Supervision of Financial Market Sentiment and Quantitative Investment Analysis Based on BP Neural Networks

Fengyi Guan[1], Chuxin Deng[1], Tianci Chen[2], Dongwu Wu[1]
[1]*School of Economics and Management, Wuyi University, Jiangmen, Guangdong, China*
[2]*College of Liberal Arts, Wuyi University, Jiangmen, Guangdong, China*

Abstract: In the complex and volatile financial market, sentiment significantly impacts investment decisions. Despite advancements using BP neural networks for sentiment supervision and quantitative investment analysis, challenges like high complexity and overfitting persist. This study employs structural equation modeling to analyze factors influencing the "digital economy" sector, utilizes Prophet time series for volume prediction, and constructs a BP neural network for index forecasting. The model achieves an 85.79% and 92.74% prediction accuracy for peak and trough prices, respectively, and demonstrates robust risk management with a 13.26% total investment return and a 2.79% information ratio. Stability tests indicate reliable model predictions, offering valuable insights for stock fund selection.

Keywords: Structural Equation Modeling; BP Neural Network; Prophet Time Series; Risk Management Model; Quantitative Analysis

1. Research Overview

1.1 Research Background

The role of quantitative investment in the global financial trading market has become increasingly important with the development of big data technology. Investors explore market dynamics and use data analysis to predict market trends and make trading decisions to achieve stable returns. The significance of quantitative investment in the global financial trading market is growing with the continuous advancement of big data technology. To obtain reliable returns, investors typically use data analysis to understand market trends and predict market movements.

Big data technology profoundly influences the digital era of global financial trading markets, leading to a thriving landscape. As the financial market becomes increasingly complex, the challenges faced by investors also grow. Against this backdrop, quantitative investment, which leverages advanced data analysis and mathematical models for investment decisions, has gradually become a significant force in the financial trading market.

Investors are increasingly aware that relying solely on traditional investment experience and intuition is insufficient for achieving stable returns in the market. They need more scientific and precise methods to grasp market dynamics and predict market trends. By analyzing large volumes of data, investors can gain deeper insights into market dynamics, identify potential investment opportunities, and formulate corresponding trading strategies. Therefore, the development of big data technology provides robust technical support and data foundation for quantitative investment, highlighting its growing importance in the global financial trading market.

1.2 Research Objectives

The financial sector is gradually shifting from traditional to quantitative investment trends. Currently, the quantitative investment market in China is developing rapidly, with the capability to quickly adopt emerging technologies and adapt to market changes, demonstrating significant growth potential. However, compared to developed countries, China is still in the emerging stage in this field, facing challenges such as fragmented and unstable market information and lagging research on relevant strategies.

Moreover, individual investors constitute a large proportion of total investors in quantitative investment. These investors are not the ideal rational economic agents; personal emotional tendencies cannot be eliminated in investment decisions. Consequently, the irrational judgments of many quantitative investors inevitably indirectly affect macroeconomic investment operations.

The field of quantitative investment requires innovative forces. Significant progress has been made in the supervision of financial market sentiment and quantitative investment analysis based on BP neural networks. For instance, some domestic researchers have used BP neural networks to analyze and predict the stock market by learning from historical data to predict stock price trends. International researchers have focused on improving BP neural network algorithms and architectures to enhance their adaptability and accuracy in the financial field.

However, there are still some deficiencies. The high complexity of BP neural networks can easily lead to overfitting, necessitating appropriate regularization methods to avoid this issue. The results of BP neural networks sometimes lack sufficient clarity in their explanations, making it challenging for investors to fully understand and trust the decision-making basis. In summary, this research aims to establish a quantitative investment model based on BP neural networks to achieve automatic monitoring and rapid response to changes in financial market sentiment, thereby improving the efficiency and accuracy of investment decisions.

1.3 Research Tasks

1. Extract the main indicators related to the "digital economy" sector based on the provided indicators.
2. Set the training set to the "digital economy" sector index recorded every 5 minutes from July 14, 2023, to December 31, 2023. Using the indicators extracted in Task 1, predict the trading volume of the "digital economy" sector index for every 5 minutes from January 4, 2024, to January 28, 2024.
3. Using the "digital economy" sector index recorded every 5 minutes from July 14, 2023, to December 31, 2023, as the training set, establish a model based on Tasks 1 and 2 to predict the closing prices of the "digital economy" sector index from January 4, 2024, to January 28, 2024.
4. Assume an initial capital of 1 million yuan and a trading commission of 0.3%. Based on the closing prices obtained in Task 3, perform trading operations on the "digital economy" sector at 5-minute intervals, and calculate the total return rate, maximum drawdown rate, and information ratio for the period from January 4, 2024, to January 28, 2024.

1.4 Research Approach

The context of the writing is as shown in Figure 1.

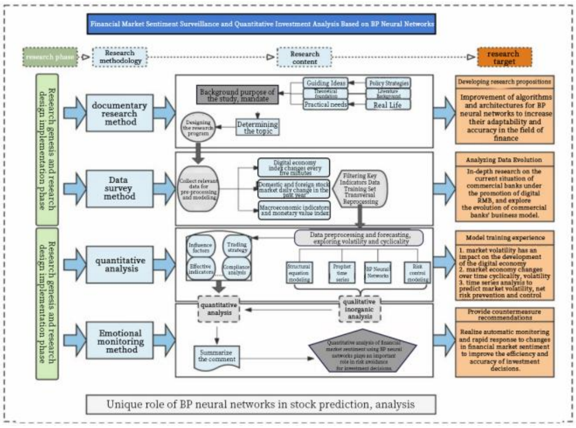

Figure 1. Technology Roadmap of the Thesis Research.

2. Model Assumptions and Analysis

2.1 Model Assumptions

- All provided parameter data is assumed to be true and valid, and related to stocks.
- It is assumed that the probability distribution of securities returns over a specific time period is the primary consideration for investors when making investment choices.
- It is assumed that investors make decisions based solely on the risk and return of securities, without considering external factors.
- It is assumed that the relevant data of adjacent stocks can be weighted to derive the stock index.
- It is assumed that the selected model evaluation indicators are reasonable and effective.

