**World Journal of Mathematics and Physics**
**ISSN: 2960-0103**

# COMBINED PREDICTION OF IMPROVED MULTIDIMENSIONAL GRAY MODEL AND SUPPORT VECTOR MACHINE

Harriet Kim, Haijun Bai*, Zandong Sun
*China University of Petroleum, Dongying 257061, Shandong, China.*

**Abstract:** Support vector machine improves the generalization ability through the principle of structural risk minimization. Now it is mostly used to solve the classification and regression problems of small samples, but when it is used for prediction, a single model has certain limitations. This paper proposes a prediction model that combines the improved multidimensional gray model and support vector machine to realize the complementary advantages of different models, avoid the limitations of a single model, and increase the stability of the model. Experimental simulation results show that the forecasting effect of the proposed combined forecasting model is significantly better than that of support vector machine and the model based on innovation priority accumulation method, and the forecasting accuracy is higher than that of the single forecasting model.
**Keywords:** Forecasting model; Multidimensional gray model; Combination of support vector machines; Forecasting model

## 1 INTRODUCTION

Professor Deng Julong created the gray theory in the early 1980s. This set of new methods and theories is specially used to study uncertain system problems [1]. This method and theory have made new contributions to the world system science research field. After years of development, gray theory has established a whole set of emerging discipline structure system, among which gray equation, gray matrix and gray algebraic operation system are the theoretical system based on the main content, the analysis system based on gray correlation space and gray cluster evaluation, The evaluation system is based on the method system with sequence operator and gray generation as the core, the gray target decision-making and gray relational decision-making, and the model system with gray system predictive modeling method as the core. The multidimensional gray model is mainly used to observe the degree of influence of various factors on the system [2], reflecting the functional relationship between the main factor (dependent variable) and the respective variable factors in the dynamic development process of the time series, and the "poor information" of the influence of multiple factors "Under the system problem modeling to get a better effect.
Support vector machine is a small-sample machine learning method [2], which avoids problems such as difficulty in determining the network structure, over-learning, under-learning, and local minima in methods such as artificial neural networks, and does not involve probability measures and large numbers. Law [3-4].
In 2002, Meyer D[5] compared the support vector machine with nine other existing machine learning methods for regression problems, and the results showed that the support vector machine has advantages in regression prediction and has advantages in specific data classes. very prominent. The combined prediction model realizes the complementary advantages of different models, can avoid the limitations of a single model, and can increase the stability of the model [6-7]. Existing research results show that: using other models or methods to combine with gray model to build a combined model can improve the prediction accuracy of the model [8-9].
Literature [10] improved the parameter estimation method of GM(0,N) model in order to improve the prediction accuracy. However, literature [11] proposed a model based on innovation priority accumulation method by improving the parameter estimation method of GM(1,N) model. In view of the literature [10][11], this paper uses the support vector machine to correct the residual error predicted by the multidimensional gray model based on the innovation priority accumulation method established in the literature [10] and literature [11], and proposes an improved multidimensional gray model based on the establishment of Combined prediction model of model and support vector machine. Finally, using the "China Statistical Yearbook" as the data source, using two examples of the research and experimental development expenditures of my country's higher education institutions and the government consumption amount of Guangxi Zhuang Autonomous Region, the accuracy of the combined prediction model established is verified, and the prediction results of the single model are compared comparing.

## 2. ESTABLISHMENT OF COMBINED MODEL

### 2.1 Support Vector Machine

Support vector machine is a machine learning algorithm proposed by Vapnik et al. in 1995. This algorithm improves the generalization ability by minimizing the structural risk. Now it is mostly used to solve the classification and regression problems of small samples [12-13]. As a small sample learning method, support vector machine is widely used in

pattern recognition, function estimation, time series prediction and other problems, and has a good performance in handwritten data recognition, face image recognition, time series prediction and so on.

Support vector machine regression prediction is to seek the corresponding nonlinear mapping relationship f between the input data sequence and the output data sequence, for a given set of training sample sets {(Xi, yi), i=1, 2,..., l}, where Xi ∈ Rn, yi ∈ R. Xi is the input data of the i-th learning sample, yi is the output data of the i-th sample, and l is the number of samples. Through the training of the training samples, the fitting optimal functional relationship y = f(X) can be obtained. When the optimal functional relationship is nonlinear, the problem is a nonlinear regression problem.

