The current issue and full text archive of this journal is available at
**www.emeraldinsight.com/0368-492X.htm**

# China's population projections based on GM(1,1) metabolic model

Lu Caimei
*School of Management, Shanxi University, Taiyuan, People's Republic of China*
Hao Yonghong
*Key Laboratory for Water Environment and Resources, Tianjin Normal University, Tianjin, People's Republic of China, and*
Wang Xuemeng
*Institute for Integrated Survey of Agriculture Resources of Shanxi, Taiyuan, People's Republic of China*

## Abstract

**Purpose** – The purpose of this paper is to apply grey system theory to population system and project China's population.

**Design/methodology/approach** – The paper applies the GM(1,1) model to China's population projections. Two key aspects of the method are crucial for obtaining best accuracy of prediction. They are the choice of the length for the original data to be used in the model and the adoption of the GM(1,1) metabolic model in prediction. The former determines what initial data to be used while the latter describes an iteration process on how to proceed to predict.

**Findings** – The results show that in 2015 China's population will reach 1.37 billion and in 2050 it will be between 1.42 and 1.48 billion, which is in accordance with the latest projections from the UN. The findings show the GM(1,1) metabolic model is an effective mathematical means in population projections.

**Research limitations/implications** – The paper suggests that GM(1,1) metabolic model can provide an effective simulation model for complicated systems with uncertainty and can be used in many fields.

**Practical implications** – The paper provides useful advice for the department of population.

**Originality/value** – Most population projections have been based on assumptions about fertility, mortality, and migration. The paper considers the population system as a grey system and introduces the GM(1,1) metabolic model to population projections.

**Keywords** China, Cybernetics, Systems theory, Population, Modelling

**Paper type** Research paper

## 1. Introduction

The population of a country has a major impact on economic and social development of the nation. Future population relates to the sustainable development of the country. Population projections provide a criterion for city planners, economists, public agencies, environmentalists and social scientists to prepare for what is to come, so the research on population projections has become an important issue in many fields. China has a huge population, which has been one of the primary factors restricting China's economic development, so the research is of great practical significance.

Kybernetes
Vol. 38 Nos 3/4, 2009
pp. 417-425

0368-492X
DOI 10.1108/03684920910944119

At present, most population projections have been based on assumptions about fertility, mortality, and migration (Huang, 2004; Song *et al.*, 1981; Wang and Ma, 2003). However, population development is affected by many factors such as traditional ideology, government policies, and natural disasters. So it must be treated as an uncertain process. Different approaches have been developed to characterize this uncertainty, which can be grouped into two main classes: scenarios and probabilistic projections. Scenarios projections assume higher and lower vital rates in dealing with the uncertainty, so this approach has a major drawback. In the high variants, fertility is assumed to be high in every year of the forecast period. Similarly, when fertility is low in one year, it is 100 per cent certain that it will be lower in the following years too (Nico and Arve, 2002). Probabilistic projections determine the vital rates and the probabilities associated with them by capturing potential structural change based on expert opinion, statistical analysis, and the analysis of errors in past projections (Brain *et al.*, 2001). However, assumptions of vital rates inevitably produce errors. As the errors are accumulated and magnified during the prediction process, population projections become more and more inaccurate.

In this study, we introduce a population projection model based on grey system theory advanced by Deng (1988), Lin and Liu (2006, 2000) and Lin *et al.* (2004). Grey system theory is an effective mathematical means of resolving problems containing uncertainty and indetermination and is used in most research (Xiong *et al.*, 2003; Peng and Cai, 2005; Song, 2005; Zhang *et al.*, 2003). We believe that the population development process can be treated as an inherently grey system. Population is regarded as an integrated result of all the factors of the population system and it implicitly contains all the information of the population development process. Without any parameter assumption, we set up a system of dynamic differential equations for population development. We use a GM(1,1) model for China's population projections.

In what follows, we first describe briefly what a GM(1,1) model is, and then how it can be applied to population projections. As a case study, the method is then applied to China's population projections.

## 2. GM(1,1) model

The grey model GM(1,1) means a grey model with first order differential time series equations in one variable. To set up the grey model, a set of disordered raw data are first transformed into a new set by means of the accumulated generating operation (AGO).