2.2 Model Analysis

2.2.1 Key Indicator Analysis

- First, select six factors as latent variables: macro market indicators, domestic stock market indicators, international stock market indicators, technical indicators, "digital economy" sector information, and other sector information. The corresponding 56 observed variables are used as manifest variables.
- Establish structural equations based on the relationships between variables in the model.
- Use the maximum likelihood estimation method for estimation and fit the goodness-of-fit

test. Consider the model's goodness-of-fit, including the likelihood ratio chi-square, CFI, NFI, IFI, and RMSEA tests.

- Perform structural equation modeling (SEM) analysis using IBM AMOS 24.0. First, use AMOS Graphics to draw the path diagram and test the causal relationships.

2.2.2 Trading Volume Analysis

- First, analyze the data from the selected indicator system from the previous question. Introduce the Prophet time series model and define the volatility function. Use the "digital economy" sector index recorded every 5 minutes from July 14, 2023, to December 31, 2023, as the training set, and the same index from January 4, 2024, to January 28, 2024, as the test set. Predict each indicator's impact on the "digital economy" sector index's 5-minute trading volume.
- To verify the effectiveness of the Prophet time series prediction and demonstrate the superiority of time series prediction index fluctuations, we also constructed a support vector regression (SVR) model using different kernel functions for analysis.
- To verify the reliability of the implementation results, we chose the HL test to validate the result set and its scheme.

2.2.3 Sector Index Analysis

- In the first and second questions, it was found that the closing prices of individual stocks exhibit regularity and periodicity in actual fluctuations.
- Introduce the BP neural network model and use a genetic algorithm. The training set consists of the "digital economy" sector index recorded every 5 minutes from July 14, 2023, to December 31, 2023, while the test set is from January 4, 2024, to January 28, 2024. Predict the closing price every 5 minutes.
- To verify the reliability of the execution results, we use the DW test to validate the result set and solution.

2.2.4 Risk Control Analysis

- Given the actual situation of China's securities industry, where short selling is not allowed, short selling is set as a constraint condition.
- Establish a risk control model based on Markovitz investment theory (Figure 2).

2.3 Definition and Symbol Description

The definition and symbol description are shown in Table 1.

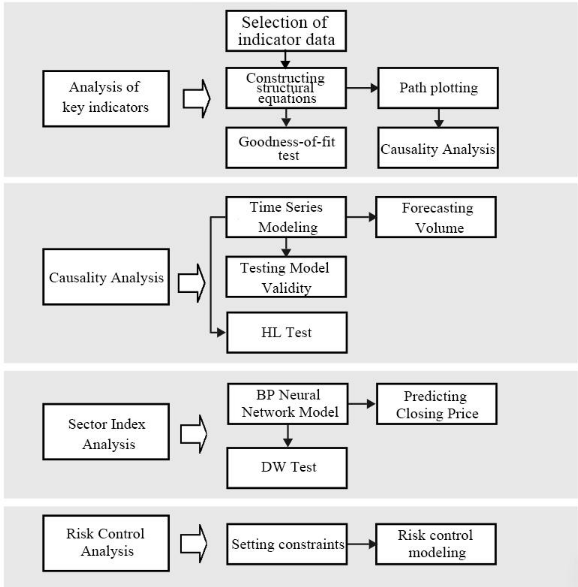

Figure 2. Model Analysis Roadmap.

Table 1. Symbol Description.

Symbol	Instructions
W	Total yield
τ_i	Day 1 same-day yieldi
φ_i	Market value of total assets at day's close
IR	Information ratio
δ	Average daily excess return
S_θ	Standard deviation of excess return
δ_i	Day 1 excess returni
T	Number of days by trading day
θ	Excess return
ρ	The yield of the CSI 500 index on the day
D_i	Net product value on Day 1i
D_j	D_iThe net worth of the following day j

2.4 Structural Equation Model Construction

The main indicators related to the "digital economy" sector are accurately screened according to the attached relevant indicator data. In this paper, six factors such as macro market index, domestic stock market index, technical index, international stock market index, "digital economy" sector information and other sector information that cannot be directly measured are discussed as potential variables, and the corresponding 56 observed variables are explicit variables (Tables 2-3).

The following three matrix equations can be used to represent the relationship between the variables in the model:

$$\mu = \alpha\gamma + \beta\lambda + \varepsilon \quad (1)$$

$$\mathrm{y} = \mathrm{r_y y} + \mathrm{c} \quad (2)$$

$$\mathbf{x} = r_x\lambda + \mathbf{p} \quad (3)$$

Formula (1) is the formula for the structural equation, which relates the potential variables using the α and β coefficient matrix and the error vector €. Formulas (2) and (3) are formulas for measuring the model, using these two sets of linear equations to connect the observed variables to the corresponding latent variables γ and λ. The equation is estimated by maximum likelihood estimation method. Then, the goodness of fit of the model is further tested by using likelihood ratio Chi-square, CFI, NFI, IFI, RMSEA tests, etc[1].

Table 3. Correspondence Table of Model Variables[2].

Latent Variables	Connotation	Measurable variables
Macro market indicators	The various indicators of the macro market can be evaluated from the following aspects	1. Purchasing managers' Index 2. Total retail sales of consumer goods 3. Gross Domestic Product 4. Consumer Price Index 5. RMB deposit rate 6. Yuan loan interest rate
Domestic stock market indicators	The index information of the domestic stock market can be measured from the following aspects	1. Volume of Shanghai Composite Index 2. Volume of Shanghai Composite Index 3. Total stock market value 4. Stock market float value 5. Csi 300 index 6. Shanghai Composite Index 7. Shanghai 50 Index 8. Shanghai A-share Index 9. Csi 500 Index 10. Chinext Index 11. Shenzhen Component Index 12. Shenzhen Composite Index Kechuang 50 index
Technical index	The common technical indicator information of the stock market can be measured from the following aspects	1. VMA 2. VMACD 3. MA 4. EXPMA 5. ARBR 6. OBV 7. BBI 8. DMA 9. MTM 10. MACD 11. BIAS 12. KDJ 13. RSI 14. BOLL
International stock market indicators	The index information of the international stock market can be measured from the following aspects	1. Dow Jones Industrial Average 2. Nasdaq Composite 3. S&p 500 index 4. Hong Kong's Hang Seng Index 5. Amex index 6. Dollar/yuan exchange rate 7. Tokyo's Nikkei 225 8. CAC40 in Paris, France 9. London's FTSE 100 index 10. Russia RTS Index 11. Italy MIB Index 12. Netherlands AEX Index 13. Eur/USD exchange rate
"Digital Economy" sector information	The stock market turnover at 5-minute intervals can be measured from the following aspects	1. The opening price every 5 minutes 2. Closes every 5 minutes 3. Highest price every 5 minutes 4. Lowest price every 5 minutes 5. Volume every 5 minutes 6. Amount per 5 minutes

Other plate information	The index information of other sectors can be measured from the following aspects	1.Index of the "digital media" sector 2."Digital Twin" sector index 3."Internet e-commerce" sector index 4."Internet" sector index 5."Fast Hand Concept" sector index

2.5 Causal Hypothesis Testing

In this paper, IBMAMOS24.0 software is used to build a structural equation model (SEM). The model is based on the existing research hypotheses, and the measurement variables and latent variables are set. First of all, the path structure diagram is drawn in AMOSGraphics, variables are edited and data is imported, which can be calculated and obtained in Table 4, as shown in the table below.