For nonlinear problems, the support vector machine maps the input data sequence to a high-dimensional feature space through a nonlinear mapping φ, and performs linear regression in this feature space:

$$f(X) = W \cdot X + b,$$

Among them, W is the method vector of the hyperplane, and b is the bias. The problem of finding the optimal hyperplane can be transformed into the following quadratic convex programming problem:

$$\min \{\frac{1}{2}\|W\_\|^2\},$$

$$\text{s. t. tormented} \begin{cases} y\_i - W\_ \cdot \varphi(Xi) - b \le \varepsilon, \\ (W \cdot \varphi(Xi) + b - yi \le \varepsilon \ (i = 1, 2, \ldots, l). \end{cases}$$

The ε-insensitive loss function is used to measure the error between the observed value and the predicted value of the function, that is, when the error between the observed value and the predicted value of the function is less than a given ε, it is considered that there is no error in the fitting of the function to the sample point at this time of [14]. Transform the risk minimization problem of ε-insensitive loss function into an optimization problem in dual form, and the expression can be transformed into:

$$L\varepsilon(y, f(X)) = \begin{cases} 0, & |the\ y - f(X)| \le \varepsilon, \\ (|the\ y - f(X)| - \varepsilon, & others. \end{cases}$$

Introducing slack variables ξi, $\xi_i^*$, the optimization problem is transformed into a quadratic convex programming problem as follows:

$$\min \{\frac{1}{2}\|W\_\|^2 + C\sum_{i=1}^{l}(\xi i\_ + \xi' i)\},$$

$$\text{s. t. tormented} \begin{cases} y\_i - W\_ \cdot \varphi(Xi) - b \le \varepsilon + \xi i\_, \\ (W \cdot \varphi(Xi) + b - yi \le \varepsilon + \xi' i (\xi i\_, \xi' i \ge 0, \ i = 1, 2, \ldots, l), \end{cases}$$

Where C is the penalty factor.

Introduce the Lagrange function:

$$L(W, b, \alpha, \varepsilon, \xi i, \xi' i) = \frac{1}{2}\|W\|^2 + C\sum_{i=1}^{l}(\xi i + \xi' i) - \sum_{i=1}^{l}\alpha i(\varepsilon + \xi i - yi + W \cdot \varphi(Xi) + b)$$

$$- \sum_{i=1}^{l}\alpha' i(\varepsilon + \xi' i + yi - W \cdot \varphi(Xi) - b) - \sum_{i=1}^{l}(\eta i \xi i + \eta' i \xi' i),$$

Among them, αi, $\alpha_i^*$, ηi, $\eta_i^*$ are Lagrangian multipliers. Substitute the Lagrangian function into the equation of the optimization problem, and simplify to obtain the dual optimization formula of the original quadratic optimization problem:

$$\max J(\alpha) = \max \{ -\frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l}(\alpha i - \alpha' i)(\alpha j - \alpha j^*)(\varphi(Xi) \cdot \varphi(Xj))$$

$$-\varepsilon\sum_{i=1}^{l}(\alpha i + \alpha' i) + \sum_{i=1}^{l} yi(\alpha i - \alpha' i)\},$$

$$\text{s. t} \begin{cases} \sum_{i=1}^{l}(\alpha i - \alpha' i) = 0, \\ (0 \le \alpha i, \alpha' i \le C \ (i = 1, 2, \ldots, l). \end{cases}$$

The analytical expression for obtaining the optimal linear regression hyperplane in the feature space is:

$$f(X) = W \cdot X + b = \sum_{\text{support vector}}(\alpha i - \alpha' i)(\varphi(Xi) \cdot \varphi(Xj)) + b,$$

Among them, φ(Xi) ·φ(Xj) is the inner product of nonlinear mapping, which can be replaced by kernel function.

