

# Grey Forecasting Model Based on Interpolation Optimization in Malignant Tumor

Li Xiaofei<sup>1</sup>, Hu Renfang<sup>2,a</sup>

<sup>1</sup>Engineering and Technology College of Yangtze University, 434020 Jingzhou Hubei, China

<sup>2</sup>Jingzhou City Second People's Hospital, 434000 Jingzhou Hubei, China

**Abstract.** People pay more and more attention to health with the development of society. In addition, owing to the uncertain pathological cause, more and more kinds of malignant tumor seriously affect our health. In order to reduce the economic loss and protect the safety of our life, it is necessary to forecast the occurrence of malignant tumor and make prevention countermeasures. Conventional grey model has usually been used for forecasting in disease prediction, but it has significant limitations. Improved grey forecasting model in malignant tumor is established based on interpolation optimization. The result of the analysis shows that the accuracy of improved grey model is significant higher than conventional model, so the improved grey model can be used for forecasting in malignant tumor. It can extend to other disease prediction.

**Keywords.** grey forecasting model; interpolation optimization; malignant tumor; prevention countermeasures

## 1 Introduction

In recent years, malignant tumor has been recognized as one of the most serious diseases for human health around the world. The amount of malignant tumor patients increases in growing numbers with rapid development of urban modernization. The amount of malignant tumor patients takes a large proportion in total amount of people in China. It will easily bring about enormous personnel's casualty, which leads to disastrous economic loss, and bad social influence. So it is extremely important for lowering personnel's casualty and economic loss to master malignant tumor occurrence regularity and make prevention countermeasures as soon as possible. Therefore, in order to take the way of sustainable development, it has the vital practical significance to provide the statistical data and the amount of malignant tumor patients forecast.

The forecasting model of malignant tumor is a typical model in the grey system theory which was pioneered by Deng <sup>[1]</sup>. The principle of grey prediction is to find regularities from discrete data, and its essence is to create new data series and establish new prediction model by dealing with raw data series and weakening its randomness. There are many complicated influencing factors for malignant tumor. Although one malignant tumor patient has randomness and contingency other than regularity, many malignant tumor patients during one period have certain probability regularity, which fluctuate within a certain range.

---

<sup>a</sup> Corresponding author: lxfei0828@163.com

Some factors are certain but others uncertain in one patient. So the forecasting model can be successfully applied in actual predication .Many models along with co-integrated techniques, regression analysis, a univariate trigonometric have been successfully applied to a number of studies to analyze and forecast. GM (1,1) model is the most commonly approach in prediction. It has the advantages of few data demand, convenient calculation and used widely consequently.

However, it also has some limitations. Therefore, in recent years, research on improved GM (1,1) model has attracted many attention. Liu <sup>[9]</sup> uses the conventional GM (1,1) model to forecast data. Tang <sup>[10]</sup> uses Lagrange interpolation formula to reconstruct the background value. Li <sup>[11]</sup> uses the Newton-Cores formula to reconstruct the background value. Wang <sup>[12]</sup> uses Gauss-Chebyshev formula to reconstruct the background value. It improves prediction accuracy of models to some extent.

The aim of this paper is to focus on forecasting for malignant tumor using the grey predictive modeling based on interpolation optimization and make prevention countermeasures as soon as possible.

## 2 Conventional GM (1,1) Forecasting Model

The original data of malignant tumor patients in certain area are:

$$x^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)\}.$$

Applying accumulated generating operation, it can get that:

$$x^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)\},$$

Where  $x^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i)$  ( $k = 1, 2, \dots, n$ ) .  $x^{(1)}(k)$  is called accumulated generating operation of  $x^{(0)}(k)$  denoted as 1-AGO. The first order linear ordinary differential equation expressed as:

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b. \quad (1)$$

Which is called whitened differential equation of GM (1,1), of which the difference form is:

$$x^{(0)}(k) + az^{(1)}(k) = b,$$

Where  $a, b$  are parameters to be identified.  $a$  is called developing coefficient, and  $b$  is called grey input. Solve it using least square method and obtain:

$$[a, b]^T = (B^T B^{-1}) B^T Y_n,$$

Where

$$Y_n = [x^{(0)}(2), x^{(0)}(3), \dots, x^{(0)}(n)], B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix}. \quad (2)$$

In (2), the background value is formulated as

$$z^{(1)}(k+1) = \frac{1}{2} [x^{(1)}(k) + x^{(1)}(k+1)] (k = 1, 2, \dots, n-1).$$

The discrete solution of (1) is:

$$\hat{x}^{(1)}(k+1) = \left( x^{(0)}(1) - \frac{b}{a} \right) \cdot e^{-ak} + \frac{b}{a}.$$

The reduction value is:

$$x^{(0)}(k+1) = x^{(1)}(k+1) - x^{(1)}(k) = (1 - e^a)(x^{(0)}(1) - \frac{b}{a}) \cdot e^{-ak}$$

for  $k = 1, 2, \dots, n-1$ .