Let $X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \ldots, x^{(0)}(n))$ be a given time sequence representing the raw data. Let $X^{(1)} = (x^{(1)}(1), x^{(1)}(2), \ldots, x^{(1)}(n))$ be the AGO sequence of $X^{(0)}$ with:

$$x^{(1)}(t) = \sum_{i=1}^{t} x^{(0)}(i), \quad t = 1, 2, \ldots, n. \tag{1}$$

Notice that the equation:

$$x^{(0)}(t) + ax^{(1)}(t) = b \tag{2}$$

gives us an equation of the grey differential type. If we replace $x^{(1)}(t)$ in equation (2) by the new background value:

$$z^{(1)}(t) = 0.5x^{(1)}(t) + 0.5x^{(1)}(t-1), t = 1, 2, \ldots, n \tag{3}$$

the mean of the entries of $X^{(1)}$, then $z^{(1)}(t)$ and the components $x^{(1)}(t)$ and $x^{(1)}(t-1)$ of the grey derivative satisfy the crucial arithmetic horizontal mapping relation and consequently equation (2) becomes the governing grey differential equation in our GM(1,1) model:

$$x^{(0)}(t) + az^{(1)}(t) = b \tag{4}$$

If we let:

$$Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \cdots \\ x^{(0)}(n) \end{bmatrix}, \quad B = \begin{bmatrix} -z^{(1)}(2), & 1 \\ -z^{(1)}(3), & 1 \\ \cdots & \cdots \\ -z^{(1)}(n), & 1 \end{bmatrix} \tag{5}$$

then the least square estimates of $a$ and $b$ from equation (4) are given by:

$$(a, b)^T = (B^T B)^{-1} B^T Y \tag{6}$$

Using these estimates of $a$ and $b$, now our whitenization equations of the grey system equation (4) are given by:

$$\frac{dx^{(1)}(t)}{dt} + ax^{(1)}(t) = b. \tag{7}$$

The solution of equation (7) can be readily found as:

$$x^{(1)}(t) = \left(x^{(1)}(0) - \frac{b}{a}\right)e^{-at} + \frac{b}{a}. \tag{8}$$

Then the approximate time response sequences of the GM(1,1) grey differential system equation (4) are set up as:

$$\hat{x}^{(1)}(t+1) = \left(x^{(1)}(0) - \frac{b}{a}\right)e^{-at} + \frac{b}{a}; \quad t = 1, 2, \ldots, n \tag{9}$$

We set $x^{(1)}(0) = x^{(0)}(1)$, then:

$$\hat{x}^{(1)}(t+1) = \left(x^{(0)}(1) - \frac{b}{a}\right)e^{-at} + \frac{b}{a}; \quad t = 1, 2, \ldots, n. \tag{10}$$

Finally, by reversing AGO, we can express the restored values of $\hat{x}^{(0)}(t+1)$ as:

$$\hat{x}^{(0)}(t+1) = \hat{x}^{(1)}(t+1) - \hat{x}^{(1)}(t), \quad t = 1, 2, \ldots, n. \tag{11}$$

In a GM(1,1) model, the parameter $a$ is known as development coefficient, which describes the development states of $\hat{X}^{(1)}$ and $\hat{X}^{(0)}$, and indicates the dynamic characteristics of the system; while the parameter $b$ is called the grey action quantity, which reflects the effects of the environment that surrounds the system. In general, the variables that act upon the system should be external or predetermined.

However, GM(1,1) is a single sequence model, which makes use of only the system's behavioral sequence without considering any external acting sequences. The grey action quantity $b$ in GM(1,1) can be derived from the background values that reflect changes contained in the data whose exact connotation is grey. Therefore, this model does not need the detailed information on the system and only uses the system's behavioral variants. It provides an effective way in building simulation models for complicated systems with deficient data such as population system.

## 3. GM(1,1) metabolic model

In the development process of a grey system, there always exists some stochastic interference or some driving forces affecting the system's development besides the temporal dimension. Thereby, by means of a GM(1,1) model, the further into the future we predict, the less accurate the results will be (Figure 1) (Wang, 1992). In fact, through a GM(1,1) prediction model, high accuracy can be obtained only for the first or the second data values that are immediately after the original value $x^{(0)}(n)$. In order to attain high accuracy of population projections, we adopt GM(1,1) metabolic model. We first use the selected $X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \ldots, x^{(0)}(n))$ to predict only one value of the population, namely $x^{(0)}(n+1)$. Then we set up a new GM(1,1) model to predict $x^{(0)}(n+2)$. The new model is based on the following new sequence, $X^{(0)} = (x^{(0)}(2), \ldots, x^{(0)}(n), x^{(0)}(n+1))$ which is obtained by inserting $x^{(0)}(n+1)$ as the last entry and deleting $x^{(0)}(1)$ from the original $X^{(0)}$. One can continue the same procedure to get the following predictions as many times as one wants.

