ISSN 2616-5902 Vol. 4, Issue 5: 35-41, DOI: 10.25236/AJBM.2022.040508

The Research on Optimal Investment Strategy Based on Grey Forecasting BP Neural Network Model

Shikai Wu, Shuo Zhao

Anhui University, Hefei, Anhui, 230601, China

Abstract: *Quantitative Trading with stable investment performance has been widely used in Europe, the United States and other countries for more than 40 years. This paper focuses on how to design a quantitative trading strategy to optimize the investment strategy, and then design use Grey Forecasting BP neural network model based on Markowitz model to facilitate the computation of different scenarios for three different types of trader investments while building a dynamic programming model later. And the robustness and sensitivity analysis of the model are tested, and the model performs stably. And then, we adopt a short-term investment model and pay attention to the rise and fall of the value curve in real time. On Fridays of the week, there is an optional scheduled vote. Through this investment method, we have achieved very desirable returns. Finally, the model is less sensitive, has good market adaptability, and has some realistic significance.*

Keywords: *Grey Forecasting; BP neural network; Quantitative investment; Trading strategy*

1. Introduction

In recent years, gold and bitcoin have been favoured by market traders as two volatile assets. Some investors would rather take a high risk to pursue returns, while conservatives would abandon some of the returns to avoid some risk. With the advent of the big data era, trading strategy has become more and more important. Outstanding trading strategies can bring maximum returns to investors. Quantitative Trading with stable investment performance has been widely used in Europe, the United States and other countries for more than 40 years.

In this paper, we seek to design a model with the best trading strategy with the idea of quantitative trading. To solve the problem, we will use Portfolio Theory, Dynamic Programming, Metabolism Gray Model, and some other ways to determine the optimal strategy.

2. Model Building and Results

2.1. Price Prediction Model

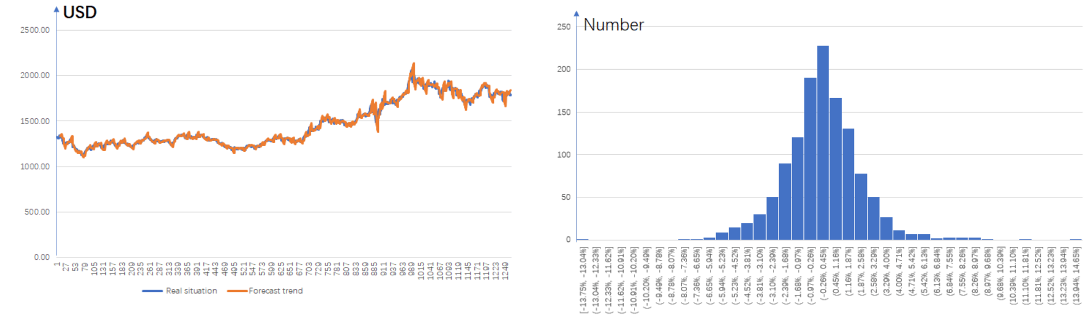
2.1.1. Metabolism GM(1,1) Model

In this case, considering that the closing price of the day can be known before the closing price of the afternoon every day, the metabolism GM model can be used to predict the future data.

By adding new information $a^{(0)}(n+1)$ and removing old information $a^{(0)}(0)$, the prediction model with $a^{(0)}=\left(a^{(0)}(2),\cdots,a^{(0)}(n),a^{(0)}(n+1)\right)$ becomes the metabolism GM(1,1) model. In order to obtain the data for the previous prediction, the data of the previous 20 days are used to start the prediction of the later data. In other words, the first 20 days do not buy or sell anything. For gold, we forecast the next five days. For bitcoin, we forecast the next seven days.

After completing the quasi-exponential law test, we predict the data and perform exponential fitting to visualize the fitting results. The prediction is shown in the Figure 1 below.