Currently commonly used kernel functions include the following [15-16] (where γ, r, d are kernel function parameters):

Linear kernel function: $K(X, Y) = X \cdot Y\_$

Polynomial kernel function: $K(X, Y) = (X \cdot Y + r)^d, \gamma > 0.$

Gaussian kernel function: $K(X, Y) = e^{-\gamma \sum_{i=1}^{no} (x_i - y_i)^2}, \gamma > 0.$

Sigmoid Kernel function: $K(X, Y) = \tanh(\gamma X \cdot Y + r).$

For the application field of support vector machine, an appropriate kernel function should be selected to replace the inner product. At this point, the expression can be transformed into:

$$f(X) = \sum_{\text{support vector}} (\alpha_i - \alpha_i') K(X_i \cdot X_j) + b.$$

### 2.2 Combined Forecasting Model of Improved Multidimensional Gray Model and Support Vector Machine

The idea of establishing the combined forecasting model based on the multidimensional gray model and support vector machine based on the innovation priority accumulation method is: first use the multidimensional gray model based on the innovation priority accumulation method to predict the data sequence, and then normalize the predicted value The preprocessed and normalized residual is used as the input data and output data of the support vector machine, and the predicted value of the residual is obtained through the support vector machine and compared with the predicted value of the multidimensional gray model based on the innovation priority accumulation method. Add to get the final prediction result. Specific steps are as follows:

(1) According to the system behavior characteristic data sequence and related factor sequence, a multi-dimensional gray model based on the innovation priority accumulation method is established, and the fitting value of the system behavior characteristic data sequence and the predicted value of the next few periods are obtained by calculating the time response of the model.

(2) The fitted value after normalization is used as the input data of the support vector machine, and the corresponding residual is preprocessed and normalized as the output data of the support vector machine, and a part of it is selected as the training sample, and the rest are test samples.

(3) Carry out learning and training through the support vector machine, select the appropriate kernel function, and determine the parameters of the support vector machine. Test whether the support vector machine can satisfy the KKT condition, if not, reselect the sample.

(4) The multi-dimensional gray model based on the innovation priority accumulation method is used to normalize the predicted values of the next few periods, which are used as the input data for the support vector machine prediction, and the corresponding output data are obtained.

(5) Pre-normalize and restore the predicted data output by the support vector machine, then the predicted value of the residual can be obtained.

(6) Add the predicted value of the residual to the predicted value of the multidimensional gray model based on the innovation priority accumulation method, and finally get the predicted value of the combined model.

## 3 EXAMPLE SIMULATION AND ANALYSIS

### 3.1 Example 1: Research and Experimental Development Expenditures of Universities

Taking the research and experimental development expenditures of my country's colleges and universities from 1998 to 2012 as the behavioral characteristic data series, and selecting the full-time equivalent of basic research personnel for research and experimental development and the full-time equivalent of applied researchers for research and experimental development in my country from 1998 to 2012. The time equivalent and the number of ordinary colleges and universities are used as the relevant factor sequence of the model to model, and the relevant data are shown in Table 1. The data source is "China Statistical Yearbook".

Establish the GM(0,4) model based on the innovation priority accumulation method, and obtain the time response formula of the model:

$$\hat{x1}^{(1)}(k) = 87.229921\, x2^{(1)}(k) - 12.542594\, x3^{(1)}(k) - 0.241735\, x4^{(1)}(k) - 121.492056,\ k = 1, 2, \ldots$$

**Table 1** 1998-2012 Research and Experimental Development Expenditures and Related Factors of Higher Education Institutions in China

| years | Higher School Research and Experimentation Development expenditure/100 million yuan | Research and Experimental Development Basis Full-time equivalent of researchers /10,000 person-years | Research and Experiment Development Application Full-time equivalent of researchers /10,000 person-years | Number of regular colleges/schools |
|---|---|---|---|---|
| 1 998 | 57.30 | 7.87 | 24.97 | 1 022 |
| 1 999 | 63.50 | 7.60 | 24.15 | 1 071 |
| 2 000 | 76.70 | 7.95 | 21.97 | 1 041 |

| | | | | |
|---|---|---|---|---|
| 2 001 | 102.40 | 7.88 | 22.60 | 1 225 |
| 2 002 | 130.50 | 8.40 | 24.73 | 1 396 |
| 2 003 | 162.30 | 8.97 | 26.03 | 1552 |
| 2 004 | 200.94 | 11.07 | 27.86 | 1 731 |
| 2 005 | 242.30 | 11.54 | 29.71 | 1 792 |
| 2 006 | 276.81 | 13.13 | 29.97 | 1 867 |
| 2 007 | 314.70 | 13.81 | 28.60 | 1 908 |
| 2 008 | 390.20 | 15.40 | 28.94 | 2263 |
| 2 009 | 468.20 | 16.46 | 31.53 | 2 305 |
| 2010 | 597.30 | 17.37 | 33.56 | 2358 |
| 2 011 | 688.84 | 19.32 | 35.28 | 2 409 |
| 2 012 | 780.56 | 21.22 | 38.38 | 2442 |