## 4. Application of GM(1,1) metabolic model to China's population projections

During the period 1952-2005, China's population development has experienced great changes. Although there have been fluctuations existing in the population development process, generally speaking, natural growth rate has declined from 20 to 6 per cent (National Bureau of Statistics, 1953-2006). Since China implemented Family Planning Policy, the excessive population growth has been effectively controlled. Based on the assumption that the population system is considered as a grey system, we apply the metabolic model to China's population projections.

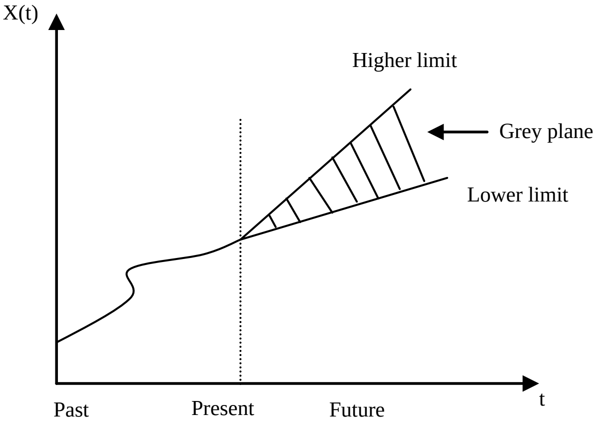

**Figure 1.**
The schematic diagram of the grey plane

In a grey system, present information not only expresses the system's current states and behaviors, but also underlines the future-developing trend. The law of system development the past information implies has been transmitted to the present information. In a GM(1,1) model, the development coefficient *a* and the grey action quantity *b* both depend on the length *n* of the original data sequence. Different parameter values for *a* and *b* give us different GM(1,1) models and different results of prediction. But it is not necessarily true that the longer the original data sequence is, the better the prediction is. So it is crucial to select the time series of appropriate lengths as the original data sequence. An appropriate time series should be consistent with the future-developing tendency of the system. Then, how long should the length of the best time series $x^{(0)}(n)$ be? This issue cannot be resolved easily. One solution is to get the appropriate dimension of the original data sequence through empirical test on the accuracy of the GM(1,1) that is applied to the known data. Therefore, we test the GM(1,1) metabolic model with different dimensions or lengths of the original data sequences to see which one gives the best possible results.

We can classify the scales of the original data sequence into two groups:

1. a long and medium dimensional sequence group with the scale of more than ten years ($n \geq 10$); and
2. a short dimensional sequence group with the scale of less than ten years ($n < 10$).

By examining errors in the two groups, respectively, we choose the best model that is suitable for predicting China's population. In Tables I and II, we have listed some of our computational results for the two groups. From these results we can see that:

- the errors of GM(1,1) metabolic models with long and medium dimensional sequence are much larger than the models with short dimensional sequences; and
- the GM(1,1) metabolic models with 5 and 6 dimensional sequences fit well with reality and their projections provide us with a lower bound and a higher bound for the true population.

Therefore, we choose GM(1,1) metabolic models with 5- and 6-dimensional sequences to project the population of China and provide a range of future population. In other words, we use the population sequences of 2001-2005 (i.e. 5-dimension) and 2000-2005 (i.e. 6-dimension) as the original data for our GM(1,1) metabolic models. By using the models, we forecast the China's population from 2006 to 2050. The projections are listed in Table III. Some calculations for the corresponding growth rate and net growth of the population are included in it.

## 5. Discussion and conclusion

From our results of computation, we can foresee the following points:

- In 2007, the population of China will reach 1.32 billion; in 2015 it will be about 1.37 billion; and in 2050 it will be between 1.42 and 1.48 billion.
- In 2015, the annual growth rate of population in China will drop to 4 per cent and the annual net growth of population will be about 5.5 million. By 2035, the annual growth rate of population will drop to 2 per cent and the annual net growth of population will be about 2 million.