As can be seen from Table 5, the P-value corresponding to the Chi-square statistic in the model is 0.206, which is not significant. The ratio of chi-square to degrees of freedom is less than 2. In addition, CFI, NFI and IFI values are all close to 1, and RMSEA values are less than 0.02. In general, all the indexes calculated by AMOS meet the requirements of model testing and goodness of fit, indicating that the research model fits the data well, indicating that the model has a good fitting effect.

In this study, Amos was used to verify each hypothesis item by item, and the path coefficients of the model were shown in Tables 5 and 6:

The numbers in Table 2 are the standard regression coefficient and significance level of each path, the size of the number indicates the strength of the influence, and the arrow indicates the direction of the influence. Table 4 is the standard total influence coefficient table, and the figures in the table indicate the influence degree of the ordinate factors on the abscissa factors. As can be seen from Table 3, all paths have significant correlation at 5% confidence level, and all pass the test, so each basic path hypothesis is accepted.

2.6 Analysis of Model Results

The structural equation model of the relationship between the impact factor and the whole "digital economy" is established in this study, and the empirical analysis shows that the model has a good imitation of cooperation, and it is deeply studied.

2.6.1 Analysis of the Relationship Between Latent Variables and the "Digital Economy" Sector

Tables 3 and 4 indicate the magnitude of influence that latent variables have on the "digital economy" sector, ranked in descending order as follows: macro market indicators (0.623), other sector indices (0.608), domestic stock market indices (0.603), "digital economy" sector information (0.507), international stock market indices (0.473), and technology indices (0.425).

The results of testing hypotheses 1, 2, 3, 4, 5, and 6 reveal a significant positive correlation between the latent variables and the "digital economy" sector. The "digital economy" sector is notably positively influenced by macro market indicators, other sector indices, and domestic stock market indices.

Specifically, the standardized regression coefficient for Hypothesis 1 (0.623) is significantly larger than those for other hypotheses, indicating that a 1% increase in the macro market indicators can directly boost the "digital economy" sector by 0.623%. This finding is highly instructive for the development of the "digital economy" sector. The pronounced impact of macro market indicators on the "digital economy" is largely attributable to the emphasis on Gross Domestic Product (GDP).

2.6.2 Analysis of the Relationship between Latent Variables and Observed Variables

The analysis reveals that certain observed variables significantly influence latent variables. These include six factors: macro market indicators, domestic stock market indicators, technology indicators, international stock market indicators, "digital economy" sector information, and other sector information. It is also possible to compare the observed variables among these factors.

Within the macro market indicators, the coefficient for GDP is the highest (0.629), followed closely by the Consumer Price Index (0.582). This underscores that GDP has a substantial impact on the "digital economy" sector. For domestic stock market indicators, the stock market liquidity value has a relatively high coefficient (0.429), highlighting it as a crucial measure for the development of the "digital economy" sector. Among the other sector information factors, the "Internet" sector index has the highest coefficient (0.553), followed by the "digital media" sector index (0.381). Clearly, the development of the "digital economy" is directly influenced by the indices of the "Internet" and "digital media" sectors.

Table 2. Basic Path Hypothesis

Basic Path Assumptions	
H1	Various indicators of the macro market have an impact on the "digital economy" sector
H2	The index of the domestic stock market has an impact on the "digital economy" sector
H3	Common stock market technical indicators have an impact on the "digital economy" sector
H4	The index information of the international stock market has an impact on the "digital economy" sector
H5	Stock market turnover at 5-minute intervals has an impact on the "digital economy" sector
H6	Index information for its sector has implications for the "digital economy" sector

Table 4. Structural Equation Fitting Index

Model	CMIN	P	CMIN/DF	CFI	NFI	IFI	RMSEA
Default model	153.44	0.206	1.044	0.991	0.9875	0.956	0.012

Table 5. Standardized Regression Coefficients.

Null hypothesis	Paths	Standard regression coefficient	Significant level	Test results
H1	Macro market indicators-The "Digital economy" segment has an impact	0.473	**	pass
H2	Domestic stock market indicators-The "digital economy" sector has an impact	0.603	*	By
H3	Technical Indicators-The "Digital Economy" section has an impact	0.425	*	By
H4	International Stock Market Indicators-The "Digital economy" sector has an impact	0.623	**	By
H4	Digital Economy "section Information-The" Digital Economy "section has influence	0.507	**	By
H5	Other sector indicators-The "Digital Economy" sector had an impact	0.608	*	By

Note: indicates a clear association at the 0.01 level (bilateral), indicates a clear association at the 0.05 level (bilateral).

Table 6. Standardized Gross Impact Factors.

Plate path	Macro market indicators	Domestic stock market index	Technical index	International stock market indicators	"Digital Economy" sector information	Other Sector Indicators	The "digital economy" sector has an impact
The "Digital economy" segment has an impact	0.623	0.603	0.425	0.473	0.507	0.608	1.000
Purchasing Managers Index	0.575	0.000	0.000	0.000	0.000	0.000	0.000
Gross Domestic product	0.000	0.629	0.000	0.000	0.000	0.000	0.000
Consumer Price Index	0.362	0.000	0.000	0.000	0.000	0.000	0.000
Total retail sales of consumer goods	0.482	0.000	0.000	0.000	0.000	0.000	0.000
RMB deposit interest rate	0.000	0.576	0.000	0.000	0.000	0.000	0.000
RMB loan interest rate	0.000	0.000	0.553	0.000	0.000	0.000	0.000
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
"Digital media" sector index	0.000	0.000	0.000	0.662	0.000	0.000	0.000
The "digital twin" sector index	0.000	0.000	0.000	0.000	0.427	0.000	0.000
"Internet e-commerce" sector index	0.000	0.000	0.000	0.000	0.476	0.000	0.000
"Fast Hand Concept" sector index	0.000	0.000	0.000	0.000	0.618	0.000	0.000

3. Forecasting Index Fluctuations with the Prophet Time Series

3.1 Data Preprocessing

In addressing this task, we use the "digital economy" sector index data recorded every 5 minutes from July 14, 2023, to December 31, 2023, as the training set. This index data, recorded at 5-minute intervals, will then serve as the test set for the period from January 4, 2024, to January 28, 2024. We extract various indicators to predict the 5-minute trading volume of the "digital economy" sector index.

