ISSN 2616-5775 Vol. 5, Issue 9: 45-51, DOI: 10.25236/AJCIS.2022.050908

Fire Rescue Model Based on Time Series Analysis

Hu Han

School of Statistics and Applied Mathematics, Anhui University of Finance and Economics, Bengbu, Anhui, China

Abstract: *In this paper, we establish relevant models for the problems in fire rescue and solve them. Using the ARIMA model to establish a prediction model for the number of fire rescue calls, and predict the number of fire rescue calls in each month in 2021. According to the characteristics of fire incidents, we divide fire incidents into 7 levels, and use the method of cluster analysis to analyze the spatial correlation of various incident densities in this area, and give the most relevant incident categories in different regions. Grey correlation analysis was used to analyze the relationship between the density of various events and population density in the area. And on this basis, based on the minimum distance between the new fire station and other areas, the Floyd algorithm is used to solve the shortest distance between each area and other areas, and the location where the new fire station should be built is analyzed.*

Keywords: *Time series analysis, K-means cluster analysis, Grey relational analysis, Floyd algorithm*

1. Introduction

With the rapid development of national economy, cities have also been rapidly expanded along with the rapid development of China's economy, which also makes various accidents and disasters frequent, security risks are constantly increasing, and the tasks undertaken by fire rescue teams also present a trend of diversification and complexity. Dynamically adjusting the number of people on duty in the fire department according to the time of disaster occurrence, and analyzing the density, time, and population of various fire incidents are of great significance in effectively arranging fire forces, predicting potential accidents, and planning fire stations. The fire rescue team needs to speed up the dispatch speed, effectively save the rescue time, significantly improve the risk analysis and safety management and control, and protect people's life and property safety to the greatest extent [1]. The author proposes to use mathematical modeling, data fusion and other aspects to analyze and construct a fire rescue situational awareness system [2]. The author points out the main problems existing in the construction of the government's full-time fire brigade from the aspects of construction scale, construction quality, security level, and management mechanism, and proposes countermeasures and suggestions for strengthening the construction of the government's full-time fire brigade [3]. Based on the above literature, previous scholars have theoretically analyzed the fire rescue model in theory, but lack of empirical research, so it is particularly important to carry out relevant empirical research.

In this paper, we establish a prediction model for the number of fire rescue calls, and predict the number of fire rescue calls in each month of 2021. Afterwards, we classified fire incidents into 7 categories. Using the method of cluster analysis, we clustered different areas of the area according to various event densities, analyzed the spatial correlation of various incident densities in the area, and gave the different most regionally relevant event category. Grey correlation analysis was used to analyze the relationship between the density of various events and population density in the area. Based on the Floyd algorithm, the shortest distance between each area and other areas is calculated, and the location where the new fire station should be built is analyzed.

2. ARIMA model

According to the data, stata software can be used to draw a scatter plot as shown in Fig. 1.

Academic Journal of Computing & Information Science
ISSN 2616-5775 Vol. 5, Issue 9: 45-51, DOI: 10.25236/AJCIS.2022.050908

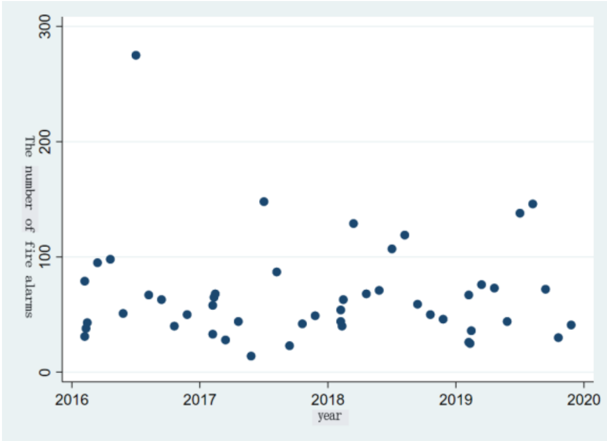

Figure 1: Scatter plot of the number of fire alarms in each month from 2016 to 2019

It can be seen from Fig. 1 that the data fluctuates around a certain mean value, and is further drawn by SPSS software as shown in Fig.2, and the graph after the first order difference is shown in Fig. 3.

