Research Journal of Applied Sciences, Engineering and Technology 5(14): 3830-3834, 2013
ISSN: 2040-7459; e-ISSN: 2040-7467

Submitted: October 17, 2012 Accepted: December 10, 2012 Published: April 20, 2013

Prediction of Chinese per Capita Grain Yield Base on Residual Modification GM (1, 1) Model

Yang Yang
Southwest University, Beibei, Chongqing 400715, China

Abstract: To build effective grain yield prediction system and predict its trend scientifically, this study, on the basis of statistics, prognostics and agricultural economics, explains and functions grey system theory. As a new method, grey system still has many shortages. On the basis of comparison in correlative prediction, we propose GM (1, 1) grey prediction method by modifying ends to improve predictive precisions. Besides, combining with historic data during 2000-2009, predict, summary and propose the research future. Research indicates, whether theoretic basis or practice, grey model is more useful and convenient. It predicts the yield in future 5 years, the increasing speed will decrease. The increasing yield is 5-6 kilos per person, less than 8-10 kilos per person during 2003-2009. Surely, grain industry includes many son industries, such as rice, corn and wheat. The biggest son industry should be found to give different financial support, in order to eliminate errors. The innovation is to solve time responding function and incandesce equation of end residual sequence of GM (1, 1) model, to eliminate error. Besides, analyze practical examples to indicate its value in economic prediction and provide references for relative areas.

Keywords: Grain yield, GM (1, 1), model, prediction, residual modification

INTRODUCTION

Economic prediction refers to a process that summary historic data and phenomenon find the internal principles, frames and methods and propose its trends. There are many methods now, qualitatively and quantitatively. Those two ways are complementary and tested mutually. The widespread qualitative ways are brainstorm and Delphi method. However, quantitative methods are linear measuring equation, markov, space state equation and grey theory, etc. All of them have their own advantages and disadvantages. Even some scholars propose to integrate those predication methods in variable weight combination model, to reduce predictive error and improve precision. For example, Zuo-Fang *et al*. (2010) modified results of grey prediction in markov method; Zhang *et al*. (2009) proposed a grey prediction model base on heritage algorithm to regulate development coefficients and generate functions. Consider heritage algorithm coefficients as grey buffer operators. Nowadays, global popular grain prediction methods have meteorological yield prediction, remote sensor prediction and statistical dynamic growth simulation, etc. They are predicted 2 months before. The error is 5%- 10% of yield. It is because it is difficult to predict in remote sensor technology before it grows to some degrees. But currently, we hardly make right prediction before more than a month. Seen from global predictions, grain yield prediction is a hot topic. Its importance is obvious. The methods are many. However, from the researches above, there are two shortages:

- Ignoring grey quality of grain production generates big error.
- Some models pay much attention to data, resulting in complex acquisition, processing and analysis and great costs, such as remote sensor technology needs a lot of techniques and money (Deng, 1987).

Grain industry includes many son industries like rice, corn and wheat. The biggest son industry should be found to make a different financial support. Currently, many scholars analyze its industrial structures. Lai (2009) fined the effects of corn fluctuation on grain yield are great and then rice and wheat are by building error modifying model on grain yield. By quantitatively analyzing national task, ensure limited stimulation to lead fastest development of grain industry. Writer says, at first, numbers of data are required by numerical economic predication methods (such as meter age regression), then due to not enough statistic papers or error, resulting in ineffective model. Markov chain stresses ineffectiveness, which is current data including all previous information, next phenomenon only relates to current data and transformation probability. It needs powerful effective market theory. But they don't exist in practice. So the imagine without practical support will lead to errors. Secondly, different predication models can improve precisions. But models are too many to use all of them to integrate coefficients. That will be too complex and the coefficients are effective only in a certain stage. Wang (2011) thinks grey model is a model of differentiate equation form, by generated and changed by discrete random numbers. It is a process from grey to

white. This study's purpose is to build an effective grain yield prediction system and predict its trend scientifically. All in all, the innovation of this study is, grey system can indicates data motion principles and needs fewer samples. It transforms the impulse sequences caused by randomly interruption from grey to white. However, any model has error, which is random disturbance term. So this study solves time responding function and incandesce equation of end residual sequence of GM (1, 1) prediction model, to eliminate error and demonstrate its economic values.