Within this model, we focus solely on the primary indicators relevant to the "digital economy" sector, disregarding actual trading volumes and other factors. It's important to note the pronounced cyclical nature of stock movements.

The overall situation across various industries can be seen in Figures 3-7. Notably, both the internet and e-commerce sectors have declined across the board, whereas the digital twin and Kuaishou concept sectors have seen a general increase. The digital media sector, however, has maintained a relatively stable trend. Figures 6 and 7 show that both the EUR/USD and USD/CNY exchange rates are in a continuous decline, with the EUR/USD nearing parity. These figures highlight the irregular individual movements of the constituent stocks, suggesting that a composite index might be more appropriate for quantification.

The final formula for calculating the composite index of constituent stocks is as follows:

$$\lambda_{ixj} = A_{ixj} + B_j \tag{4}$$

Using the above component index calculation method, the change trend of the component composite index of the attached stocks is as shown in Figure 8.

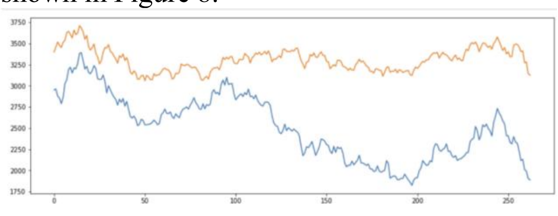

Figure 3. Digital Media and the Internet

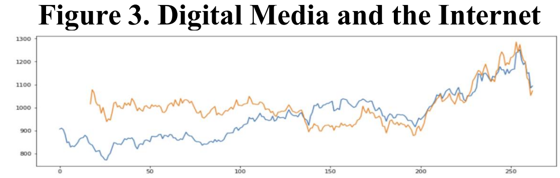

Figure 4. Digital Twins and the Fast Hand Concept

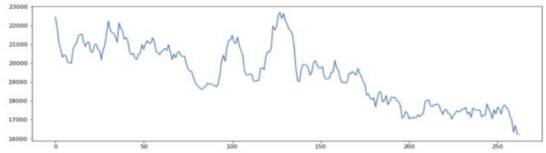

Figure 5. Internet E-commerce

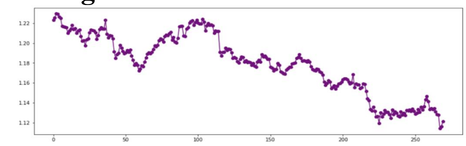

Figure 6. Exchange Rate of the Euro to the United States dollar

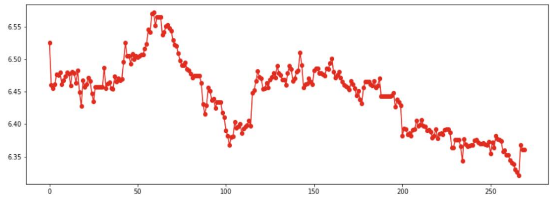

Figure 7. Exchange Rate of the United States Dollar to the Renminbi

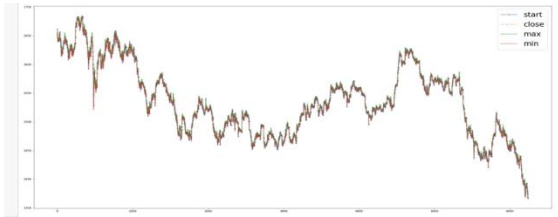

Figure 8. Stock Market Opening Price, Closing Price, Highest and Lowest Price

3.2 Basic Approach

We thoroughly cleaned the data and extracted features from the dataset to assess index fluctuations comprehensively. To forecast future index fluctuations, we employed a time series model.

3.3 Model building

Step1: Considering the volatility and periodicity of the cleaned data, we assume a time series, for any, the autocovariance function of the series is defined as follows:

$$\{X_t, t \epsilon T\} t, s \epsilon T \{X_t\} \gamma(t, s) \tag{5}$$

$$\gamma(t,s) = E(X_t - \mu_t)(X_s - \mu_{ts})$$

The autocorrelation coefficient (t,s) is defined as follows: ρ

$$\rho(t,s) = \frac{\gamma(t,s)}{\sqrt{DX_t \cdot DX_s}} \tag{6}$$

Step2: We construct the general structure as follows:

$$y(t) = g(t) + s(t) + h(t) + \varepsilon \tag{7}$$

Step3: Set a series of transition points in the time series, the growth rate will change, is the change in time, so as to build a vector, $S_j, j =$

$1, \cdots, S \delta_j t_j a_j(t) \epsilon \{0,1\}$

$$a_j(t) \begin{cases} 1, t \geq s_j \\ 0, t < s_j \end{cases} \forall jT, \begin{cases} \delta_j = 0, W.P., \frac{S}{T} \\ \delta_j, Laplace(0, \lambda), W.P., \frac{S}{T} \end{cases} \quad (8)$$

step4: The bias function m changes accordingly according to the change of growth rate, so that it is connected with the tail of the time segment. At the turning point j, the bias amount is adjusted appropriately. The formula is as follows:

$$g(t) = \frac{C(t)}{1 + \exp(-(k + a(t)^T)\delta)(t - m + a(t)^T)\gamma)} \quad (9)$$

When the scope of the prediction problem is still expanding (that is, it has not reached the saturation state), then the segmented continuous growth model works well in some times, the formula is as follows:

$$g(t) = (k + a(t)^T\delta)t + (m + a(t)^T\gamma) \quad (10)$$

Step5: Using Bayesian architecture, a posterior distribution can be obtained by setting multi-layer prior distribution, and maximum likelihood estimation of growth rate scaling parameter can also be used:

$$\lambda = \frac{1}{s}\sum_{j=1}^{s} |\delta_j| \quad (11)$$

Step6: Set P as a time series of regular cycle length. The estimated value of any smoothing period effect is as follows:

$$\mathbf{s(t)} = \sum_{n=1}^{N}\left(\mathbf{a_n cos}\left(\frac{\mathbf{2\pi nt}}{\mathbf{p}}\right) + \mathbf{b_n sin}\left(\frac{\mathbf{2\pi nt}}{\mathbf{p}}\right)\right) \quad (12)$$

3.4 Model Solution

We utilized machine learning to obtain the following fitted results from the time series analysis (Figure 9):

The results indicate that the "digital economy" sector index showed significant volatility every 5 minutes from January 4, 2024, to January 28, 2024. The highest trading volume was 22,579,060 shares, while the lowest was 5,044,510 shares. The peak volumes occurred at 15:35 on January 13, 2023, and 19:55 on January 22, 2023, respectively.

