Prediction of Population Development Trend in Huizhou Based on Grey Discrete Incremental Model

Jing Wang[1] and Shuhong Peng[2](✉)

[1] School of Arts and Design, Huizhou University, Huizhou 516007, China
wangjing@hzu.edu.cn
[2] School of Computer Science and Engineering, Huizhou University, Huizhou 516007, China
pengsh@hzu.edu.cn

Abstract. The population issue is related to regional economic development, and population changes will affect the formulation of the regional basic policies, the arrangement of labor and employment. Huizhou is located in the Pearl River Delta Economic Belt of Guangdong Province in China, with a high population density. In recent years, the population has gradually increased. Based on the grey incremental model GM (1,1) method, this paper uses the population data from 2000 to 2022 to construct a population prediction model for Huizhou and predict the population in the next 15 years. The prediction results show that the urban population of Huizhou will still increase year by year from 2023 to 2037, which will have an impact on the sustainable development of the city. Finally, the countermeasures for population development in Huizhou are proposed for regional development.

Keywords: Grey incremental model · population quantity · population prediction

1 Introduction

China is a total population of 1411.78 million according to the seventh population census in 2021. According to economists' estimates, with a population growth of 1%, it is necessary to increase by about 3% to maintain the living standard of the current population, which still exerts tremendous pressure on China's society and economy [1–3]. In China's modernization process, it is necessary to achieve coordinated and sustainable development of population and economy, society, resources, and environment, further control the population quantity, improve the population quality, improve the population structure, guide the population distribution, and stabilize low fertility levels.

The grey system was founded by Chinese scholar Professor Deng Julong in, and its research object is "small sample" and "poor information" uncertain systems with "some information known and some information unknown". It understands and understands the real world through the generation and development of "part" of known information, achieving a correct understanding and description of the system's operational behavior and evolution laws [4–8]. In 1982, Professor Deng, a Chinese scholar, established

N. Akhtar et al. (Eds.): PMIS 2023, AHIS 8, pp. 1351–1358, 2023.
https://doi.org/10.2991/978-94-6463-200-2_145

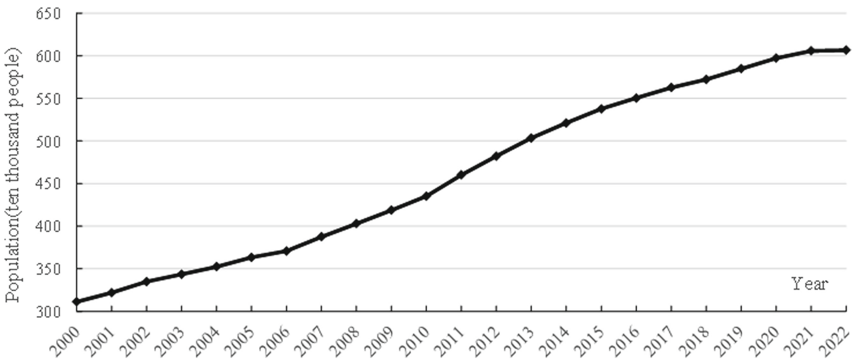

Fig. 1. Population Change Trend of Huizhou from 2000 to 2022

the grey system theory, which is a new method for studying uncertain problems with little data and poor information [9]. The grey system theory takes "small samples" and "poor information" uncertain systems with "partially known and partially unknown information" as the research object. It mainly extracts valuable information through the generation and development of partially known information, and realizes the correct description and effective monitoring of system operation behavior and evolution rules. The grey system model has no special requirements and limitations on experimental observation data, so its application fields are very broad [7–13].

The study area is Huizhou, a prefecture level city in Guangdong Province, located on the southeast bank of the Guangdong Hong Kong Macao Greater Bay Area, in the northeast of the Pearl River Delta, bordering Dongguan, Shenzhen, with a total area of about 11000 square kilometers. With the continuous advancement of industrialization and urbanization in China, the building density has increased, and the population has gradually increased.

The main topic of this paper is to construct a population prediction model for Huizhou based on the grey incremental model GM (1,1) method and the population data from 2000 to 2022 (Fig. 1), to predict the population of Huizhou in the next 15 years. Therefore, scientifically and accurately predicting the future population development of a region has extremely important practical significance for strategic decision-making of regional economic and social development [14–16].

2 Grey Discrete Incremental Model

2.1 Modeling

To study a system, it is generally necessary to first establish a mathematical model of the system, and then conduct a specific quantitative study of the overall function and coordination function of the system, as well as the correlation and dynamic relationship between various factors in the system. This research must be led by qualitative analysis, with a close combination of quantitative and qualitative analysis. The establishment of a system model generally involves five steps: thought development, factor analysis, quantification, dynamism, and optimization, so it is called five-step modeling [17–20].