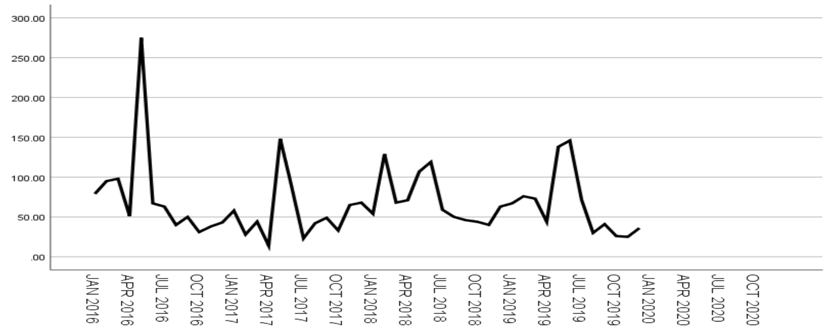

Figure 2: The line chart of the number of fire alarms in each month from 2016 to 2019

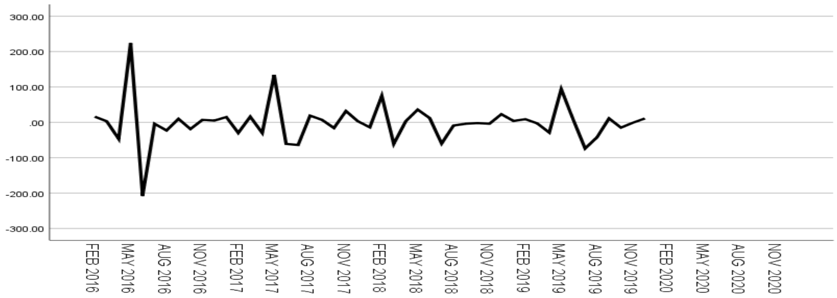

Figure 3: First-order difference diagram of the number of fire alarms in each month from 2016 to 2019

It can be clearly seen from Figure 3 that the curve roughly fluctuates around 0, which is very similar to the time series model. The time series model can be used to fit it and use it for prediction. A numerical sequence in which the index values of a phenomenon are arranged in chronological order. Time series analysis can be roughly divided into three parts, which are describing the past, analyzing laws and predicting the future. Two models are commonly used in time series analysis: exponential smoothing method and ARIMA model. The basic formula of exponential smoothing method is:

$$S_t = a \cdot y_t + (1 - a)S_{t-1} \quad (1)$$

where S_t is the predicted value or exponential smoothing value of the t period, y_t is the actual value of the t period, S_{t-1} is the predicted value of the t-1 period, a is the smoothing index, and its value range is [0,1]. To put it simply, the exponential smoothing value of any period is the weighted average of the actual observed value of the current period and the exponential smoothing value of the previous period, which can also be understood as the predicted value of the data of the next period is related to the actual value of the current period and the predicted value of the previous period.

The general form of the ARIMA model is the ARIMA (p, d, q) form, and the basic form is [4]:

ISSN 2616-5775 Vol. 5, Issue 9: 45-51, DOI: 10.25236/AJCIS.2022.050908

$$\begin{cases} y_t{}' = a_0 + \sum_{i=1}^{p} a_i y_{t-i}{}' + \varepsilon_t + \sum_{i=1}^{q} \beta_i \varepsilon_{t-i} \\ (1 - \sum_{i=1}^{p} a_i L^i)(1-L)^d y_t = a_0 + (1 + \sum_{i=1}^{q} \beta_i L^i)\varepsilon_t \end{cases} \tag{2}$$

where $y_t{}'$ represents a time series, $(1-L)^d$ represents the d-order difference, ε_t represents the random interference item, $(1 - \sum_{i=1}^{p} a_i L^i)$ represents the AR(p) model, $(1 + \sum_{i=1}^{q} \beta_i L^i)\varepsilon_t$ represents the MA(q) model.

Here, SPSS software is used to fit the above models with exponential smoothness model and ARIMA model respectively. In the output results, we see that the fitting effect of the ARIMA (0,1,0) model is better, and the results shown in the software as shown in table 1.

Table 1: Model statistics

Model statistics						
Model	number of predictors	Model fit statistics	Ljung-Box test (18)			Number of outliers
		Stationary R-square	Statistics	DF	salience	
Number of alarms - model	0	0.651	13.809	18	0.741	2

The goodness of fit of the model is 0.651, indicating that the model fits well, and the p value obtained by the Q test on the residual is 0.741, that is, we cannot reject the null hypothesis and consider the residual to be a white noise sequence, which is shown in Fig.4.