We also predicted the accuracy of the forecast, with the results as shown in Table 7.

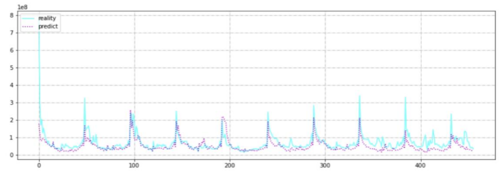

Figure 9. Forecasting Results of Time Series. Analysis

Table 7. PROPHET Time Series Accuracy Prediction.

PREDICTION TIME (MINUTES)	MAD	MSD	MAPE
0	111.952	5.23	0.0879%
5	247.809	6.72	0.8090%
10	161.805	11.25	1.6600%
15	234.247	3.28	0.0000%
20	138.142	6.37	0.1663%
25	256.591	7.52	0.6888%
30	138.189	11.36	0.1805%
35	291.321	5.28	1.6749%
40	178.522	11.62	0.4002%
...	...	...	...
...	...	...	...
...	...	...	...
35970	213.438	11.58	0.7995%
35975	259.531	9.36	0.6946%
35980	294.321	15.32	1.1469%
35985	133.279	3.23	1.3808%
35990	314.470	6.27	0.1561%
35995	234.261	8.26	0.4209%
36000	176.529	1.51	0.0079%

Using the time series model, the prediction accuracy for the highest and lowest values were 83% and 90%, respectively. This indicates a good fit, with both the mean absolute error and mean percentage error being relatively low, thereby further validating the accuracy of the predictions.

4. BP Neural Network Prediction Model

4.1 Data Preprocessing

In this task, we use the "digital economy" sector index data recorded every 5 minutes from July 14, 2023, to December 31, 2023, as the training set, and the same index data from January 4, 2024, to January 28, 2024, as the test set. Based on selected indicators and trading volumes, we construct a BP neural network model to predict the "digital economy" closing price at 5-minute intervals[3] （Figures 10-12）.

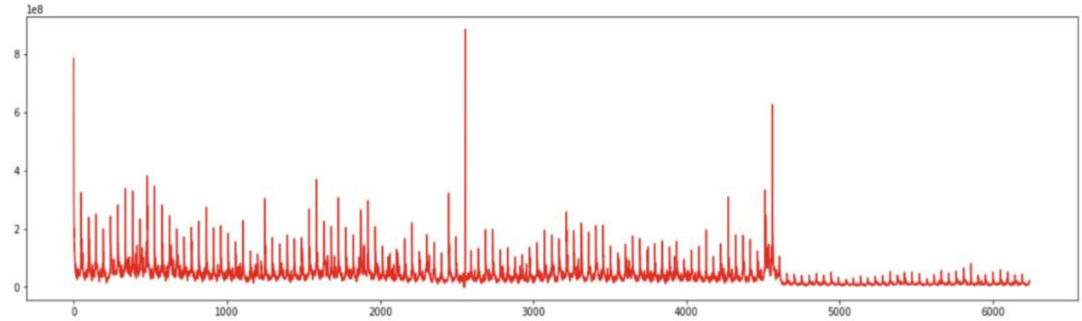

Figure 10. Volume.

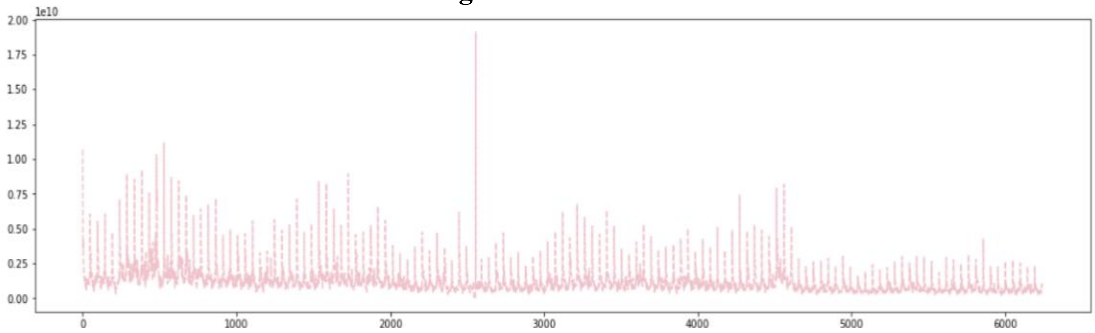

Figure 11. Turnover.

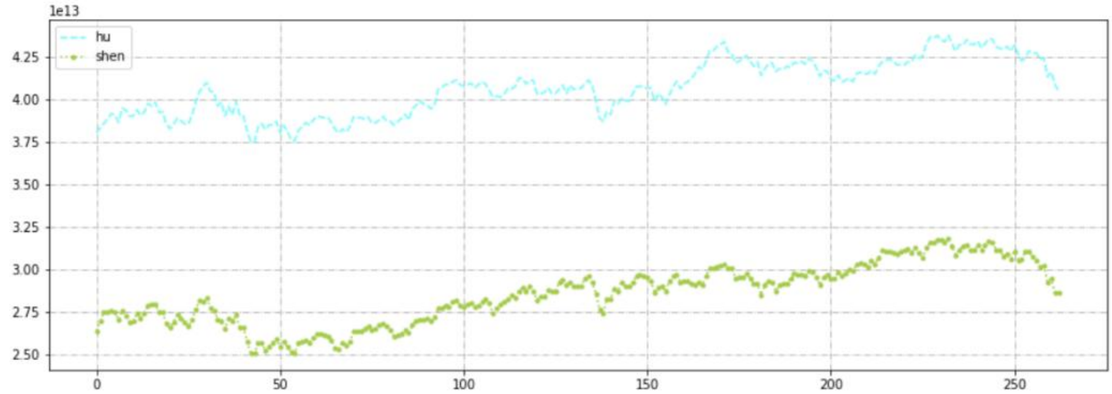

Figure 12. Turnover Market Capitalization of Shanghai and Shenzhen Stock Markets.