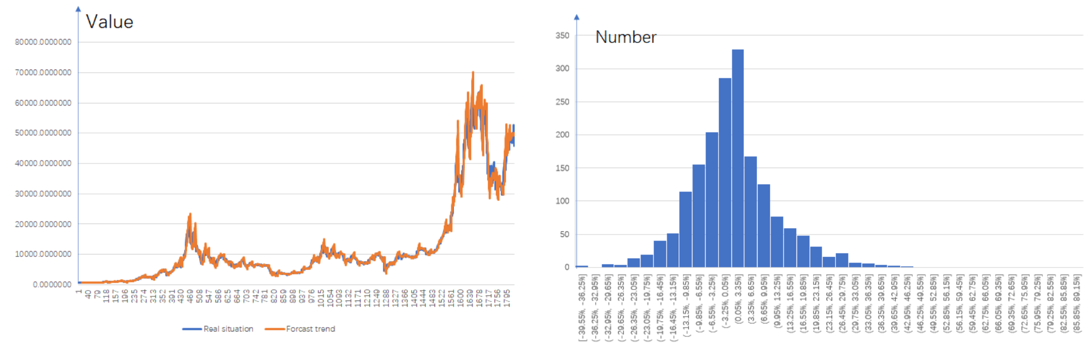
ISSN 2616-5902 Vol. 4, Issue 5: 35-41, DOI: 10.25236/AJBM.2022.040508



(a) Comparison of predicted and true gold value *(b) Gold's forecast error rate*

Figure 1: Forecast of Gold.

Known from the analysis of Figure 1, forecast movements in the price of gold is relatively stable in the first years and second year. Forecast price movements of bitcoin at the beginning of the second and fifth year there were large fluctuations.



(a) Comparison of predicted and true Bitcoin value *(b) Bitcoin's forecast error rate*

Figure 2: Forecast of Bitcoin.

From the analysis in Figure 2, we can know that the most of the prediction for bitcoin error rate within 10% and most of prediction for gold error is within 5%.

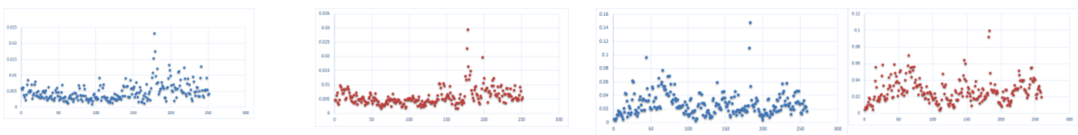
2.1.2. Metabolism GM(1,1) Model Test

Using GM(1,1) to predict the data, the predicted effect can be tested in the following three ways:

(1) Residual test.

(2) Stage ratio deviation test:

For data calculation through MATLAB, we visualized the residual and stage ratio deviation obtained each time, as shown in Figure 3 and Figure 4. As it can be seen from the Figure 3, except for some outliers, 99% of the predicted residuals and stage ratio deviations are within 0.02, which meets our requirements. As it can be seen from the Figure 4, except for some outliers, 99% of the predicted residuals and stage ratio deviations are within 0.1, which meets our requirements.



(a) Residual-GOLD *(b) Class ratio dispersion-GOLD* *(c) Residual-Bitcoin* *(d) Class ratio dispersion-Bitcoin*

Figure 3: Error Detection

2.1.3. BP neural network prediction data

BP network [5], also known as back-propagation neural network, is a supervised learning algorithm with strong self-adaptive, self-learning, and non-linear mapping capabilities. It can better solve the problem of less data, poor information, and uncertainty problem, and is not limited by non-linear models. In this problem, we use a neural network to predict the ups and downs of trading time periods. We first train our neurons with data before the trading day, then use it to predict our future value trends.

In the process of our data to predict, we only have one input, through the analysis of time data to train our neurons, Log Sigmoid is activated to make use of our neural network. Neurons work as follows:

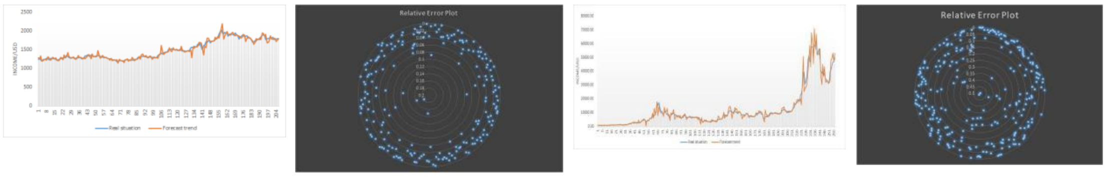
(1) The input terminal accepts the input signal xi.

(2) Find the weighted sum of all inputs. $net=\sum_{i=1}^{n} w_i x_i$.

(3) Output the result after nonlinear transformation of the net. That is, $y=f(net)$.

In this problem, we use multiple layers of neurons to make predictions on the data. First, we partitioned the dataset. Then we use a two-layer hidden layer neural network (15*15). Among them, the activation functions in the two hidden layers are both Log Sigmoid functions.