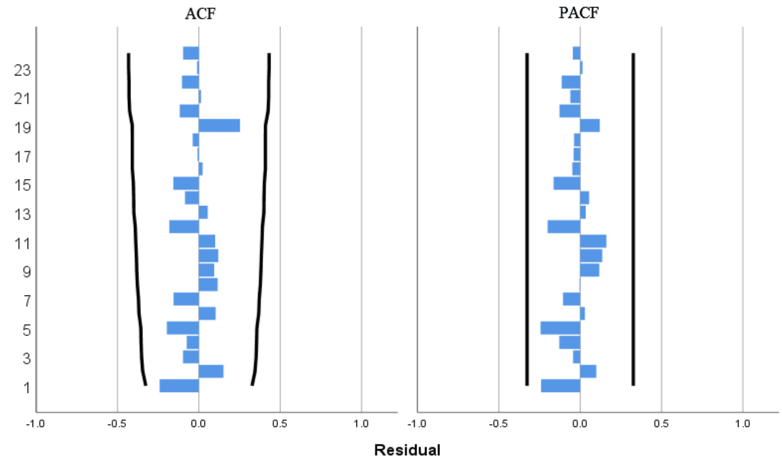

Figure 4: Model residual map

In addition, from the data obtained in the software, it can be seen from the ACF and PACF graphs of the residuals that the autocorrelation coefficients and partial autocorrelation coefficients of all lag orders are not significantly different from 0 , so the model can well identify the alarm data in this example. Combined with the requirements of the topic, the data from January 1, 2020 to December 31, 2020 is used as the validation dataset of the model, and the validation results are shown in Figure 5.

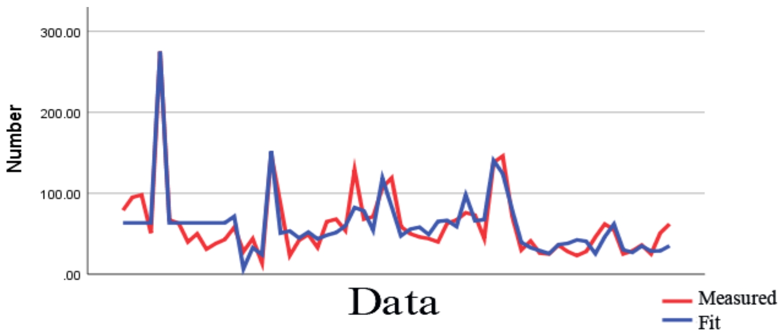

Figure 5: Model fitting curve

In Fig.5, the red line represents the measured line, and the blue line represents the fitted line. It is not difficult to see that the trend of the measured curve from January 1, 2020 to December 31, 2020 is very close to the fitted curve. Therefore, the time series model can be used to predict the number of fire rescue calls in each month in 2021. The predicted results are shown in table 2.

ISSN 2616-5775 Vol. 5, Issue 9: 45-51, DOI: 10.25236/AJCIS.2022.050908

Table 2: Predicted value of the number of fire rescue calls for each month in 2021

Time	Number of alarms
January 2020	31
February 2020	29
March 2020	31
April 2020	41
May 2020	41
June 2020	46
July 2020	30
August 2020	32
September 2020	44
October 2020	30
November 2020	44
December 2020	50

3. K-means clustering algorithm

The K-means clustering algorithm is a simple iterative clustering algorithm that uses distance as a similarity measure to divide a given data set into k classes. The specific process is as follows:

(1) Randomly select k samples as the initial clustering center;

(2) Calculating the distance between each sample and the cluster center;

(3) Regressing each sample to the nearest cluster center;

(4) Calculating the mean of each class as the new cluster center;

(5) If the cluster center remains unchanged or the number of iterations is reached, the algorithm ends, otherwise go back to step (2) [5].

Based on this principle, we can first obtain the following area and event density table by processing the data, as shown in table 3.