The fitting value of the research and experimental development expenditure of my country's higher education institutions from 1998 to 2012 is:

^

X 1 (0) = (4.77, 101.15, 166.27, 107.78, 85.09, 80.80, 197.75, 200.80, 318.11, 384.70, 433.31, 483.14, 524.24, 660.44, 779.32).

The residual sequence is:

r=(52.53, -37.65, -89.57, -5.38, 45.41, 81.50, 3.19, 41.50, -41.30, -70.00, -43.11, -14.94, 73.06, 28.40, 1.24).

The residual sequence is preprocessed as follows:

$$r'(k) = \frac{r(1) + r(2) + \ldots + r(k)}{k},$$

have to:

r'=(4.77, 52.96, 90.73, 94.99, 93.01, 90.98, 106.23, 118.05, 140.28, 164.72, 189.14, 213. 64, 237.53, 267.74, 301.84).

The fitted value after normalization processing is used as the input data of the support vector machine, and the corresponding preprocessed residual is normalized as the output data of the support vector machine, and the years 1998-2001 and 2003 are selected -The data from 2006, 2008 to 2011 are used as training samples, and the data from 2002, 2007, and 2012 are used as testing samples. The kernel function of support vector machine is Gaussian kernel function. After training and learning, the value of parameters is: $C = 17.0829$, $\gamma = 12.1926$, $\varepsilon = 0.001$.

The GM (0, 4) model based on the innovation priority accumulation method is used to study and test the colleges and universities in China from 2013 to 2015. The predicted value of the experimental development expenditure is normalized, and it is used as the input data of the support vector machine to obtain the corresponding output data. Denormalize the output data and restore the preprocessing, and compare it with the GM(0, 4) based on the innovation priority accumulation method The prediction values of the models are added to obtain the final prediction value of the combination model. Comparing the prediction value with the prediction result of the support vector machine and the GM (0, 4) model based on the innovation priority accumulation method, table 2 is obtained.

**Table 2** Predicted values and relative errors of the three models

| years | actual value /billion | Support vector machine | | GM (0, 4) model based on innovation priority accumulation method | | Combination prediction model in this paper | |
|---|---|---|---|---|---|---|---|
| | | Predicted value/100 million | relative error/% | Predicted value/billion yuan | Relative error/% | Predicted value/100 million | relative error/% |
| 1998 | 57.3 0 | 78.70 | 37.35 | 4.77 | 91.68 | 14.08 | 75.42 |
| 1999 | 63.5 0 | 73.77 | 16.17 | 101.15 | 59.28 | 106.79 | 68.18 |
| 2000 | 76.70 | 71.35 | 6.97 | 166.27 | 116.78 | 170.05 | 121.71 |
| 2 0 01 | 102.40 | 80.90 | 21.00 | 107.78 | 5.26 | 118.47 | 15.69 |
| 200 2 | 130.50 | 109.35 | 16.21 | 85.09 | 34.80 | 94.60 | 27.51 |
| 200 3 | 162.30 | 140.78 | 13.26 | 80.80 | 50.22 | 89.07 | 45.12 |
| 2 0 04 | 200.94 | 218.18 | 8.58 | 197.75 | 1.59 | 190.11 | 5.39 |
| 200 5 | 242.30 | 257.27 | 6.18 | 200.80 | 17.13 | 206.00 | 14.98 |
| 2 0 06 | 276.81 | 309.22 | 11.71 | 318.11 | 14.92 | 303.97 | 9.81 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 2007 | 314.70 | 314.07 | 0.20 | 384.70 | 22.24 | 375.85 | 19.43 |
| 2008 | 390.20 | 411.69 | 5.51 | 433.31 | 11.05 | 425.77 | 9.12 |
| 2009 | 468.20 | 489.88 | 4.63 | 483.14 | 3.19 | 473.70 | 1.18 |
| 2010 | 597.30 | 562.61 | 5.81 | 524.24 | 12.23 | 515.18 | 13.75 |
| 2011 | 688.84 | 667.61 | 3.08 | 660.44 | 4.12 | 627.30 | 8.93 |
| 2012 | 780.56 | 803.40 | 2.93 | 779.32 | 0.16 | 747.47 | 4.24 |
| 2013 | 856.70 | 877.13 | 2.38 | 848.62 | 0.94 | 826.66 | 3.51 |
| 2014 | 898.10 | 953.39 | 6.16 | 931.56 | 3.73 | 904.69 | 0.73 |
| 2015 | 998.59 | 1077.39 | 7.89 | 1049.99 | 5.15 | 1012.94 | 1.44 |
| 2013-2015 _ Forecast average error/% | | | 5.48 | | 3.27 | | 1.89 |