After setting the error parameters, we start to train and predict the data in a loop, and the data set for each training will be updated continuously with the increase of the known data size. Finally, we have the data predicted by the neural network. To justify our data, we plotted and analysed the data. The real and predicted data are then plotted within the same image. By looking at the movement of the two curves, we can clearly see that our forecast data fits real data well.



(a) Comparison of predicted and true gold value *(b) Gold's forecast error rate* *(c) Comparison of predicted and true Bitcoin value* *(d) Bitcoin's forecast error rate*

Figure 4: Forecast Values.

From the analysis of Figure 4, it can be seen that the gold price trend in the first two years is relatively stable, while the bitcoin price trend is predicted to fluctuate greatly at the beginning of two years and last year. Analysis of Figure 4 shows that most of the predicted bitcoin error rates are within 10%, and most of the gold forecasts error is less than 5%.

The following formula is used to measure prediction algorithm for the whole training set of prediction accuracy. It's essentially adding up the "losses" for each sample and averaging them. The loss function for the whole training set is called cost function. Its calculation results, the greater the cost is larger, the forecast is not accurate. The loss function is defined as follows:

$$J(w,b)=\frac{1}{m}\sum_{i=1}^{m} L\left(\hat{y}^{(i)},y^{(i)}\right)=-\frac{1}{m}\sum_{i=1}^{m}\left[\left(y^{(i)}\log\left(\hat{y}^{(i)}\right)+\left(1-y^{(i)}\right)\log\left(1-\hat{y}^{(i)}\right)\right)\right] \tag{1}$$

To sum up, based on prediction-error Figure 5, our BP neural network prediction can better forecast data, and the error is small. But when the data set is small, we can't get good fitness.

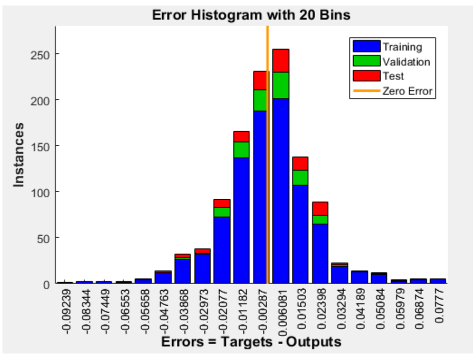


Figure 5: Neural Network Prediction Fitting Degree and Error Histogram.

2.1.4. Comparison of Models

Based on the above analysis, we can easily see that the prediction accuracy of the metabolism GM(1,1) is slightly better than that of the neural network on the whole. The prediction error rate is within a reasonable range and it has also passed the grade ratio and residual test. The prediction effect of neural

ISSN 2616-5902 Vol. 4, Issue 5: 35-41, DOI: 10.25236/AJBM.2022.040508

network model is very poor when the early data is less, and the data fluctuation is severe. In general, we choose metabolism GM(1,1) model as the model for predicting data, then carry out decision-making analysis in the next step.

2.2. Dynamic Programming Model

2.2.1. Model Establishing

Non-linear programming [6] is a new subject formed in the 1950s. Firstly, it is necessary to select the appropriate target variable and decision variable, and establish the functional relationship between the target variable and decision variable, which is called the objective function. Then, some equations or inequalities that the decision variables should satisfy are obtained by abstracting the constraints. The general mathematical model of nonlinear programming problems can be expressed as the unknown quantities $x_1, x_2, \cdots, x_n$, so that the constraint conditions are met:

$$\begin{aligned} g_i(x_1, \cdots, x_n) \geq 0, i = 1,2, \cdots, m. \\ h_j(x_1, \cdots, x_n) = 0, j = 1,2, \cdots, p. \end{aligned} \tag{2}$$

And make the objective function $f(x_1, \cdots, x_n)$ reaches the minimum (or maximum) value. Where f, g_i and h_i are all real-valued functions defined on some subset D (do-main) of n-dimensional vector space R_n, and at least one of them is nonlinear.

The above model can be abbreviated as:

$$\begin{aligned} &\min\ f(x) \\ &s.t. g_i(x) \geq 0, i = 1,2, \cdots, m. \\ &s.t. h_j(x) \geq 0, j = 1,2, \cdots, p. \end{aligned} \tag{3}$$

Where $x = (x_1, x_2, \cdots, x_n)$ belongs to the domain D, the symbol min stands for “minimize” and the symbol s.t. Means constrained.