During the above GM (1,1) modeling process, it can be found that developing coefficient  $a$  and grey input  $b$  have great influences on simulation and prediction accuracy of the model. However, it shows that the values of  $a$  and  $b$  depend on the construction of the background value. Therefore, to construct a new background value form is the key to optimize and modify conventional GM (1,1) model.

### 3 Improvement GM (1,1) Model Based On Interpolation Optimization

In the conventional model, using consecutive neighbors means

$$z^{(1)}(k+1) = \frac{1}{2} [x^{(1)}(k) + x^{(1)}(k+1)]$$

trapezoidal area with exponential curve side of  $x^{(1)}(t)$  is replaced by straight-edge trapezoidal area. The shortcoming is that, as the exponential increases, data sequence varies greatly and the prediction result will have big error ( $\Delta S$ ). That will influence the applicability of the model to some extent.

In order to overcome this shortcoming, we first present a new construction method of the background value combinative interpolation optimization method, which uses numerical approximation idea combining interpolation algorithm to resolve this problem and reduce error.

**Step1:** Divide the interval  $[k, k+1]$  into three equal intervals:

$$(k + \frac{1}{3}, x^{(1)}(k + \frac{1}{3})), (k + \frac{2}{3}, x^{(1)}(k + \frac{2}{3})), (k + 1, x^{(1)}(k + 1)).$$

$$x^{(1)}(k + \frac{i}{3}), i = 1, 2.$$

**Step2:** Solve

Establish quadric Lagrange interpolation polynomial  $P_2(t)$  which is submit to

$$P_2(t) \approx x^{(1)}(t), t \in [k, k+1],$$

then

$$x^{(1)}(k + \frac{i}{3}) \approx P_2(k + \frac{i}{3}) (i = 1, 2),$$

where

$$P_2(t) = x^{(1)}(k) \frac{(t-k-1)(t-k-2)}{(-1) \cdot (-2)} + x^{(1)}(k+1) \frac{(t-k)(t-k-2)}{1 \cdot (-1)} + x^{(1)}(k+2) \frac{(t-k)(t-k-1)}{1 \cdot 2}.$$

Then we get :

$$x^{(1)}(k + \frac{1}{3}) \approx P_2(k + \frac{1}{3}) = \frac{5}{9} x^{(1)}(k) + \frac{5}{9} x^{(1)}(k+1) - \frac{1}{9} x^{(1)}(k+2);$$

$$x^{(1)}(k + \frac{2}{3}) \approx P_2(k + \frac{2}{3}) = \frac{2}{9} x^{(1)}(k) + \frac{8}{9} x^{(1)}(k+1) - \frac{1}{9} x^{(1)}(k+2).$$

**Step3:** Obtain the piecewise interpolation function in the interval  $[k, k+1]$ .

$$\text{If } k \leq t \leq k + \frac{1}{3},$$

$$\text{then } S_k(t) = 3[x^{(1)}(k + \frac{1}{3}) - x^{(1)}(k)]t + x^{(1)}(k) - 3k[x^{(1)}(k + \frac{1}{3}) - x^{(1)}(k)].$$

$$\text{If } k + \frac{1}{3} \leq t \leq k + \frac{2}{3}, \text{ then } S_k(t) = 3[x^{(1)}(k + \frac{2}{3}) - x^{(1)}(k + \frac{1}{3})]t + x^{(1)}(k + \frac{1}{3})$$

$$- 3(k + \frac{1}{3})[x^{(1)}(k + \frac{2}{3}) - x^{(1)}(k + \frac{1}{3})].$$

$$\text{If } k + \frac{2}{3} \leq t \leq k + 1,$$

$$\text{then } S_k(t) = 3[x^{(1)}(k + 1) - x^{(1)}(k + \frac{2}{3})]t + x^{(1)}(k + \frac{2}{3})$$

$$- 3(k + \frac{2}{3})[x^{(1)}(k + 1) - x^{(1)}(k + \frac{2}{3})].$$

**Step4:** Calculate the numerical integration of the background value.

$$z^{(1)}(k+1) = \int_k^{k+1} x^{(1)}(t) dt$$

$$\int_k^{k+1} x^{(1)}(t) dt \approx \int_k^{k+1} S_k(t) dt = \int_k^{k+\frac{1}{3}} S_k(t) dt$$

$$+ \int_{k+\frac{1}{3}}^{k+\frac{2}{3}} S_k(t) dt + \int_{k+\frac{2}{3}}^{k+1} S_k(t) dt = \frac{1}{6} x^{(1)}(k+1)$$

$$+ \frac{1}{3} x^{(1)}(k + \frac{2}{3}) + \frac{1}{3} x^{(1)}(k + \frac{1}{3}) + \frac{1}{6} x^{(1)}(k).$$

(3)

Substitute the conclusion of Step2 into (3) and get the optimized background value.

$$z^{(1)}(k+1) = \int_k^{k+1} x^{(1)}(t) dt \approx \int_k^{k+1} S_k(t) dt$$

$$= \frac{23}{54} x^{(1)}(k) + \frac{35}{54} x^{(1)}(k+1) - \frac{2}{27} x^{(1)}(k+2).$$

(4)

(4) is the novel background value of GM(1,1) obtained from the combinative interpolation optimization idea.