1. The first step is to develop ideas and form concepts. Through qualitative analysis and research, the direction, objectives, approaches, and measures of research are clarified, and the results are expressed in accurate and concise language. This is the language model.
2. The second step is to analyze the factors in the language model and the relationship between them, and identify the antecedents and consequences that affect the development of things.
3. The third step is to conduct quantitative research on the causal relationship of each link, and initially obtain a low-level approximate quantitative relationship, which is called a quantitative model.
4. The fourth step is to further collect the input and output data of each link, and use the obtained data sequence to establish a dynamic GM model, that is, a dynamic model. Dynamic model is a high-level quantitative model that more profoundly reveals the quantitative relationship or conversion rules between input and output, and is the foundation of system analysis and optimization.
5. The fifth step is to conduct systematic research and analysis on the dynamic model. Through adjusting the structure, mechanism, and parameters, the system is reorganized to optimize the configuration and improve the dynamic quality of the system. The resulting model is called an optimization model.

The whole process of five-step modeling is the process of establishing five models in five different stages:

Language model → Network model → Quantitative model → Dynamic model → Optimization model.

During the modeling process, the model gradually tends to be improved [21, 22].

2.2 GM (1,1) Model

(1) Cumulative generation, set X^0 as the original series

$$X^{(0)} = [x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(\mathrm{n})] \tag{1}$$

Accumulate X^0 as the original series once to generate a sequence:

$$X^{(1)} = [x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(\mathrm{n})] \tag{2}$$

where:

$$x^{(1)}(k) = \sum\nolimits_{i=1}^{k} x^{(0)}(i), k = 1, 2, \cdots, n \tag{3}$$

(2) Modeling, let X^1 be the construction background value sequence

$$Z^{(1)} = [z^{(1)}(2), z^{(1)}(3), \cdots, z^{(1)}(n)] \tag{4}$$

where:

$$z^{(1)}(k) = \alpha x^{(1)}(k-1) + (1-\alpha)x^{(1)}(k), k = 2, 3, \cdots, n \quad (5)$$

Take: $\alpha = 0.5$ Establish shadow equation:

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b \quad (6)$$

This is called the original formation of the GM (1,1) model.

By discretizing the above equation and differential variation, the GM (1,1) grey differential equation is obtained as follows:

$$x^{(0)}(k) + az^{(1)}(k) = b \quad (7)$$

It is called the basic form of the GM (1,1) model. Where a and b are undetermined coefficients, called development coefficients and gray action quantities, respectively. The effective interval of a is (−2,2).

(3) Solve parameters. The least square method can be used to obtain:

$$\hat{a} = (a, b)^T = (B^T B)^{-1} \cdot B^T \cdot Y_n \quad (8)$$

Including:

$$B = \begin{pmatrix} -1/2(x^{(1)}(1) + x^{(1)}(2)), & 1 \\ -1/2(x^{(1)}(2) + x^{(1)}(3)), & 1 \\ \cdots & \cdots \\ -1/2(x^{(1)}(n-1) + x^{(1)}(n)), & 1 \end{pmatrix} \quad (9)$$

$$Y_n = [x^{(0)}(2), x^{(0)}(3), \cdots, x^{(0)}(\mathrm{n})] \quad (10)$$

In actual gray modeling, not all of the original sequence data of the system may be used for modeling. When modeling with different dimension length sequences, the values of the parameters a and b obtained are different, and thus the prediction effects of the models are different. They form a prediction gray interval. The development coefficient a reflects the development trend of the system. According to gray theory, when a is negative, the greater its absolute value, the faster the system develops, and vice versa. The gray action quantity b is mined from the background value, which reflects the relationship between data changes, and its exact connotation is gray. The grey action quantity is the concrete embodiment of connotation extension, and its existence is the watershed between grey modeling and general input output modeling. It is also an important symbol to distinguish between the grey system perspective and the grey box perspective.

(4) Establish prediction formula

$$\begin{cases} \hat{x}^{(1)}(k+1) = (x^{(0)}(1) - \frac{b}{a})e^{-ak} + \frac{b}{a} \\ \hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k) \end{cases} \quad (11)$$

(5) Model verification.