**Table I.**
Errors of GM(1,1) metabolic model with long and medium dimensional sequence

| | | Forecast values and relative errors | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 2002 | | 2003 | | 2004 | | 2005 | |
| Dimensions of original sequence | Period of original sequence | Value ($10^4$) | Error (per cent) | Value ($10^4$) | Error (per cent) | Value ($10^4$) | Error (per cent) | Value ($10^4$) | Error (per cent) |
| 50 | 1952-2001 | 136,523 | − 6.28 | 137,691 | − 6.55 | 138,797 | − 6.78 | 139,833 | − 6.54 |
| 40 | 1962-2001 | 134,726 | − 4.88 | 135,624 | − 4.95 | 136,434 | − 4.96 | 137,198 | − 4.93 |
| 30 | 1972-2001 | 131,928 | − 2.71 | 132,986 | − 2.91 | 133,982 | − 3.07 | 134,914 | − 3.18 |
| 20 | 1982-2001 | 130,945 | − 1.94 | 131,788 | − 1.98 | 132,516 | − 1.94 | 133,156 | − 1.84 |
| 10 | 1992-2001 | 129,213 | − 0.59 | 129,966 | − 0.57 | 130,979 | − 0.76 | 131,996 | − 0.95 |

**Table II.** Errors of GM(1,1) metabolic model with short dimensional sequence

| | | Forecast values and relative errors | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 2002 | | 2003 | | 2004 | | 2005 | |
| Dimensions of original sequence | Period of original sequence | Value ($10^4$) | Error (per cent) | Value ($10^4$) | Error (per cent) | Value ($10^4$) | Error (per cent) | Value ($10^4$) | Error (per cent) |
| 4 | 1998-2001 | 128,276 | 0.14 | 128,986 | 0.19 | 129,636 | 0.27 | 130,344 | 0.32 |
| 5 | 1997-2001 | 128,392 | 0.05 | 129,114 | 0.09 | 129,829 | 0.12 | 130,574 | 0.14 |
| 6 | 1996-2001 | 128,570 | − 0.09 | 129,371 | − 0.11 | 130,121 | − 0.10 | 130,873 | − 0.09 |
| 7 | 1995-2001 | 128,752 | − 0.23 | 129,560 | − 0.26 | 130,232 | − 0.19 | 131,022 | − 0.20 |
| 8 | 1994-2001 | 128,861 | − 0.32 | 129,694 | − 0.36 | 130,352 | − 0.28 | 130,998 | − 0.19 |
| 9 | 1993-2001 | 129,013 | − 0.44 | 129,834 | − 0.47 | 130,494 | − 0.39 | 1311,20 | − 0.28 |

**Table III.**
Population projections of China from 2006 to 2050

| | Population ($10^4$) | | Growth rate (per cent/year) | | Net growth ($10^4$) | |
|---|---|---|---|---|---|---|
| Year | 5-dimension | 6-dimension | 5-dimension | 6-dimension | 5-dimension | 6-dimension |
| 2006 | 131,414 | 131,488 | 5.03 | 5.60 | 658 | 732 |
| 2007 | 132,076 | 132,188 | 5.04 | 5.32 | 662 | 700 |
| 2008 | 132,702 | 132,872 | 4.74 | 5.17 | 626 | 684 |
| 2009 | 133,292 | 133,540 | 4.45 | 5.03 | 590 | 668 |
| 2010 | 133,852 | 134,186 | 4.20 | 4.84 | 560 | 646 |
| 2011 | 134,386 | 134,810 | 3.99 | 4.65 | 534 | 624 |
| 2012 | 134,890 | 135,416 | 3.75 | 4.50 | 504 | 606 |
| 2013 | 135,372 | 136,008 | 3.57 | 4.37 | 482 | 592 |
| 2014 | 135,828 | 136,576 | 3.37 | 4.18 | 456 | 568 |
| 2015 | 136,264 | 137,128 | 3.21 | 4.04 | 436 | 552 |
| 2016 | 136,672 | 137,662 | 2.99 | 3.89 | 408 | 534 |
| 2017 | 137,064 | 138,180 | 2.87 | 3.76 | 392 | 518 |
| 2018 | 137,432 | 138,684 | 2.68 | 3.65 | 368 | 504 |
| 2019 | 137,780 | 139,172 | 2.53 | 3.52 | 348 | 488 |
| 2020 | 138,116 | 139,640 | 2.44 | 3.36 | 336 | 468 |
| 2021 | 138,436 | 140,096 | 2.32 | 3.27 | 320 | 456 |
| 2022 | 138,736 | 140,540 | 2.17 | 3.17 | 300 | 444 |
| 2023 | 139,020 | 140,968 | 2.05 | 3.05 | 284 | 428 |
| 2024 | 139,296 | 141,384 | 1.99 | 2.95 | 276 | 416 |
| 2025 | 139,552 | 141,784 | 1.84 | 2.83 | 256 | 400 |
| 2026 | 139,800 | 142,176 | 1.78 | 2.76 | 248 | 392 |
| 2027 | 140,032 | 142,552 | 1.66 | 2.64 | 232 | 376 |
| 2028 | 140,256 | 142,920 | 1.60 | 2.58 | 224 | 368 |
| 2029 | 140,464 | 143,276 | 1.48 | 2.49 | 208 | 356 |
| 2030 | 140,664 | 143,616 | 1.42 | 2.37 | 200 | 340 |
| 2035 | 141,512 | 145,176 | 1.13 | 2.02 | 160 | 292 |
| 2040 | 142,160 | 146,520 | 0.90 | 1.81 | 128 | 264 |
| 2045 | 142,556 | 147,640 | 0.67 | 1.41 | 86 | 208 |
| 2050 | 142,853 | 148,608 | 0.45 | 1.29 | 64 | 192 |