Table 3: Correspondence table of area and event density

Area/Event Density	①	②	③	④	⑤	⑥	⑦
A	0.21	0.07	0.24	0.21	0.01	0.03	0.97
B	0.16	0.03	0.13	0.06	0.00	0.02	0.58
C	0.14	0.01	0.16	0.10	0.02	0.00	0.61
D	0.25	0.12	0.33	0.17	0.03	0.03	1.11
E	0.23	0.04	0.57	0.08	0.02	0.03	1.25
F	0.31	0.01	0.29	0.11	0.02	0.01	1.05
G	0.18	0.02	0.29	0.17	0.03	0.02	1.04
H	0.16	0.06	0.12	0.12	0.02	0.00	0.65
I	0.05	0.01	0.08	0.05	0.00	0.01	0.39
J	0.27	0.18	0.38	0.16	0.01	0.11	0.99
K	0.15	0.02	0.24	0.08	0.00	0.02	0.86
L	0.04	0.02	0.18	0.11	0.02	0.05	0.38
M	0.13	0.05	0.42	0.10	0.03	0.05	1.04
N	0.15	0.06	0.34	0.07	0.00	0.01	0.49
P	17.68	5.05	23.74	12.93	1.92	28.59	80.91

It can be seen from table 3 that there are 15 regions and 7 event categories. The purpose of clustering is to cluster these 15 regions into 7 categories according to event categories. Here, K-means clustering analysis of SPSS software can be used to analyze each region. According to the event density, the event category with the strongest correlation in each region can be obtained, and the obtained results are shown in table 4.

ISSN 2616-5775 Vol. 5, Issue 9: 45-51, DOI: 10.25236/AJCIS.2022.050908

Table 4: Clustering result table

Cluster member		
Case number	Cluster	Distance
A	6	0.122
B	3	0.052
C	3	0.038
D	6	0.095
E	5	0.000
F	6	0.126
G	6	0.084
H	3	0.056
I	2	0.063
J	6	0.151
K	7	0.000
L	2	0.063
M	6	0.147
N	1	0.000
P	4	0.000

It can be seen from table 4 that region A has the strongest correlation with event (6), region B has the strongest correlation with event (3), region C has the strongest correlation with event (3), region D has the strongest correlation with event (6), and region E has the strongest correlation with event(5), region F has the strongest correlation with event (6), region G has the strongest correlation with event (6), region H has the strongest correlation with event (3), region I has the strongest correlation with event (2), and region J has the strongest correlation with event (6), region K has the strongest correlation with event (7), region L has the strongest correlation with event (2), region M has the strongest correlation with event (6), region N has the strongest correlation with event (1), and region P has the strongest correlation with event (4).

To analyze the relationship between various event densities and population densities in this area, due to the small amount of data, grey correlation analysis method can be used to analyze the relationship between various event densities and population densities. First, the population density of each region is calculated, and the results are shown in table 5.

Table 5: Relationship between event density and population density in each region

Area/Event Density	(1)	(2)	(3)	(4)	(5)	(6)	(7)	Population density
A	0.21	0.07	0.24	0.21	0.01	0.03	0.97	0.07
B	0.16	0.03	0.13	0.06	0	0.02	0.58	0.07
C	0.14	0.01	0.16	0.1	0.02	0	0.61	0.06
D	0.25	0.12	0.33	0.17	0.03	0.03	1.11	0.08
E	0.23	0.04	0.57	0.08	0.02	0.03	1.25	0.08
F	0.31	0.01	0.29	0.11	0.02	0.01	1.05	0.08
G	0.18	0.02	0.29	0.17	0.03	0.02	1.04	0.07
H	0.16	0.06	0.12	0.12	0.02	0	0.65	0.07
I	0.05	0.01	0.08	0.05	0	0.01	0.39	0.06
J	0.27	0.18	0.38	0.16	0.01	0.11	0.99	0.07
K	0.15	0.02	0.24	0.08	0	0.02	0.86	0.07
L	0.04	0.02	0.18	0.11	0.02	0.05	0.38	0.05
M	0.13	0.05	0.42	0.1	0.03	0.05	1.04	0.08
N	0.15	0.06	0.34	0.07	0	0.01	0.49	0.08
P	17.68	5.05	23.74	12.93	1.92	28.59	80.91	1.61

4. Grey relational analysis

Grey relational analysis is to study the degree of closeness between the selected object and the recognized object, quantify it, calculate the degree of relatedness of its closeness, and evaluate the degree of influence of the object to be identified by comparing the degree of relatedness [6]. The first step of grey relational analysis is to determine the analysis sequence, including subsequence and parent sequence.