2.2.2. Risk Assessment

According to Markowitz’s portfolio theory, if a portfolio selects N assets from the security pool and combines them in a certain proportion of investment, the return of the portfolio in a certain investment cycle is quantified by the weighted average sum of the return rates of all assets, which can be calculated by Formula:

$$r_p = \sum_{i=1}^{N} x_i \overline{R}_i \tag{4}$$

r_p is the expected return rate of the portfolio, x_i is the capital proportion of the *i*th asset in the portfolio, meeting the budget constraint condition $\sum_{i=1}^{N} x_i = 1, x_i > 0, \overline{R}_i$ is the expected return rate of the asset.

In a real portfolio, securities tend to be correlated. They may have the same or opposite rate of return. We use covariance to express this correlation. $R = (R_1, R_2, \cdots, R_N)^T$ is assumed to be the actual return rate of each security in the portfolio, and the covariance between *i* and *j* is calculated as formula:

$$\sigma_{ij} = \text{cov}(R_i, R_j) = \frac{1}{T} \sum_{t}^{T} (r_i^t - \overline{R}_i)(r_j^t - \overline{R}_j) \tag{5}$$

T is the period of time.

2.2.3. Solution of Non-linear Programming Model

As mentioned earlier, we only make decisions on Fridays when the market of gold and Bitcoin are open. Let’s pick out Friday’s data and do the following. If Friday is T, then next Friday is T+1. Combined

ISSN 2616-5902 Vol. 4, Issue 5: 35-41, DOI: 10.25236/AJBM.2022.040508

with Markowitz's portfolio theory, when we are standing in T phase and forecasting T+1 phase based on the previous data of T phase, we need to make a decision at the present moment to ensure that our benefits are maximized and our risks are minimized as much as possible. Therefore, we can write the objective function as:

$$f_1(X_{i+1}) = G_{i+1} \times V_{G_{i+1}} + V_{B_{i+1}} \times B_{i+1} + C_{i+1} \tag{6}$$

$$f_2(X_{i+1}) = \sigma_{gi} \times G_{i+1} \times V_{G_{i+1}} + \sigma_{bi} \times B_{i+1} \times V_{B_{i+1}} \tag{7}$$

f1 represents the benefit term and f2 represents the risk term. According to Markowitz's portfolio theory model, we set the weights of both benefits and risks at 50%. From formula(8) and formula(9), the formula(10) is obtained by calculation. The overall objective function is:

$$-0.5 f_1(X_{i+1}) + 0.5 f_2(X_{i+1}) = F(X_{i+1}) \tag{8}$$

Constraint condition:

$$Get_i = V_{G_i} \times G_i + V_{B_i} \times B_i + C_i \tag{9}$$

$$\begin{gathered} Get_i = C_{i+1} + G_{i+1} \times V_{G_i} + V_{B_i} \times B_{i+1} + \alpha_G \times |G_{i+1} - G_i| \times V_{G_i} + \alpha_B \times |B_{i+1} - B_i| \times V_B \\ [C_i, G_i, B_i] \geq 0 \end{gathered} \tag{10}$$

Get_i is the total assets that we had at time T. In the constraint conditions, it is necessary to ensure that the three components of vector **[C G B]** are greater than or equal to 0 and to ensure that the total assets today are divided into two parts:

(1) total assets of tomorrow.

(2) the cost generated in the process of transaction.

After planning, [C_i,G_i,B_i] becomes [C_{i+1},G_{i+1},B_{i+1}]. We iterate repeatedly and update [C_{i+1},G_{i+1},B_{i+1}] constantly, recording the real assets of each period as Get_i .

2.3. Results

If we do not consider the risk of investment, we will make extreme decisions. Behind the high returns, we actually bring high risks. As shown in Figure 12, the orange curve represents the case where risk is not considered and the blue curve represents the case where risk is considered. In both cases, returns increase over time. At the same time, it's easy to see that the blue curve is smoother than the orange curve. On the surface, our returns are much higher than the risks, but that may be a special case in this case, and the returns would be much lower if there were other price increases. The decisions we made were not just about maximizing returns, but about making our investments sound.

As it can be seen from the Table 2, our income increases with the growth of years and the growth rate in the last three years is greater than that in the first two years. The reason is related to the price surge of bitcoin in the last two years, so our income also increases rapidly.

After calculation, as of September 10, 2021, according to the investment plan of this strategy, we finally get more than $13,000. The annual yield is around 69%.