## 4 Application Of Forecasting Model In Malignant Tumor

The main goal of this study is to predict the amount of malignant tumor patients of Jingzhou city in Hubei province in China using the improved GM(1,1). According to the international classification standard of diseases, we extract the number of malignant tumor patients of Jingzhou city from 2002 to 2007 in table 1.

**Table 1.** The actual data of the malignant tumor patients from 2002 to 2007

Year	2002	2003	2004	2005	2006	2007
Actual amount	13567	15947	16546	16637	17821	19079

The forecasting results are as shown in table 2 compared with the actual data from 2008 to 2011.

**Table 2.** The forecasting data of the malignant tumor patients from 2008 to 2011

Year	Actual amount	Forecasting data(conventional model)	Random Error(%)	Forecasting data(improvement model)	Random Error(%)
2008	18812	19615	4.27	18867	2.92
2009	19245	20505	6.55	19528	1.47
2010	20685	21436	3.63	20211	2.29
2011	19797	22410	13.19	20919	5.67
Average RE(%)			11.53		3.08

We find that the prediction accuracy of this model is 3.08% from the table 2. So the precision is very high and the model is correct and applicable to predict. The data of the amount from 2012 to 2016 is obtained in table 3.

**Table 3.** The forecasting data of the malignant tumor patients from 2012 to 2016

Year	2012	2013	2014	2015	2016
Forecasting data	21652	22410	23195	24007	24848

## 5 Discussion

The forecasting result of grey forecasting model cannot be completely precise because there are many complex reasons external or internal. As time goes on, because of the deterioration of the environment and increasing in population, we will face more and more serious problems for human health. One side, we should pay attention to food safety. There are many shortcomings and problems in Chinese law on food safety system. There are many serious illegal activities in Chinese food industry at present. These will bring very large living space in unsafe food production. Food insecurity leads to the disease. In order to exclude the foods in danger in time, the government should establish a complete food safety guarantee system, formulate corresponding measures to protect our health and carry out the security inspection regularly. Medical institutions will improve the detection techniques to ensure food safety to provide technical support. The relevant departments should strength supervision and stop the illegal acts at the same time. We should control the unsafe food to a minimum. On the other side, we will enhance tumor knowledge education. Let people master the knowledge of tumor in common and arrange them to do medical examination regularly. We should

strengthen the physical exercise to prevent the occurrence of disease especially the malignant tumor at the same time. Timely detection, and timely treatment.

## 6 Acknowledgment

This work is supported by the twelve five year plan project in Hubei province (GN: 2013B308) and the foundation for research development of engineering and technology college, Yangtze University (GN: 13J0802). The authors would like to thank the referees for their helpful suggestions which essentially improved the presentation of this paper.

## References

1. Deng Julong. Grey modeling resource theory and GM(1,1,bk). *The Journal of Grey System*, 2005,17(3): 201-206.
2. Diyar Akay, Mehmet Atak. Grey prediction with rolling mechanism for electricity demand forecasting of Turkey. *Energy*, 2007,(32):1670-1675.
3. Wang Tengjun, Lai Bailian. Research on landslide prediction based on GM(1,1) data fusion algorithm. *Bulletin of Surveying and Mapping*, 2012,(5):63-65.
4. Deng Qi, Wang Qi. Application of GM(1,1) in forecasting quantity of industrial solid waste. *Environmental Science & Technology*, 2012,35(6) : 180-183.
5. Yu Wei,Wu Ziyin. Prediction on the coastline change based on the grey model GM(1,1) and matlab. *Marine Science Bulletin*, 2012,31(4):404-408.
6. Gorucu FB, Geris Pu, Gumrah F. Artificial neural network modeling for forecasting gas consumption. *Energy Sources*,2004,26:299-307.
7. Wang H Q, D Hu. The construction and application of combination forecasting model in Chinese energy consumption system. *Statistics and Decision*,2008,25:64–66.
8. Zhou P, B Wang, K L Poh. A trigonometric grey prediction approach to forecasting electricity demand. *Energy*, 2006,31: 2839-2847.
9. Liu Sifeng, Guo Tianbang, Dang Yaoguo. *Grey system theory and application*, Beijing, Press of Science, 2008.
10. Tang Wangmei, Xiang Changhe, The improvement of forecasting method in GM(1,1) model based on quadratic interpolation, *Chinese Journal of Management Science*, 2006,6:109-112.
11. Li Junfeng, Dai Wenzhan. A new approach of background value-building and its application based on data interpolation and Newton-Cores formula, *Systems Engineering-Theory & Practice*, 2004,10:122-126.
12. Wang Xiaojia. Forecasting modeling and analysis of power engineering in China based on Gauss-Chebyshev formula. *Systems Engineering Procedia*,2012,5:131-136.