The stock closing price exhibits clear cyclical changes, as seen in the figure. Key influencing factors in stock prediction include time, gross domestic product (GDP), stock market liquidity, the "Internet" sector index, and the "Digital Media" sector index. While the closing and opening prices exhibit some volatility and approximate periodicity, their statistical characteristics differ, and the manifestation varies among different stocks.

4.2 Basic Approach

Using forward analysis based on the network topology, we employ data from Appendix 3 as input signals to obtain actual output values. Error signals, which are feedback signals following the network topology, are used to adjust the network weights continuously. This process minimizes the difference between the actual output and the desired output, thereby refining the actual output to closely match the expected output[4].

4.3 Model Building

Step1: Initialize the network, determine the number of input layers, the number of hidden layer nodes and the number of output layers.

Step2: Calculate the output of the hidden layer. X When the input variable is, the connection weight of the input layer and the hidden layer is the threshold value of the hidden layer, and the hidden layer output is set to:$w_{ij}aH$

$$H_j = f(\sum_{i=1}^{n} w_{ij}x_i - a_j)j = 1,2,3\cdots 1 \qquad (13)$$

Where l is the hidden layer incentive function with various manifestations. The incentive function selected in this paper is as follows:

$$f(x) = \frac{1}{1 + e^{-x}} \qquad (14)$$

Step3: Calculate the output layer result. H,In the above formula, the output of the hidden layer is the weight value is and the threshold value is, then the output result of the network can be expressed as the following formula.wijbO_k

$$O_k = \sum_{j=1}^{1} H_j w_{jk} - b_k k = 1,2,3 \cdots m \qquad (15)$$

Step4: Calculate the error, calculate the prediction error according to the output result predicted by the network and the output expected by the targetOYe_k

$$e_k = Y_k - O_k k = 1,2,3 \cdots m \qquad (16)$$

Step5: Adjust the weights. According to the network established by each system, the connection weight between each network layer should be adjusted according to the error e.

$$w_{ij} = w_{ij} + \gamma H_j (1 - H_j) x(i) \sum_{k=1}^{m} w_{jk} e_k \qquad (17)$$

In formula,$i = 1,2,3 \cdots n; j = 1,2,3 \cdots l$

$$w_{jk} = w_{jk} + \gamma H_j e_k j = 1,2 \cdots l; k = 1,2 \cdots m \qquad (18)$$

Where is the learning rate.γ

Step6: Adjust the threshold. a, b Knowing the prediction error of the system model, the threshold value of each network node can be calculated.

$$a = a_j + \gamma H_j (1 - H_j) \sum_{k=1}^{m} w_{jk} e_k] = 1,3 \cdots l \qquad (19)$$

$$\mathbf{b} = \mathbf{b_k} + \mathbf{e_k k} = 1,2 \cdots \mathbf{m} \qquad (20)$$

The neutrals of the formula are the thresholds of the network nodes of each layer.$k_a k_b$

Step7: Check whether the algorithm is iterated, if not completed, go back to Step2 to continue the cycle operation.

4.4 Model Solution

Using machine learning, we obtained the following BP neural network prediction results (Figure 13):

The results indicate that the trading volume from January 4, 2024, to January 28, 2024, mostly exhibits periodicity, with peaks at 12:00 noon and 19:00 in the evening. However, both the highest and lowest prices show a downward trend, reaching almost the lowest point since January. The closing prices for the entire month range from 1,586.7526 to 1,365.1680 yuan, with the highest and lowest values occurring on January 25, 2023, and January 6, 2023, respectively.

We also predicted accuracy, with the results as shown in Table 8.

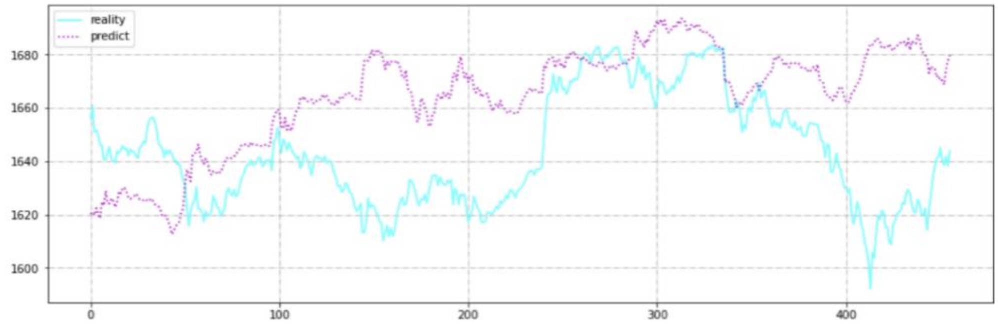

Figure 13. BP **Neural Network Prediction and Fitting Results.**

Table 8. Accuracy Prediction of BP Neural Network.

Prediction time (minutes)	MAD	MSD	MAPE
0	151.952	5.26	0.3454%
5	247.809	3.21	0.4548%
10	161.805	5.25	0.1368%
15	234.247	1.39	0.8876%
20	138.142	4.28	0.2800%
25	456.591	1.35	0.3626%
30	388.189	4.23	0.6885%
35	271.321	1.39	0.4030%
40	158.522	8.36	0.6529%
...	...	...	...
...	...	...	...
...	...	...	...
35970	293.438	7.25	0.0822%
35975	359.531	6.32	0.5173%
35980	234.321	5.34	0.1004%
35985	363.279	1.39	0.1106%
35990	174.470	1.02	0.0260%
35995	354.261	1.96	0.0427%
36000	106.529	2.46	0.7292%

5. Markowitz Risk Control Model

5.1 Model Construction

According to the regulations of China's securities market, short selling of stocks is prohibited, so the constraint conditions of the distribution coefficient can be known as:

$$w_i \geq 0 \qquad (21)$$

Therefore, the investment return maximization model is as follows:

$$\max = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i r_j \quad (22)$$

$$\begin{cases} \sum_{i=1}^{n} w_i \sigma_i^2 \le \sigma_0^2 \\ w_i \ge 0 \\ \sum_{i=1}^{n} w_i = 1 \end{cases} \quad (23)$$

The main goal of mathematical models built under the guidance of Markowitz's investment theory is to maximize investment returns. The risk control is mainly determined by maintaining the variance parameter of the investment stock price at a certain level. While controlling the risk, the aim is to maximize the return on investment. The goal is to find the best portfolio strategy, so consider introducing additional optimization goals for reducing investment value at risk in this article, i.e., minimizing investment value at risk[5].