- An increase of some 200 million people can be expected between 1995 and 2025 in China; and an additional 50 million between 2025 and 2050.

According to the medium variant of projections from the Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat (the 2006 revision), China's population will reach 1.38 billion for the year 2015 and 1.41 billion for 2050 (Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat, 2007). Obviously, our findings are in accordance with the projections from the UN.

From all the above, we can conclude the following observations and comments.

The GM(1,1) model provides an effective simulation model for complicated systems with uncertainty such as population system. Even though the GM(1,1) model is an exponential growth equation, the metabolic process in our method automatically gives rise to the convergence properties for the growth rate and the net growth of population. These properties are consistent with the real population development process. In contrast, in order to ensure these properties, other methods need to pre-impose some conditions on vital rates. The population development process is changing with time.

To take different lengths of population sequences as original sequence can lead to different results. However, the best results can be obtained by the appropriate dimensional sequence of the original data, which can be found by means of a model test of the GM(1,1) or an empirical analysis of the accuracy of the GM(1,1).

## References

Brain, C., Deborah, B., Melanie, B. and Markos, E. (2001), "A guide to population projections", *Demographic Research*, No. 4, pp. 203-88.

Deng, J.L. (1988), *Grey System*, China Ocean Press, Beijing (in Chinese).

Huang, R.Q. (2004), "Thinking on the methods of population projection", *Population Research*, No. 1, pp. 35-7 (in Chinese).

Lin, Y. and Liu, S.F. (2000), "A systemic analysis with data (I)", *International Journal of General Systems*, Vol. 29, pp. 989-99.

Lin, Y. and Liu, S.F. (2006), "Solving problems with incomplete information: a grey systems approach", *Advances in Imaging and Electron Physics*, Vol. 141, Elsevier, Oxford, pp. 77-174.

Lin, Y., Chen, M.-Y. and Liu, S.-F. (2004), "Theory of grey systems: capturing uncertainties of grey information", *Kybernetes: The International Journal of Systems & Cybernetics*, Vol. 33 No. 2, pp. 196-218.

National Bureau of Statistics (1953-2006), *China Statistical Yearbook*, China Statistics Press, Beijing (in Chinese).

Nico, K. and Arve, H. (2002), "Why population forecasts should be probabilistic – illustrated by the case of Norway", *Demographic Research*, No. 6, pp. 410-54.

Peng, Z.W. and Cai, S.Q. (2005), "Grey system forecast and analysis of petroleum futures prices", *Advances in Systems Science and Applications*, Vol. 5 No. 1, pp. 118-24.

Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat (2007), *World Population Prospects: The 2006 Revision. Highlights*, United Nations, New York, NY.

Song, J., Tian, X., Yu, J. and Li, G. (1981), *Population Forecasts and Population Control*, People's Press, Beijing (in Chinese).

Song, Z.M. (2005), "The analysis on grey element series", *Advances in Systems Science and Applications*, Vol. 5 No. 2, pp. 234-8.

Wang, X.M. (1992), *An Introduction to Grey System*, Chengdu Sci. & Tech. University Press, Chengdu, pp. 40-1 (in Chinese).

Wang, Y. and Ma, B. (2003), "The mathematical model of China population after 1980 and projection", *Journal of Peking University*, Vol. 39, pp. 28-33 (Science edition, supplement).

Xiong, H.J., Chen, D.J. and Chen, M.Y. (2003), "Study on gray method of data mining", *Advances in Systems Science and Applications*, Vol. 3 No. 2, pp. 309-13.

Zhang, S.H., Zhang, J.L. and Chen, M.Y. (2003), "A new preprocessing method of raw data and the modified GM (1,1) model", *Advances in Systems Science and Applications*, Vol. 3 No. 2, pp. 182-7.

## Corresponding author

Lu Caimei can be contacted at: lucaimei@sxu.edu.cn