ISSN 2616-5775 Vol. 5, Issue 9: 45-51, DOI: 10.25236/AJCIS.2022.050908

The subsequence is a data sequence composed of factors that affect the behavior of the system, and the parent sequence is a data sequence that can reflect the behavior of the system, which can be expressed as:

$$\mathcal{X}(\mathrm{m}) = (\mathcal{X}_{\mathrm{m}}(1), \mathcal{X}_{\mathrm{m}}(2) \ldots)' \tag{3}$$

$$\mathcal{X}(0) = (\mathcal{X}_0(1), \mathcal{X}_0(2) \ldots \ldots)' \tag{4}$$

Calculating the absolute value of the difference between each subsequence and the parent sequence, and find the minimum and maximum values from them, which can be expressed as:

$$\mathrm{a} = \min(\min|\mathcal{X}_0(\mathrm{k}) - \mathcal{X}_{\mathrm{m}}(\mathrm{k})|) \tag{5}$$

$$\mathrm{b} = \max(\max|\mathcal{X}_0(\mathrm{k}) - \mathcal{X}_{\mathrm{m}}(\mathrm{k})|) \tag{6}$$

Calculating the correlation degree, which can be expressed as:

$$\mathrm{y}(\mathcal{X}_0(\mathrm{k}), \mathcal{X}_{\mathrm{m}}(\mathrm{k})) = \frac{\mathrm{a}+\rho\mathrm{b}}{|\mathrm{x}_0(\mathrm{k})-\mathcal{X}_{\mathrm{m}}(\mathrm{k})|+\rho\mathrm{b}} \tag{7}$$

where ρ is called the resolution coefficient, and the general value is 0.5. The larger the value of $\mathrm{y}(\mathcal{X}_0(\mathrm{k}), \mathcal{X}_{\mathrm{m}}(\mathrm{k}))$, the higher the degree of correlation [7-8].

The results obtained by solving the problem with Matlab are shown in table 6.

Table 6: Grey correlation degree between events and population density

The grey correlation degree of each indicator in the subsequence						
0.7165	0.7963	0.8043	0.6613	0.6901	0.8542	0.8287

It can be seen from table 6 that event ⑦ and event ③ and ⑥ have the greatest correlation with population density, followed by event ② and event ①, and event ④ and event ⑤ have the smallest correlation with population density. Therefore, it can be considered that the degree of correlation between events ⑦, event ③ and event ⑥ in this area is more significant with population density, followed by event ② and event ①, and event ④ and event ⑤ are less correlated with population density.

For the site selection of a new fire station, the distance between the area and other areas should be kept to a minimum. Therefore, the Floyd algorithm can be used to find the shortest distance between each area and other areas, and then analyze the location of the new fire station. The relevant regions are first analyzed, as shown in Fig. 6.

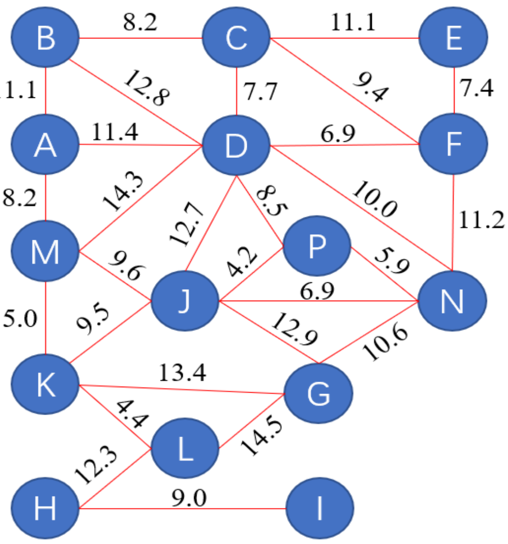

Figure 6: The relationship between regions

We can use the Floyd algorithm to solve the distance between each area and other areas [9]. After solving through Matlab, we can get the distance matrix between each point, as shown in table 7.