In order to ensure that the established grey model has a high accuracy and is applied to prediction practice, the following steps are generally required for testing:

1) Find x^0 (k) Residual error e (k), relative error Δ k, and average relative error Δ_k of:

$$e(k) = x^{(0)}(k) - \hat{x}^{(0)}(k) \tag{12}$$

$$\Delta_k = \left| \frac{e(k)}{x^{(0)}(k)} \right| \times 100\% \tag{13}$$

$$\overline{\Delta} = \frac{1}{n} \sum\nolimits_{k=1}^{n} \Delta_k \tag{14}$$

2) Find the average value x of the original data, Average residual value $\bar{e}$:

$$\bar{x} = \frac{1}{n} \sum\nolimits_{k=1}^{n} x^{(0)}(k) \tag{15}$$

$$\bar{e} = \frac{1}{n-1} \sum\nolimits_{k=2}^{n} e^{(0)}(k) \tag{16}$$

3) Calculate the original data variance S12, the ratio of residual variance S22 to mean square deviation C, and the small error probability p,

$$s_1^2 = \frac{1}{n} \sum\nolimits_{k=1}^{n} [x^{(0)}(k) - \bar{x}]^2 \tag{17}$$

$$s_2^2 = \frac{1}{n-1} \sum_{k=2}^{n} [e^{(0)}(k) - \bar{e}]^2 \tag{18}$$

$$C = s_2/s_1 \tag{19}$$

$$p = P\{e^{(0)}(k) - \bar{e} \mid < 0.6745 s_1\} \tag{20}$$

4) Find x ^ 0 (k) and x^^ Absolute correlation degree of 0 (k) γ.

Generally, the smaller the values of e (k), Δk, and C, γ, The higher the P value, the better the accuracy of the model.

3 Results

3.1 Grey Incremental Model Prediction

The data is sourced from the "Huizhou Statistical Yearbook" from 2010 to 2021. In actual gray modeling, the original sequence data of the system may not be all used for modeling. The values of the parameters obtained by modeling with different dimensions or length series are different, so the prediction effect of the model is also different. Here, we select the long series one year (2000–2022) and the medium long series one year (2015–2022) to simulate and predict the population of Huizhou, respectively, the best prediction sequence is obtained through comparative analysis.

Table 1. Prediction of population in 2022

Model	GM	GM(1,1)
Actual value (10000 persons)	606.65	606.65
Estimate	605.32	605.78
Residual	1.53	0.89
Relative error (‰)	0.19	0.12
Mean square deviation ratio C	0.0021	0.0013
Small probability error P	1	1

Based on the simulation prediction of the population in 2022 using the long series grey increment model, the simulation prediction formula of the ordinary grey increment model is as follows:

$$X_1(t+1) = -43275.255 * \exp(-0.0238632^* t) + (655371.383), t = 1, 2, 3, \ldots \quad (21)$$

Simulation prediction formula of discrete grey incremental model:

$$X^{(1)}(k+1) = 0.95329X^{(1)}(k) + 1843.4521, \mathrm{k} = 1, 2, 3, \ldots \quad (22)$$

Prediction results of permanent population in 2022:

3.2 Future Population Forecast

The population of Huizhou in the next 15 years is predicted. The prediction results show that the urban population of Huizhou will still increase year by year from 2023 to 2037.

The results show that the population will continue to grow in the next 15 years, and the net increase in population and growth rate both decreased after reaching their peak in 2031.

Table 2. Population Forecast from 2023 to 2037

Year	**2023**	**2024**	**2025**	**2026**	**2027**	**2028**
Population (10000 people)	621.90	632.95	651.99	662.96	682.83	695.48
Net increase	15.40	11.05	19.04	10.97	19.87	12.65
Growth rate (‰)	24.76	17.46	29.20	16.55	29.09	18.19
Year	**2029**	**2030**	**2031**	**2032**	**2033**	
Population (10000 people)	711.51	728.04	747.24	760.17	771.26	
Net increase	16.02	16.54	19.20	12.92	11.09	
Growth rate (‰)	22.52	22.71	25.70	17.00	14.38	

4 Conclusion

This paper describes the research situation of population development trend prediction based on the gray discrete incremental model, using the GM (1,1) model theory to achieve the method. According to the prediction results, the total population will reach 7.7126 million in 2033, 2031 indicates that population growth will tend to slow down.

Taking the permanent population of Huizhou as an example, the permanent population data of Huizhou was released in 2022, and the permanent population of Huizhou continued to grow, exceeding 6 million. What is more appropriate for housing supply? Against this background, it is necessary to predict the future number of permanent residents in Huizhou in order to better plan the housing supply. Therefore, the gray discrete incremental model has a great prospect and future, so it is hoped that scientists around the world can further make better research.

