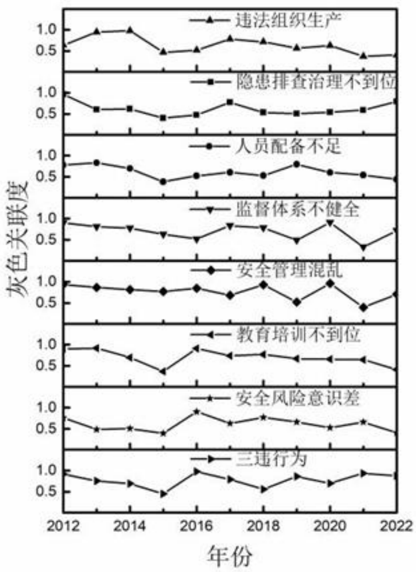

Figure.3 Schematic diagram of the correlation of the influencing factors

Table 3 Correlation between the number of fatalities in major coal mine accidents and the causes of each accident

numbe rof people kiled	Three violatio ns(of the law)	Poor awarene ssof security risks	Inadequ a te educatio n andtrain ing	Disor g anize d secur ity mana g emen t	Inadequ ate monitor i ng system	Inad eq uate staffi ng	Inadequa te identifica ti on and manage ment of hidden dangers	Illega l organiz ati on of produc tio n

2012	1.0000	0.9135	0.7599	0.9000	0.9344	0.9016	0.7784	0.9616	0.6389
2013	1.0000	0.7557	0.4864	0.9169	0.8671	0.8092	0.8322	0.6026	0.9433
2014	1.0000	0.6914	0.5074	0.6959	0.8137	0.7787	0.6984	0.6205	0.9762
2015	1.0000	0.4520	0.3937	0.3681	0.7672	0.6331	0.3860	0.4054	0.4690
2016	1.0000	0.9745	0.9062	0.9116	0.8468	0.5204	0.5256	0.4786	0.5228
2017	1.0000	0.7891	0.6295	0.7400	0.6814	0.8337	0.6059	0.7739	0.7808
2018	1.0000	0.5595	0.7691	0.7704	0.9276	0.7842	0.5301	0.5368	0.7155
2019	1.0000	0.8589	0.6614	0.6710	0.5225	0.4904	0.7944	0.5091	0.5664
2020	1.0000	0.6976	0.5297	0.6572	0.9607	0.9118	0.6038	0.5424	0.6296
2021	1.0000	0.9317	0.6604	0.6462	0.3989	0.3333	0.5439	0.5931	0.3765
2022	1.0000	0.8733	0.4075	0.4135	0.7040	0.7187	0.4398	0.7940	0.4088
relatedness	1.0000	0.7725	0.6101	0.6992	0.7658	0.7014	0.6126	0.6198	0.6389

4. Gray forecasting

The number of fatalities in major and above accidents in coal mines nationwide from 2012 to 2022, and the number of fatalities from 2023 to 2025 are predicted. The raw data columns were processed according to the □ values taken and the results are shown in the table below.

Table 4: Model prediction results for different values of variable weights

particular year	number of people killed	□=0.2	□=0.4	□=0.6	□= 0.8
2012	245	204	163	121	80
2013	254	211	168	125	82
2014	229	191	153	115	77
2015	344	283	222	161	100
2016	194	163	132	101	70
2017	104	91	78	65	52
2018	87	78	68	58	48
2019	57	54	50	46	42
2020	52	50	47	44	41
2021	64	59	54	49	44
2022	38	38	38	38	38

The original gray GM (1, 1) model, i.e., when□= 0, is established as a gray GM (1, 1) model for predicting the number of deaths in heavy accidents in coal mines, which is obtained by substituting the original data x_0 = [245, 254, 229, 344, 194, 104, 87, 57, 52, 64, 38] into Eq. (3) and Eq. (4):

□ □372 1□ □254□
□□ 613.5 1□□ □□229□□
□ 1□ □344□
□ □900 □ □ □

□ □1169 1□ □194□ 1□
□ □104□
□ □1318 □ Y □ □ □
B □ □ 1□ □ 87 □
□ □1413 1□□ □□ 57 □□
□ □1485 1□ □ 52 □
□ 1□□ □□ 64 □□
□ □1540 1□□ □□ 38 □□
□ □1598
□
□□
□1649

obtained using the least squares method:	□(1) x (t □1) □ 388.6520exp(□0.2043t)
Similarly, when□ = 0.2	□(1) x (t □1) □ 316.5593exp(□0.1924t)
When□ = 0.4	□(1) x (t □1) □ 245.0722exp(□0.1762t)
When□ = 0.6	□(1) x (t □1) □173.7632exp(□0.1504t)
When□ = 0.8	□(1) x (t □1) □103.8251exp(□0.1039t)

The predictions are different for different values and to get the best□value, genetic algorithm search can be used.