From Table 2 that the combination of the GM (0, 4) model based on the innovation priority accumulation method established in this paper and the support vector machine The prediction effect of the prediction model is better than the two single models of support vector machine and GM (0, 4) model based on the innovation priority accumulation method In other words, the forecasting accuracy of the combined forecasting model established is higher than that of the single forecasting model.

### 3. 2 Example 2: Government Consumption Amount

Taking the government consumption amount of Guangxi Zhuang Autonomous Region from 1997 to 2012 as the behavioral characteristic data sequence of the model, and The general budget revenue and residents' consumption of Guangxi Zhuang Autonomous Region from 1997 to 2012 were selected as the model. Modeling is carried out on the sequence of related factors, and the relevant data are shown in Table 3.

**Table 3** Government consumption and related factors in Guangxi Zhuang Autonomous Region from 1997 to 2012

| years | Amount of government consumption/100 million yuan | General budget revenue of local finance/100 million yuan | Amount of household consumption/100 million yuan |
|---|---|---|---|
| 1997 | 319.11 | 99.16 | 936.65 |
| 1998 | 356.04 | 119.67 | 952.63 |
| 1999 | 364.07 | 133.56 | 978.81 |
| 2000 | 357.32 | 147.05 | 1091.00 |
| 2001 | 436.00 | 178.67 | 1159.34 |
| 2002 | 448.80 | 186.73 | 1250.93 |
| 2003 | 499.30 | 203.66 | 1360.25 |
| 2004 | 559.16 | 237.77 | 1537.99 |
| 2005 | 655.05 | 283.04 | 1808.47 |
| 2006 | 772.60 | 342.58 | 2006.99 |
| 2007 | 917.58 | 418.83 | 2425.83 |
| 2008 | 932.35 | 518.42 | 2947.82 |
| 2009 | 1006.03 | 620.99 | 3369.86 |
| 2010 | 1196.39 | 771.99 | 3745.84 |
| 2011 | 1353.29 | 947.72 | 4248.30 |
| 2012 | 1612.19 | 1166.06 | 4905.76 |

Establish the GM (1, 3) model of the innovation priority accumulation method, and obtain the time response formula of the model:

$$\hat{x}^{(1)}(k+1) = \{319.11 - \frac{1}{1.195883}[-0.881864 x^{(1)}(k+1) + 0.583026 x^{(1)}(k+1)]\} e^{-1.195883 k} + \frac{1}{1.1\ 58\ 8}[-0.881864 x^{(1)}(k+1) + 0.583026 x^{(1)}(k+1)], k = 1, 2, \cdots$$

After normalization processing is used as the input data of the support vector machine, and the corresponding preprocessing residual is carried out After normalization processing, it was used as the output data of the support vector machine, and the data from 1998-2001, 2003-2006, and 2008-2011 were selected as training samples, and the data from 2002, 2007, and 2012 were selected as training samples.

The kernel function of the support vector machine is Gaussian kernel function, after training and learning, the value of the parameters is: $C = 10.8616$, $\gamma = 7.8063, \varepsilon = 0.001$.