Acknowledgements. This research has been supported by the Special Fund of Guangdong Education Department (2020KZDZX1196).

References

1. Aguila-Leon, J., et al., Solar photovoltaic Maximum Power Point Tracking controller optimization using Grey Wolf Optimizer: A performance comparison between bio-inspired and traditional algorithms. Expert Systems with Applications, 2023. **211**: p. 118700.
2. Batty, M. and Y. Xie, *Self-organized criticality and urban development.* Discrete Dynamics in Nature and Society, 1999. **3**(2-3): p. 109-124.
3. Britton, J.R., et al., *From introduction to fishery dominance: the initial impacts of the invasive carp Cyprinus carpio in Lake Naivasha, Kenya, 1999 to 2006.* Journal of Fish biology, 2007. **71**: p. 239-257.
4. Ding, S., et al., *The optimization of grey Verhulst model and its application.* The Journal of Grey System, 2015. **27**(2): p. 1-13.
5. Gao, M., et al., *A novel fractional grey Riccati model for carbon emission prediction.* Journal of Cleaner Production, 2021. **282**: p. 124471.
6. Hu, R., et al., *Using a Mathematical Model to Analyze the Population Increase from Dual Income, No Kid Families.* Recent Advances in Computer Science and Information Engineering: Volume 6, 2012: p. 807-810.
7. Ikram, M., Q. Zhang, and R. Sroufe, *Future of quality management system (ISO 9001) certification: novel grey forecasting approach.* Total Quality Management & Business Excellence, 2021. **32**(15-16): p. 1666-1693.
8. Kaur, R., et al., *Discrete multiobjective grey wolf algorithm based optimal sizing and sensitivity analysis of PV-wind-battery system for rural telecom towers.* IEEE Systems Journal, 2019. **14**(1): p. 729-737.
9. Mansfield, C. and J. Cartwright, *Fault growth by linkage: observations and implications from analogue models.* Journal of Structural Geology, 2001. **23**(5): p. 745-763.
10. Ma, W., X. Zhu, and M. Wang, *Forecasting iron ore import and consumption of China using grey model optimized by particle swarm optimization algorithm.* Resources Policy, 2013. **38**(4): p. 613-620.

11. Pan, W., et al., *A health indicator extraction and optimization for capacity estimation of Li-ion battery using incremental capacity curves.* Journal of Energy Storage, 2021. **42**: p. 103072.
12. Precup, R.-E., et al., *Grey wolf optimizer-based approach to the tuning of PI-fuzzy controllers with a reduced process parametric sensitivity.* IFAC-PapersOnLine, 2016. **49**(5): p. 55-60.
13. Rehman, E., et al., *Sectoral-based CO 2 emissions of Pakistan: A novel Grey Relation Analysis (GRA) approach.* Environmental Science and Pollution Research, 2020. **27**: p. 29118-29129.
14. Saxena, A., *Grey forecasting models based on internal optimization for Novel Corona virus (COVID-19).* Applied Soft Computing, 2021. **111**: p. 107735.
15. Seyyedabbasi, A. and F. Kiani, *I-GWO and Ex-GWO: improved algorithms of the Grey Wolf Optimizer to solve global optimization problems.* Engineering with Computers, 2021. **37**(1): p. 509-532.
16. Xiao, Q., et al., *A novel grey Riccati–Bernoulli model and its application for the clean energy consumption prediction.* Engineering Applications of Artificial Intelligence, 2020. **95**: p. 103863.
17. Xie, F., et al., *Grey forecast model for accurate recommendation in presence of data sparsity and correlation.* Knowledge-Based Systems, 2014. **69**: p. 179-190.
18. Yang, Y., et al., *An enhanced multi-objective grey wolf optimizer for service composition in cloud manufacturing.* Applied Soft Computing, 2020. **87**: p. 106003.
19. Ye, L., et al., *An enhanced multivariable dynamic time-delay discrete grey forecasting model for predicting China's carbon emissions.* Energy, 2022. **249**: p. 123681.
20. Zhang, Y., et al., *A novel grey Lotka–Volterra model driven by the mechanism of competition and cooperation for energy consumption forecasting.* Energy, 2023. **264**: p. 126154.
21. Zhang, Z., Q. Tang, and L. Zhang, *Mathematical model and grey wolf optimization for low-carbon and low-noise U-shaped robotic assembly line balancing problem.* Journal of Cleaner Production, 2019. **215**: p. 744-756.
22. Zhou, D., *A new hybrid grey neural network based on grey verhulst model and BP neural network for time series forecasting.* International Journal of Information Technology and Computer Science, 2013. **5**(10): p. 114-120.