Construct the fitness function:

1 □11 ^ (0) $^{(0)}(i)$ (14)

where x (i) and $x^{(0)}(i)$ are the ith G(1, 1) model fit based on the variable-weighted reinforcement buffer operator and the modeled data values after the variable-weighted reinforcement operator, respectively.

Using matlab to write the calculation program of the fitness function, after 100 iterations, the output of the optimal□* = 0.073 (so that the fitness function value is minimized), at this moment the gray model is:

x (t □1) □ 379.9966exp(□0.2032t)

The fit is shown in Figure 4.

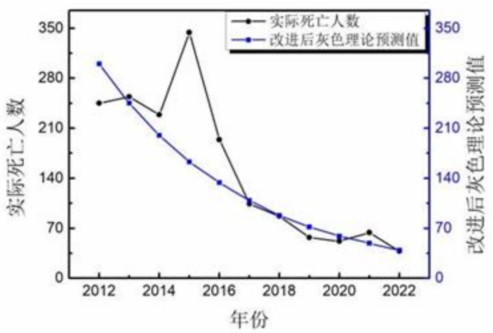

Figure.4 Model prediction results

Based on Eq. (7)-Eq. (13) the average relative error, the posterior residual ratio, and the probability of a small error for this model can be derived as shown in Table 5.

Table 5 Model Prediction Accuracy

mould	Average relative error	Posterior residual c	The small error probability P
gray model	0.173	0.17	1
Improved gray model	0.1628	0.113	1

As can be seen from Table 5, both the gray model and the improved model have a prediction accuracy rating of "good", and the residuals after the improved model are smaller, which results in higher accuracy, and the prediction of the number of fatalities in major accidents in the next five years is shown in Table 6.

Table 6: Forecast of fatalities in major coal mine accidents, 2022-2026

mould	2022	2023	2024	2025	2026
gray theory	42	35	29	24	20
Improved Gray Theory	39	32	26	21	17

It is known that the number of deaths in coal mine major accidents in 2022 is 38, as can be seen from the above table, the predicted value of the gray theoretical model is 42, with a relative error of 9.5%, while the predicted value of the model improved by variable-weighted buffer arithmetic is 39, with a relative error of only 2.5%, so it can be predicted by the gray model improved by buffer arithmetic.

5. Conclusion

1) This study used gray correlation analysis to ascertain the gray correlation degree of each index with the goal of identifying the contributing causes to significant accidents in coal mines. The analysis's findings indicate that the three violations have the highest correlation with the number of coal mine accident fatalities. These three violations are followed by ineffective supervision systems, chaotic safety management, and inadequate education and training. Illegal production organization, insufficient staffing, inadequate management of hidden dangers, and a lack of awareness of safety risks are the next most common violations.In order to avoid this, it is first essential to reinforce the safety education and training of the coal mine operators, increase the operators' safety awareness,reducing "three violations" during coal mining operations; Second, the department in charge of safety oversight shall properly carry out its duties and forbid the unlawful arrangement of production; In order to help lower casualties and accident losses, coal mines must also boost the implementation of hidden hazard inspection and rectification, improve their safety management institutions, and ensure that all management people are fully equipped in accordance with rules.
2) The number of deaths in major coal mine accidents in China from 2023 to 2025 is predicted using the gray theoretical model of variable weight buffer operator in this paper. The predicted and actual values are compared, and the results indicate that the gray model improved by variable weight buffer operator (optimized by genetic algorithm) has higher prediction accuracy and smaller relative error, which can serve as a basis for the relevant departments to formulate corresponding measures. However, the model cannot take into account other influencing factors and has certain defects.