GM (1, 1) MODEL WITH RESIDUAL MODIFICATION

Nowadays, many researches are about grey prediction. This study, ignoring previous models, stresses economic significances and variances' functions. GM (1, 1) model is single-variance grey prediction model under one-order condition. There are basic and primary forms.

Primary form is $x_0(k)+ax_1(k) = b$. $x_0(k)$ is a primary sequence. $x_1(k)$ is a first accumulation sequence. So a and b can be solved by simple OLS. a is development coefficient, which is the effects of a unit change of accumulation sequence on primary sequence. Imagine $x_0(k)$ is a simple sequence increasing timely, such as $x_0(k)$ = Ck, c is a constant. Then a and b are linear related. Basic form is to calculate neighboring average sequence $z_1(k)$ of $x_1(k)$. Then replace $x_1(k)$ to form $x_0(k)+az_1(k) = b$. Its purpose is to control trends expansion. In other words, buffer operates.
The regression matrix can be solved:

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

and

$$Y = \begin{bmatrix} x_0(2) \\ x_0(3) \\ \cdots\cdots \\ x_0(n) \end{bmatrix}, B = \begin{bmatrix} -z_1(2) & 1 \\ -z_1(3) & 1 \\ \cdots\cdots & \cdots\cdots \\ -z_1(n) & 1 \end{bmatrix}$$

Then minimal two multiple estimating parameter sequence of $x_0(k)+az_1(k) = b$ of GM (1, 1) meets the requirements.

The time responding function sequence of $x_0(k)+az_1(k) = b$ is:

$$x_1(k+1) = (x_0(1) - \frac{b}{a})e^{-ak} + \frac{b}{a}, k = 1,2,\ldots\ldots n$$

Get reduction sequence:

$$x_0(k+1) = \alpha_1 \hat{x}_1(k+1) = \hat{x}_1(k+1) - \hat{x}_1(k)$$
$$= (1-e^a)(x_0(1) - \frac{b}{a})e^{-ak}$$

X_0 has error, the time responding formula of GM (1,1) of:

$$\varepsilon^1 = (\varepsilon^1(k_0), \varepsilon^1(k_0+1), \ldots\ldots, \varepsilon^1(n))$$

is:

$$\hat{\varepsilon}(k+1) = (\varepsilon_0(k_0) - \frac{b_\varepsilon}{a_\varepsilon})e^{-a_\varepsilon(k-k_0)} + \frac{b_\varepsilon}{a_\varepsilon}, k \geq k_0$$

The predictive sequence after K0 is:

$$\hat{\varepsilon} = (\hat{\varepsilon}(k_0), \hat{\varepsilon}(k_0+1), \ldots\ldots, \hat{\varepsilon}(n))$$

Of which

$$\hat{\varepsilon}_0(k+1) = -a_\varepsilon(\varepsilon_0(k_0) - \frac{b_\varepsilon}{a_\varepsilon})e^{-a_\varepsilon(k-k_0)}, k \geq k_0$$

If $\hat{x}_0(k) = \hat{x}_1(k) - \hat{x}_1(k-1)$, the prediction equation after modification iS:

$$x_0(k+1) = (1-e^a)(x_0(1) - \frac{b}{a})e^{-ak} \pm a_\varepsilon(\varepsilon_0(k_0) - \frac{b_\varepsilon}{a_\varepsilon})e^{-a_\varepsilon(k-k_0)}$$

The basic idea of residual modification GM (1, 1) is: since grey process can get values by internal principles of primary sequences, part of error is essential to be true. In the same way, processing by time formulas, overcome the incomplete information got by residual in previous methods. It should be noted, K0 is the start value of cutting error term, which doesn't mean the last half section must be got from n residual sequence. It can be selected by comparing prediction errors in different points. The following example sets K0 as 2, reduces error in the whole. Of course we can set K0 according to real situations.