The final value-at-risk minimization objective function is:

$$\min\emptyset = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i \sigma_j \quad (24)$$

For the calculation of value at risk for each stock, we can quantify it using the value at risk model. We can also measure the discrete volatility of each stock through the variance of the return rate of each stock, that is, the risk degree of each stock. In this paper, the variance of return rate is chosen to solve. Minimizing the VAR objective function can be simplified as:

$$\min\emptyset = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \operatorname{Cov}(Bi, Bj) \quad (25)$$

Where, Bi represents the average return of the i th stock, and the final multi-objective optimization model is as follows:

$$\min\emptyset = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \operatorname{Cov}(Bi, Bj) \quad (26)$$

$$\max r = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i r_j \quad (27)$$

$$\begin{cases} \sum_{i=1}^{n} w_i \sigma_i^2 \le \sigma_0^2 \\ \sum_{i=1}^{10} w_i \mu_i \ge \mu_0 \\ w_i \ge 0 \\ \sum_{i=1}^{n} w_i = 1 \end{cases} \quad (28)$$

5.2 Model Solving

To solve the above model, this paper can use the main objective method and hierarchical sequence method to transform the multi-objective optimization model into a single objective model. Converting multi-objective optimization multi-objective into single objective solution can convert the model into the following two cases[6]:

Step1: Calculate the total return rate.

$$\max r = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i r_j \quad (29)$$

$$\begin{cases} \sum_{i=1}^{n} w_i \sigma_i^2 \le \sigma_0^2 \\ w_i \ge 0 \\ \sum_{i=1}^{n} w_i = 1 \\ W = [(1 + \sum_{i=1}^{n} \tau_i) - 1] \times 100\% \\ \tau_i = \frac{\varphi_i - \varphi_{i-1}}{\varphi_{i-1}} \end{cases} \quad (30)$$

Where, is the total return rate, represents the return rate on the day of the day, represents the market value of the total assets at the close of the day, represents the market value of the total assets at the close of the previous day.$W\tau_i i\varphi_i\varphi_{i-1}$

Step2: Calculate the information ratio

$$\begin{cases} IR = \frac{\beta}{S_\theta} \\ \beta = \frac{\beta_i}{T} \\ \theta = \tau_i - \gamma \times 90\% \end{cases} \quad (31)$$

Where IR represents the information ratio, represents the average daily excess return, represents the standard deviation of the excess return, represents the day 1 excess return, represents the total number of days on a trading day basis, represents the excess return, represents the index return.$\beta s_\theta \beta_i i T\theta\gamma$

Step3: Calculate the maximum retracement rate

$$D = \max\frac{D_i - D_j}{D_i} \quad (32)$$

Where, denotes the net value of the product on day 1, and denotes the net value of the following day j.$D_i i D_j D_i$.

5.3 Model Result Analysis

The results showed that this part of the asset was used more efficiently and effectively, and the overall return rate reached 13.26 percent. The information ratio of 2.79% indicates that the investment manager has the ability to obtain the excess return under the same risk above the market level. The maximum decline is 11.83%, and the decline is less than 20%, indicating a stable return (Figure 14).

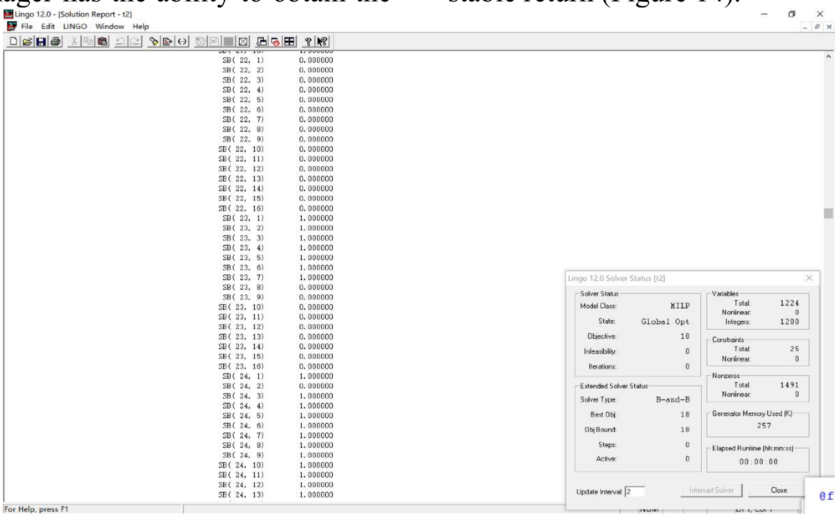

Figure 14. Calculation Results of Risk Control Model.

6. Analysis and Evaluation of the Model

6.1 DW Model Testing

Because in the processing of time series data, it will cause series-related problems, i.e. $\varepsilon_t = \rho\varepsilon_{t-1} + u_t$Therefore, it is necessary to carry out the series correlation test, that is, the DW test, to ensure the validity of the predicted model results.

The statistic DW is defined as:

$$\mathrm{DW} = \sum_{t=2}^{n} \frac{(e_t - e_{t-1})^2}{\sum_{t=2}^{n} e_t^2} \tag{33}$$

By, therefore, and. $\mathrm{DW} \approx 2(1-\hat{\rho}), -1 \leq \hat{\rho} \leq 1, 0 \leq \mathrm{DW} \leq 4$If it is near 0, then DW is near 2, and the autocorrelation is weak (or non-existent); $\hat{\rho}\varepsilon_t$If it is near 0, then DW is close to 0 or 4, and the autocorrelation is strong.$\hat{\rho} \pm 1\varepsilon_t$

The DW sum is calculated using the relevant program as follows (Figure 15):$\hat{\rho}$