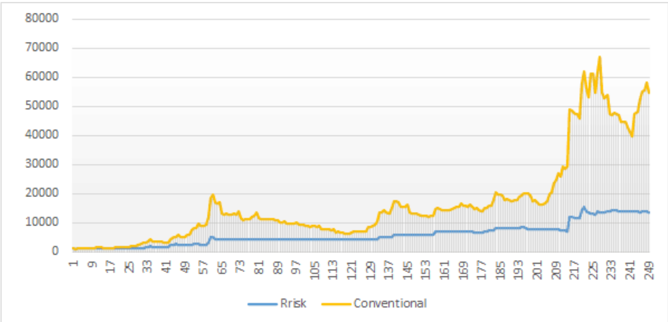


Figure 6: Comparison of Risk Assessment.

ISSN 2616-5902 Vol. 4, Issue 5: 35-41, DOI: 10.25236/AJBM.2022.040508

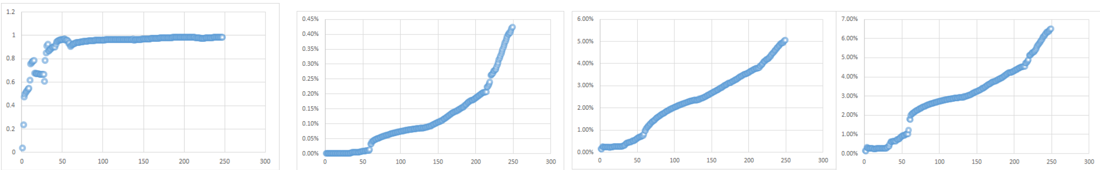
Table 1: Profit Per Year.

Total annual assets	Consider the Risk	Not Considering Risk
2016 → 2017	2511.9164	5639.828604
2016 → 2018	4139.1512	9123.306207
2016 → 2019	5865.3149	12614.3277
2016 → 2020	7897.8762	17560.67737
2016 → 2021	13551.931	54758.5948

2.4. Model Testing

2.4.1. Validating the Model

After obtaining the data results of the previous question, the rationality of the decision is demonstrated through some indicators below.



(a) R^2 (b) MSE (c) MAE (d) RMSE

Figure 7: Evaluation indicators.

The real income and the forecast income are normalized. P_{Fi} represents the forecast income normalized by T period, and P_{Ti} represents the real income normalized by T period. The following results can be obtained through MATLAB calculation. For MAE, as the gap between P_{Fi} and P_{Ti} goes on, the error will be accumulated, so MAE gradually increases. It can be seen from the figure that MAE is less than 0.5%, and the errors are all within a reasonable range. We can see from the table that the real data a in the early stage is relatively small, so the real goodness of fit R^2 deviates from 1, but with the update of the date, we get more real data P_{Ti} , and the goodness of fit of the data is gradually approaching 1.

Combining all the above data, we can derive our Sharpe ratio of 1.17 and annual return of 69.05%. With risk taken into account, our investments are more robust and can guarantee a good return every year.

2.4.2. Sensitivity analysis of decision models

In the above, we set the transaction cost as a fixed value (GOLD: 1% BITCOIN: 2%), and then analyse the sensitivity of total investment return to transaction cost in the portfolio decision model. Keep other parameters and problem solving unchanged, change the size of the transaction cost, and the result is as shown in the Figure 14.

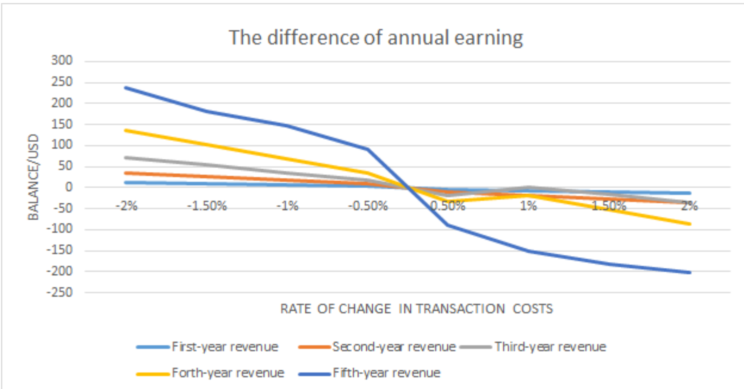
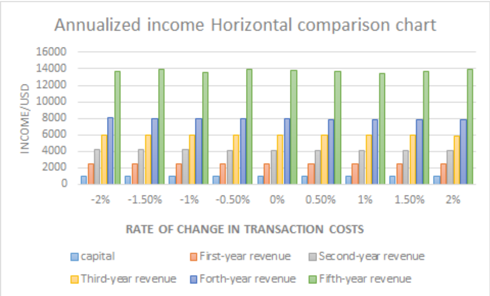


Figure 8: Sensitivity Analysis of Portfolio Strategy to Transaction Cost.