EXAMPLE ANALYSIS

Statistic analysis of chinese grain yield: According to grey model and residual modification above, select 2000-2009 Chinese per capita output as object. The data comes from 2009 China Statistical Yearbook. Function GM (1, 1) prediction and solve time responding function of 1-AGO sequence and modify the prediction model. As a traditionally agricultural country, China never ignores grain safety. But its yield reached summit since 1998, due to less land area, lots of agricultural labors' loss and too low technical basis, the yield in 2003 reduced, which caused much terror. But series of three agriculture policies enhanced the policy's inclination and investment, overcame the natural disaster and realized a big harvest. But it should be known, although the yield increases year by year, the increasing magnitude is small. The increasing magnitude in 2009

Table 1: Total output and per capita output of our country's grain

Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Yield (million tons)	46217	45264	45706	43070	46947	48402	49746	50150	52850	53082
per capita yield (kg)	364.66	354.66	355.82	333.29	361.16	370.17	378.46	379.63	398.12	387.69

Table 2: Prove of non-negative smoothing sequence and quasi-exponential principle

Years	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
P (k)		0.972	0.495	0.310	0.344	0.188	0.177	0.151	0.137	0.118
P (k+1)/ P(k)			0.509	0.627	1.111	0.546	0.941	0.852	0.911	0.856
O(k)		0.9726	1.0033	0.9367	1.0836	1.0249	1.0224	1.0031	1.0487	0.9738

base on 2000 is 14.85%, the mean increasing ratio is less than 1.5%, which is not only less than that of agricultural investment but also that of farmer's salary. So the improvement of grain yield is a systematic engineering, which can be successful with all people's endeavor. The current policies should be improved. Too much population is our basic situation. The increasing population doesn't mean the increasing labors. So model development can not be added as labor in new classic economy. In reality, per capita resource is reduced. Some scholar proved the grain yield returns to the start line in 24 years after reform and opening-up, which is shown as Table 1 shows. Per capita output depends on total output and population, which indicates the increasing of grain yield and the functioning of family planning.

GM (1, 1) prediction of grain yield: From the analysis of basic model above we can see that the robotic system may be decomposed into basic analysis model and this is the basis for problem solving. If we want to increase the system's self-learning ability and intelligence, traditional passive modules have not met the need of intelligence. Therefore agent is introduced to take place of module for solving the generalization problem. According to this idea, the basic unit of robotic architecture is considered as an agent, which decouples among hardware and software, function and modules. Corresponding, the architecture of agent is also divided into three typical architecture including deliberative, reactive and hybrid. Like this, every agent may design independently. And these agents are independent and of self-awareness, so the strategy of solving practical problems is performed through cooperation and competition among them.

Grey prediction is a process for historic data sequence changing from grey to white. According to theoretic narration in first section, predict the per capita output in the following steps:

Accumulation sequence 1-AGO generation. According to behavior sequence:

X (0) = (X (0) (1), X (0) (2), ….. X (0) (10)) = (364. 66, 354.66, 355.82, 333.29, 361.16, 370.17, 378.46, 379.63, 387.69)

Get 1-AGO sequence:

X (1) = (X (1) (1), X (1) (2), …X (1)(10) = (364.66, 719.32, 1075.14, 1048.43, 1969.59, 2139.76, 2518.22, 2897.85, 3295.97 and 3683.66)

The basic requirement is X (0) is non-negative quasi-smoothing sequence. X (1) has quasi-exponent quality. Prove in the following. At first, function X (0) smoothing ratio to P (k) calculation, whose result is shown as Table 2. P (k) in 2001 in Table 2 is 0.972. Other annual fees are less than 0.5. All P (k+1) /P (k) in eight time points can be tested. So prove is done. In the same way, O (k) can be solved, which is also shown as Table 1. Fluctuate area is [0.9367, 1.0836] and extreme value <0.5. So the data sequence is quasi-exponential.