The GM (1, 3) model based on the new information priority accumulation method is used to analyze the government consumption in Guangxi Zhuang Autonomous Region from 2013 to 2015. The predicted value of the fee amount is normalized as the input data of the support vector machine prediction, and the input data of the support vector machine prediction The data is denormalized and restored to obtain the predicted value of the residual, which is compared with the new information priority accumulation method. The predicted value of the GM (1, 3) model is added to obtain the predicted value of the combined model. Combine the predicted value with support vector machine and innovation-based optimization The prediction results of the GM (1, 3) model of the cumulative method are compared. It can be seen that the combination of the GM (1, 3) model based on the innovation priority accumulation method established in this paper and the support vector machine The prediction effect of the prediction model is obviously better than that of the support vector machine and the GM (1, 3) model based on the innovation priority accumulation method. The prediction accuracy of the model is higher than that of the single prediction model.

## 4 CONCLUSION

In this paper, a combined forecasting model based on the multidimensional gray model and support vector machine based on the innovation priority accumulation method is established, and through The GM(0, 4) model based on the innovation priority accumulation method is established for the research and experimental development expenditures of colleges and universities in China and the expenditure The combined forecasting model of support vector machine and the combined forecasting model of GM (1, 3) model based on innovation priority accumulation method and support vector machine are established for the government consumption of Guangxi Zhuang Autonomous Region, and the predicted results of the combined forecasting model established are with a single model for comparison. The results show that the forecasting accuracy of the combined forecasting model established in this paper is improved compared with the single model. So this article The combination prediction model of the multidimensional gray model and support vector machine based on the innovation priority accumulation method is an effective prediction model. The model further improves the prediction accuracy of the multidimensional gray model.

## COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

## REFERENCES

[1] Wu Huaan, Zeng Bo, Peng You, Zhou Meng. Prediction of urban population density based on multidimensional gray system model. Statistics and Information Forum, 2018, 33(08): 60-67.

[2] Luo Yiyong, Zhang Hao, Zhang Liting. Research on Multidimensional Gray Deformation Prediction Model Based on Genetic Support Vector Machine. Journal of Zhejiang University of Technology, 2010, 3 8 (01): 79-83.

[3] Awad M, Khanna R. Efficient learning Machines. Apress, 2015: 203-224.

[4] Zhang Haoran, Han Zhengzhi, Li Changgang. Support Vector Machines. Computer Science, 2002, 29(12): 135-137.

[5] Meyer D, Leisch F, Hornik K. Benchmarking Support vector Machines. Austria： Vienna University of Economics and Business Administration, 2002.

[6] Yang Guangxi. Application of combined forecasting method in economic forecasting. Statistics and Forecasting, 1998(6): 17-20.

[7] Tang Xiaowo. Optimal combination forecasting method and its application. Mathematical Statistics and Management, 1992(1): 31- 35.

[8] Chen Huayou. Combination forecasting method validity theory and its application. Beijing: Science Press, 2008: 1-7.

[9] Zeng Bo, Liu Sifeng, Fang Zhigeng. Gray combination forecasting model and its application. Chinese Management Science, 2009, 17(5): 150-155.

[10] Yuan Quan, Zeng Xiangyan. The GM(0, N) model based on innovation priority accumulation method and its application. Statistics and Decision Making, 2018, 34(12): 79-81.

[11] Yuan Quan, Zeng Xiangyan. The GM (1, N) model based on innovation priority accumulation method and its application. Journal of Guilin University of Electronic Technology, 2017, 37(04): 332-336.
[12] Gu Yaxiang, Ding Shifei. Research progress of support vector machine. Computer Science, 2011, 38(2): 14-17.
[13] Xu Jianhua, Zhang Xue Work, Li Yanda. The new development of support vector machine. Control and Decision Making, 2004, 19(5): 481-484.
[14] Sun Deshan. Research on Support Vector Machine Classification and Regression Method. Changsha: Central South University, 2004.
[15] Yang Junyan, Zhang Youyun, Zhu Yongsheng. ε- insensitive loss function support vector machine classification performance research. Journal of Xi'an Jiaotong University, 2007, 41(11): 1315-1320.
[16] Feng Guohe. SVM classification kernel function and parameter selection comparison. Computer Engineering and Application, 2011, 47(3): 123-124.