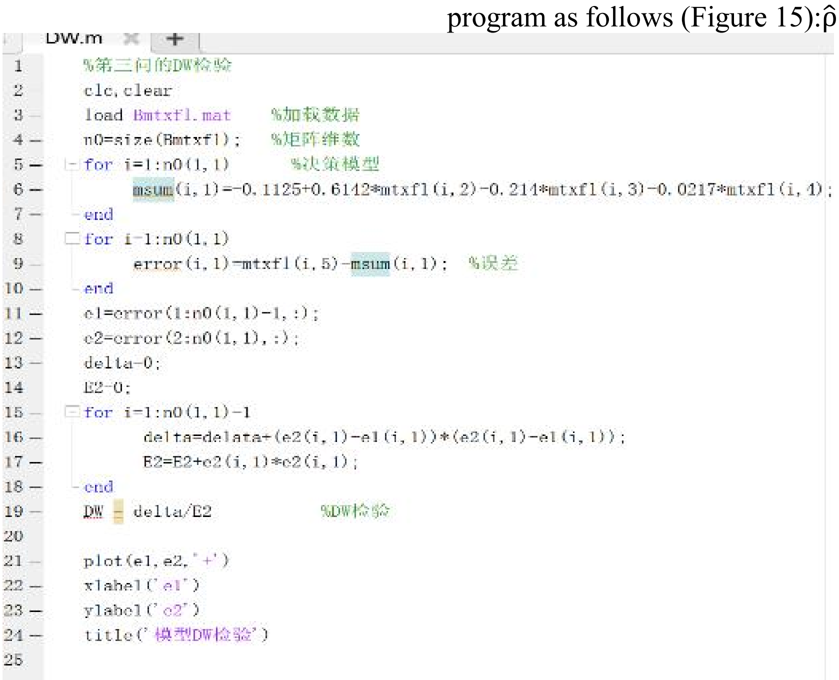

Figure 15. Diagram of the DW Inspection Runtime Program

Run the program, after DW inspection can be obtained:

$\mathrm{DW} = 0.01$，$\hat{\rho}$ =0.95

To get this distribution map (Figure 16).

In addition to the large prediction error of individual positions, it is still good and reasonable on the whole.

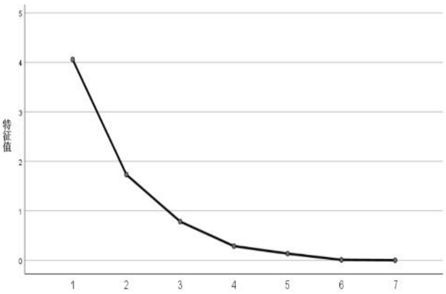

Figure 16 Lithotripsy Diagram

6.2 HL test of Time Series

To calculate the number of observations we expect, the HL test takes the average of the predicted probabilities in the group and multiplies it by the number of observations in the group (Table 9).

Step1: Group the observations according to the predicted probabilities of the first group.

Step2: Secondly, according to the grouped data, set the number of observation indicators in group i; n_jCalculate the number of observations in group i and the mean of the prediction probability.

Step3: HL statistic formula is as follows:

$$HL = \sum_{j=1}^{n} HL(j) = \sum_{j=1}^{n} \frac{[y(j) - n_j p(j)]^2}{n_j p(j)[1 - p(j)]} \quad (34)$$

According to the HL test table, since p=0.037<0.05, it is considered that the prediction of time series model is relatively stable, and this result is consistent with intuitive imagination.

6.3 Model Strengths and Weaknesses

6.3.1 Practicality of the Model

In addressing issues pertaining to stock trading volumes and indices within the digital economy sector, we successfully applied both the PCA decision model and Prophet time series model. This seamless integration facilitated practical implementation in computer applications. Moreover, the model demonstrated commendable performance in handling multi-output problems, indicating a degree of fault tolerance.

Table 9. HL Checklist

j	**Minimum**	**Maximum**	$y(j)$	$n_jp(j)$	n_j-$y(j)$	n_j-$n_jp(j)$	n_j	**HL (j)**
1	0.041	0.068	3	0.957	15	16.043	13	6.720
2	0.155	0.265	4	1.800	18	12.253	16	3.013
3	0.221	0.301	2	3.426	15	11.876	18	0.565
4	0.253	0.568	6	4.038	23	12.084	15	6.421
5	0.264	0.363	7	4.675	16	13.325	18	1.228
6	0.288	0.583	4	4.747	23	12.253	17	0.163
7	0.311	0.802	7	5.124	16	13.325	17	0.983
8	0.344	0.041	8	5.918	12	16.043	19	0.906
9	0.375	0.155	3	6.066	17	12.253	16	0.442
10	0.465	0.013	3	7.504	30	11.876	18	0.680
Total			47	44.255	183	131.331	167	21.121
HL statistic			14.347		**P value**		0.037	

6.3.2 Algorithm Effectiveness

Our utilization of methods such as the Prophet time series model effectively mitigated the mutual influence among original data components. Consequently, the resulting model and its solutions were conducive to computational implementation, exhibiting desirable traits of randomness and scalability.

6.3.3 Model Limitations

Despite the establishment of the PCA model, its parallel mechanism failed to fully exploit its potential. Additionally, principal components with low contribution rates may harbor crucial information regarding sample differences. Furthermore, the Prophet time series model exhibited certain errors, indicating areas for refinement and improvement.

6.3.4 Model Generalization and Future Prospects

By reducing fundamental assumptions, the relevant models can be further applied to various aspects of stock funds. This suggests that the model's conceptual framework, algorithms, and results possess a degree of universality and can be extended to broader contexts within the field, promising further advancements and applications.

References

[1] Chen Q, Qi F, et al. Review of supplier selection research [J]. *Foreign Economics &*

Management 2009,31 (5) :31-37.

[2] Shi, L. The application of decision model in teaching quality assessment [J]. *Educational Informatization* 2006 (5) :54-55.

[3] Deng J. Grey Theory Foundation. Wuhan: Huazhong University of Science and Technology Press, 2002:35-38.

[4] Niu H. Dynamic relationship between industrial structure and economic growth in Jiangsu Province: Based on Grey correlation model [J]. *Technical Economics and Management Research* 2010(6):31-34.

[5] He Y, Li N, Bai C. Quan. Evaluation Method of 26Supply chain performance based on G1-DEA and TOPSIS-Grey correlation Analysis: Taking Enterprise Internal Supply Chain as an example. *Science and Technology Management Research* 2016 (16): 68-73.

[6] Si S, Sun X. Mathematical Modeling algorithms and applications [M]. Beijing: National Defense Industry Press, 2011.