References

[1] Jing Li, Tianbao Song, Yaran Wang. Prediction of coal mine accident fatalities based on improved unbiased gray Markov model[J]. Coal Technology,2017,36(08):318-320.DOI:10.13301/j.cnki.ct.2017.08.127.
[2] Sifeng Liu,Yi Lin. On measures of information content of gray numbers[J]. Kybernetes,2006,35(6).
[3] Zhong Yang, Yulan Ding, Chaoyi Zhao. Gray correlation analysis and trend prediction of safety accidents in Kailuan coal mine[J]. Coal Journal,2003(01):59-63.
[4] Yousuo Joseph Zou, Jun Steed Huang, Tong Xu, Anne Zou, Bruce He, Xinyi Tao, Sam Zhang. Prediction of Coal Mine Accidental Deaths for 5 Years Based on 14 Years Data Analysis[C]//.Proceedings of the 9th International Conference on Computer Engineering and Networks(CENet2019). ,2019:250-257.DOI:10.26914/c.cnkihy.2019.048094.

[5] Zequan Lan, Benfu Fu, Dongmei Tian, Fumei Song. Gray prediction and analysis of special major accidents in coal mines[J]. Coal Technology,2020,39(07):81-83.DOI:10.13301/j.cnki.ct.2020.07.023.

[6] Jiguang Huang, Hanpeng Ma, Chunjiao Fan, Rongrong Yao. Statistical analysis and prediction of coal mine safety accidents in China[J]. Shaanxi Coal,2020,39(03):34-39+6.

[7] Yanlong Bai, Yu Chen, Changjiang Bai, Jianming Liang, Long Xi, Yubi Xue, Jiajia Peng, Xin Wang. Research on predicting the number of coal mine accidents based on neural network[J]. Coal and Chemical

[8] Industry,2020,43(09):91-94+97.DOI:10.19286/j.cnki.cci.2020.09.026.

[9] Peisen Zhang, Xiaole Zhang, Yuhang Dong, Daqiang Xu.Research on the analysis and prediction of the accident pattern of China's coal mines from 2008 to 2021[J]. Mining Safety and Environmental

[10] Protection,2023,50(02):136-140+146.DOI:10.19835/j.issn.1008-4495.2023.02.024.

[11] Chunxia Zhao. Statistical analysis and prediction of production safety accidents in coal mines in China[D]. Southwest University of Science and Technology,2017.

[12] Jiangrong Wang, Shuo Liu, Cuncheng Jin. Application of gray G(1,1) model based on variable weight buffer operator in metro energy consumption prediction[J]. Practice and Understanding of Mathematics,2020,50(07):90-96.

[13] Minjie Yang, Ji Wang, Yutong Pu, Wenjue Su, Mengqin Li. Medium- and long-term railroad passenger traffic forecasting based on variable weight buffer multivariate gray model[C]//.Proceedings of the World Conference on Transportation 2022 (WTC2022) (Transportation Planning and Interdisciplinary

[14] Part). ,2022:614-619.DOI:10.26914/c.cnkihy.2022.019326.

[15] Aiping Jiang. Construction and application of a class of weakened buffer operators[J]. Journal of Langfang Normal College(Natural Science Edition),2014,14(06):17-19.

[16] Jing Li, Ruixin Zhang, Zequan . Application of improved gray model in the prediction of safety hazards in surface coal mines[J]. Journal of North China Institute of Science and Technology,2019,16(03):67-72.

[17] Liang Wang. Research on the prediction and application of coal mine million-ton mortality rate based on gray system theory[J]. Mining Technology,2018,18(06):99-101.DOI:10.13828/j.cnki.ckjs.2018.06.032.

[18] Jiangrong Wang, Shuo Liu, Cuncheng Jin. Application of gray G(1,1) model based on variable weight buffer operator in metro energy consumption prediction[J]. Practice and Understanding of Mathematics,2020,50(07):90-96.

[19] Fei Tang, Yungang Wang, Bingcheng Du, Xiangyan Kong. Prediction of coal mine accidental deaths based on optimized Markov model [J]. Chinese Journal of Safety Science,2022,32(04):122-128.DOI:10.16265/j.cnki.issn1003-3033.2022.04.018.

[20] Kai Wang, Jiahui Liang. Research on the prediction of coal mine gas accidents in China based on gray system theory[J]. Modern Mining,2016,32(11):188-190.

[21] Li Li, Ming Yu. Research on China's export trade of electromechanical products to RCEP member countries - Based on gray correlation analysis and GM(1,1) model [J]. China Business Journal, 2023, (24): 85-89. DOI:10.19699/j.cnki.issn2096-0298.2023.24.085