Neighboring mean Z (1) generation and prediction model estimation. The study above has already proposed the condition of grey modeling. The following generates neighboring mean sequence: Z (1) = (364. 66,541, 897.23, 1061.78, 1509, 2054, 2328, 2708, 3096 and 3489).

Due to X (t)+aZ(t) = b, the estimated matrix form is:

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

$$\left[\begin{pmatrix} -541 & 1 \\ -897.23 & 1 \\ -1061.78 & 1 \\ -1509 & 1 \\ -2054 & 1 \\ -2328 & 1 \\ -2708 & 1 \\ -3096 & 1 \\ -3489 & 1 \end{pmatrix}^T \begin{pmatrix} -541 & 1 \\ -897.23 & 1 \\ -1061.78 & 1 \\ -1509 & 1 \\ -2054 & 1 \\ -2328 & 1 \\ -2708 & 1 \\ -3096 & 1 \\ -3489 & 1 \end{pmatrix}\right]^{-1} \begin{pmatrix} -541 & 1 \\ -897.23 & 1 \\ -1061.78 & 1 \\ -1509 & 1 \\ -2054 & 1 \\ -2328 & 1 \\ -2708 & 1 \\ -3096 & 1 \\ -3489 & 1 \end{pmatrix}^T \begin{pmatrix} 354.66 \\ 355.82 \\ 333.29 \\ 361.16 \\ 370.17 \\ 378.46 \\ 379.63 \\ 398.12 \\ 389.67 \end{pmatrix} = \begin{pmatrix} 341.023 \\ 0.0149 \end{pmatrix}$$

The development coefficient is 0.0149. The abs less than 2 indicates GM model is stable. And according to [-2/ (n+1), 2/ (n+1)], it has high precision.

The time function can be solved: $x_1(k+1)$ = 23245.164304 exp (0.014905t)-22880.504304. Get fitting sequence after calculating in DSP software, which is shown as Table 3. The third column is prediction error and the fourth column is relative error. The fitting is bad. The sum of residual square C is 0.7046, mean relative error P = 0.5556. The prediction of future 5 stages is X (t+k) = 399.16069, 405.15455, 411.23841, 417.41363 和 423.68157, K = 1, 2, 3, 4, 5.

Table 3: Prediction results

No	Observation value	Fitting value	Error	Relative error%
X(2)	366.7185	350.4491	5.8034	1.6284
X(3)	392.5754	378.3863	13.4717	3.8891
X(4)	334.3219	372.2009	-29.1804	-8.2674
X(5)	373.4395	392.7555	-3.9701	-1.1484
X(6)	387.9657	383.0924	16.4022	4.4342
X(7)	391.7883	401.631	-13.6033	-3.7429
X(8)	386.4998	387.0498	3.2671	0.8571
X(9)	410.0636	424.6317	11.4087	2.8925
X(10)	393.1177	406.6262	-5.944	-1.4557

Table 4: Prediction results after residual modification

No	Observation value	Fitting value	Error	Relative error%
X(2)	378.7769	359.3588	0.2828	0.0791
X(3)	380.4453	379.8809	16.3703	4.4535
X(4)	344.3117	369.6997	-10.3082	-3.0957
X(5)	388.6082	375.1327	-13.3267	-3.5760
X(6)	415.7879	395.1288	4.7926	1.2463
X(7)	391.0614	395.4974	-7.1097	-2.0206
X(8)	414.9869	400.7756	-2.7164	-0.7385
X(9)	425.1921	419.5420	9.5420	2.4147
X(10)	414.0529	385.0517	4.3184	1.1580