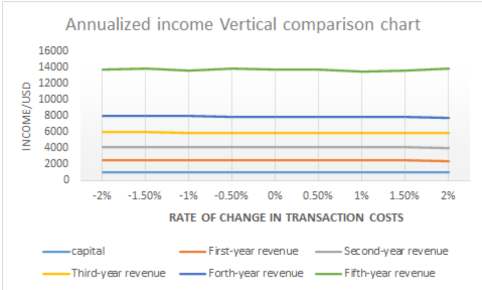
The data we obtained under the constraints of different transaction costs with the principal of $1000 is roughly the same as the first question, which is about $13,800. We calculate the difference between the income under different costs (For example, we increase or decrease the cost by 0.5%, 1%, 1.5%, 2%) and the income in the first question, and draw a broken line chart for them. By observing the line chart of the difference, it can be clearly seen that the profit difference under different costs in five years is almost 0. However, compared with the previous four years, the income of the fifth year in the line chart fluctuates significantly. Combined with the changes in the value of bitcoin and gold in the past five years,

ISSN 2616-5902 Vol. 4, Issue 5: 35-41, DOI: 10.25236/AJBM.2022.040508

we can have a reasonable explanation for the fluctuation. In the fifth year of income mainly depends on the volatility of the market, which led to dramatically change our forecast deviation. The deviation will be relatively increased.



(a)



(b)

Figure 9: Year-Over-Year and Sequential Analysis of Annualized Revenue at Different Costs.

By observing the Figure 15, we can know that under different costs, our revenue increases every year and presents an exponential growth trend. As the revenue gets higher and higher, the risk of holding the revenue also increases. According to the broken line Figure 15, through a horizontal comparison of the changes in returns under different transaction costs, it can be found that our sequential returns almost maintain a horizontal line.

In general, small changes in transaction costs will have little impact on our earnings. It shows that our model can be applicable to investments under different transaction costs. In other words, our portfolio model has good robustness.

3. Summary

Thorough consideration and analysis of the realities of different types of traders, this paper use grey forecasting model and BP neural network to forecast the data. Using the above two methods, we have a good cycle to predict the rise and fall of the value of gold and bitcoin, and the obtained data fits the rise and fall curves very well. And then, we use the Markowitz model to assess risk and rationally regulate our investment decision portfolio. Besides, we adopt a short-term investment model and pay attention to the rise and fall of the value curve in real time. On Fridays of the week, there is an optional scheduled vote. Through this investment method, we have achieved very desirable returns.

And also, by replacing the respective transaction costs of gold and bitcoin in the model, we test that both the final value and the number of transactions in the strategy move inversely to the transaction costs, and the model is not very sensitive.

However, our investment decisions fail to pay attention to more detailed changes, such as volatility over a five-day period, or an overall trend over a longer period of time. So we can try to use more prediction results and give different weights to better optimize our decisions in the future.

References

[1] Hu Y C, Jiang P, Chiu Y J, et al. Incorporating Grey Relational Analysis into Grey Prediction Models to Forecast the Demand for Magnesium Materials [J]. Cybernetics and Systems, 2021(1):1-11.
[2] Alhumaid Y, Khan K, F Alismail, et al. Multi-Input Nonlinear Programming Based Deterministic Optimization Framework for Evaluating Microgrids with Optimal Renewable-Storage Energy Mix [J]. Sustainability, 2021, 13(11): 5878.
[3] Steinbach M C. Markowitz revisited: Mean-variance models in financial portfolio analysis [J]. SIAM review, 2001, 43(1): 31-85.
[4] HH A, Zta B, Jl C, et al. Exploiting fractional accumulation and background value optimization in multivariate interval grey prediction model and its application[J]. Engineering Applications of Artificial Intelligence, 104.
[5] C. E. Scales Introduction to Nonlinear Optimization. NY:Springer-Verlag, 1985.
[6] El-Sobky B, Abo-Elnaga Y, Mousa A, et al. Trust-Region Based Penalty Barrier Algorithm for Constrained Nonlinear Programming Problems: An Application of Design of Minimum Cost Canal Sections [J]. Mathematics, 2021, 9.