ISSN 2616-5775 Vol. 5, Issue 9: 45-51, DOI: 10.25236/AJCIS.2022.050908

Table 7: The shortest distance in each area

Distance	A	B	C	D	E	F	G	H	I	J	K	L	M	N	P
A	0	11.1	19.1	11.4	25.7	18.3	30.7	39.9	48.9	17.8	23.2	27.6	8.2	21.4	19.9
B	11.1	0	8.2	12.8	19.3	17.6	33.4	51	60	25.5	34.3	38.7	19.3	22.8	21.3
C	19.1	8.2	0	7.7	11.1	9.4	28.3	46.6	55.6	20.4	29.9	34.3	22	17.7	16.2
D	11.4	12.8	7.7	0	14.3	6.9	20.6	38.9	47.9	12.7	22.2	26.6	14.3	10	8.5
E	25.7	19.3	11.1	14.3	0	7.4	29.2	51.7	60.7	25.5	35	39.4	28.6	18.6	22.8
F	18.3	17.6	9.4	6.9	7.4	0	21.8	44.3	53.3	18.1	27.6	32	21.2	11.2	15.4
G	30.7	33.4	28.3	20.6	29.2	21.8	0	26.8	35.8	12.9	13.4	14.5	22.5	10.6	16.5
H	39.9	51	46.6	38.9	51.7	44.3	26.8	0	9	26.2	16.7	12.3	31.7	33.1	30.4
I	48.9	60	55.6	47.9	60.7	53.3	35.8	9	0	35.2	25.7	21.3	40.7	42.1	39.4
J	17.8	25.5	20.4	12.7	25.5	18.1	12.9	26.2	35.2	0	9.5	13.9	9.6	6.9	4.2
K	23.2	34.3	29.9	22.2	35	27.6	13.4	16.7	25.7	9.5	0	4.4	15	16.4	13.7
L	27.6	38.7	34.3	26.6	39.4	32	14.5	12.3	21.3	13.9	4.4	0	19.4	20.8	18.1
M	8.2	19.3	22	14.3	28.6	21.2	22.5	31.7	40.7	9.6	15	19.4	0	16.5	13.8
N	21.4	22.8	17.7	10	18.6	11.2	10.6	33.1	42.1	6.9	16.4	20.8	16.5	0	5.9
P	19.9	21.3	16.2	8.5	22.8	15.4	16.5	30.4	39.4	4.2	13.7	18.1	13.8	5.9	0

It can be seen from table 7 that the distance between the newly built fire station and other points should be kept to the minimum. Here, the arithmetic mean of the distance between this point and other points can be used for comparison. Since J and N have established fire stations. Therefore, the J and N areas should be removed for the newly built fire station. The location of the new fire station obtained by calculation should be point P. In 2021-2029, a new fire station will be built every 3 years. The above idea can still be used, that is, the distance between this point and the rest of the points should be kept the smallest, so that in 2021 -The newly built fire stations in 2029 are points D, K and M in order.

5. Conclusion

This paper starts with actual data, establishes ARIMA model to fit and analyze the number of fire rescues, and predicts the number of fire alarms in the future. The correlation between the area and the population density is analyzed. Finally, based on the Floyd algorithm, taking the actual situation of a certain place as an example, the best fire station establishment location is analyzed. Feasibility analysis has been carried out on the rational allocation of firefighting resources, which has certain reference and practical value for other regions.

References

[1] Wei Lijun. Analysis on the improvement of firefighting and rescue team's ability to improve firefighting and rescue in the first battle [J]. Fire Industry (Electronic Edition), 2022, 8(06): 43-45.

[2] Qu Zuodong, Shi Chengtao, Li Youdong, Zhang Wei. Research on the construction method of fire rescue situational awareness system [J]. Fire Science and Technology, 2022, 41(04): 559-561+565.

[3] Zhang Quanling, Wang Dong. Analysis and countermeasures for the construction and development of the government's full-time fire brigade [J]. Fire Science and Technology, 2022, 41(04): 551-554.

[4] Ding Haifeng, Li Liqing. Research on the time series prediction analysis of total health expenditure in the Yangtze river delta region based on ARIMA model [J]. China Medical Management Science, 2022, 12(02): 4-10.

[5] Zhao Wei, Ni Lulin, Li Feng. Research on ShadowscoksR traffic identification based on K-means cluster analysis [J]. Communication Technology, 2020, 53(08): 1970-1975.

[6] Wang Hanqi, Wu Weili, Gu Xuefeng, Chen Weilai. Grey relational analysis of principal components of sustainable cotton blended yarns [J]. Modern Textile Technology, 2021, 29(6): 55-61.

[7] Li Bingsong, Wang Junyong. The impact of intergenerational economic mutual support on the mental health of the elderly from the perspective of change: Based on grey relational analysis [J]. China's rural health service management, 2022, 42(4):286-290.

[8] Yan Huiguo, Dong Zhiling, Li Yu, Ma Tingting, Luo Hui, Ye Hua, He Wenping. Analysis of the relationship between morphological characteristics and body mass of Yanyuan carp based on path analysis and grey correlation [J]. Journal of Southwest University (Natural Science) Version). 2022, 44(5): 74-81.

[9] Qin Kepeng. Classical shortest path Algorithm and implementation [J]. China New Communications, 2019, 21(19): 137-139.