RESIDUAL SEQUENCE MODIFICATION ODEL

According to current situation, such results are bad. Get their abs, then:

$|\varepsilon_0|$ =（5.8034, 13.4717, 29.180, 3.970, 16.40, 13.60, 3.267, 11.4, 5.94） Build GM (1, 1) model and get ε_0 1-AGO sequence and ε_1 time responding formula:

x (t+1) = -175.067530 exp (-0.101709t) + 186.653550 of which a = 0.101709, b = 18.984327. Its difference is ε_0 (k+1) = (-0.101709*175.06753 exp (-0.101709t) =17.80 exp (-0.101709t)

The prediction model after residual modification:

X_0(k+1) =（1-exp(-0.0149) -17.80 exp (-0.101709t) = x(t+1) = -175.067530 exp(-0.101709t) + 186.653550

Table 4 shows fitting value, residual sequence and relative mean error. Residual square summary increases by 0.8889. It is good. The prediction of future 5 years: 394.35298, 399.69133, 405.18307, 410.82343 and 416.60824 kilo per capita. Two problems exist: one is due to obvious increasing of yield, residual expansion in simulation is the results of trends expansion. The prediction after modifying has better convergence. The results show the fluctuation is smaller.

Grey theory holds that, when two fluctuating sequences and trends are close, their relation is grey related. Figure 1 shows the prediction sequence after modifying is more next to real sequence. The reason for its difference from traditional prediction methods is: grey theory considers behavior sequence (real value) isn't true sequence. The task of grey process is to get true sequence from behavior sequence and eliminate disturbance. But other methods are opposite. Most methods consider observation as true sequence.

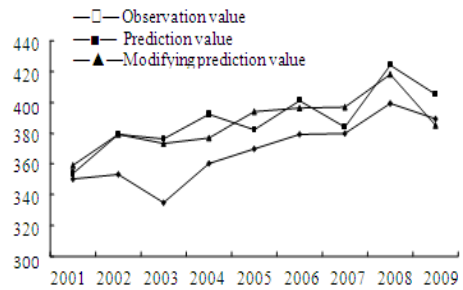

Fig. 1: Comparison of prediction effects

CONCLUSION

The following conclusions can be got after discussion and prove base on residual modification on GM (1, 1) prediction model:

- No matter theoretic basis or practice, grey model is more useful and convenient without loss of logical basis. The social activities are complex. A model can't include all factors and due to common lack of data, buffering operator makes primary data processing more regulative, controllable and systematical. It can regulate primary data according to real economic phenomenon and trends. It discards the error, which holds observation data is true data.
- Grey system residual modification overcomes the shortage that effective information can't be retrieved from random disturbance term. The more important is it won't bring the fake regression and different variance in regression method. And it can be functioned in many times, until the information

is perfectly retrieved. The result proves residual modification improves the precision greatly, which is through relative test. Besides, buffer operator makes the combination with grey algorithm and other algorithms come true.

- The prediction of grain yield in future 5 years, the increasing ratio of per capita will reduced. It will be 5-6 kilo/people, less than 8-10 kilo/people during 2003-2009. Certainly, the result base on current development. If policies and technologies are well improved, the situation will change. Especially the measures like grain safety, building grain base, will result in change of systematic behaviors.

REFERENCES

Deng, J., 1987. Essential Methods of Grey System. Huazhong University of Science and Technology Press, Huhan.

Lai, H., 2009. Analysis of fluctuation and structure of Chinese grain yield. Agric. Technol. Econ., vol. 5.

Wang, M., 2011. Contrast of predictions of three grey systems. Statist. Decision, Vol. 8.

Zhang, D., H. Zhu, W. Li and J. Zhang, 2009. Application of grey model base on heritage algorithm on railway capacity. Statist. Decision, vol. 24.

Zuo-Fang, Y., L. Xing-Tu and F. Yang, 2010. Application of markov modification grey model on grain yield prediction. Geogragh. Sci., Vol. 3.