ISSN 2616-5805 Vol. 4, Issue 4: 45-50, DOI: 10.25236/AJMS.2023.040407

Birth Population Prediction and Influencing Factors Analysis Based on Gray Verhulst Mathematical Model

Jing Ran

Mathematics Teaching Department, Sichuan University Jinjiang College, Meishan, Sichuan, 620000, China

Abstract: *This study is based on the grey Verhulst mathematical model and utilizes the birth population data of Chengdu City from 2013 to 2022 to predict the future five-year birth population in Chengdu City. Four factors influencing the birth population are selected: per capita disposable income, average proportion of elderly population, number of kindergartens, and number of hospitals and health clinics. The entropy weight method and grey relational analysis method are employed to calculate the grey weighted relational degree, thereby analyzing the impact of each factor on the birth population.*

Keywords: *Grey Verhulst Model, Grey Correlation Analysis, Entropy Weight Method, Birth Population*

1. Introduction

Predicting the birth population is an important topic in population studies, providing decision-making basis for policy formulation, social development planning, resource allocation, economic growth, and research on social issues. Accurate birth population prediction contributes to promoting sustainable development and improving population health. Numerous scholars have established mathematical models to forecast the birth population. For example, Yanru Li and Xingji Huang analyzed the birth population trends in Shanghai and Anhui Province using the GM(1,1) grey prediction model [1], while Buchuan Zhou and others established the Leslie matrix extension model for population prediction in China. This study establishes the grey Verhulst model to predict the birth population in Chengdu City. Compared to traditional statistical models, the grey Verhulst model combines grey system theory, which enables better handling of data uncertainty and incompleteness. Since its proposal [2], this prediction model has been widely applied in various fields such as industry, agriculture, transportation, medicine, and military [3-7]. Researchers have conducted extensive and systematic studies on the initial value, background value, modeling mechanism, model properties, and model combinations of grey prediction models [8-10], promoting the development and improvement of the grey prediction model theoretical system. The grey model has also been extended to new categories of prediction models such as GM(1, n), DGM(1,1), NDGM(1,1), and CM(1,1) power models [11-14]. Furthermore, grey prediction models have also been studied in population forecasting [15]. In this study, the grey Verhulst model is utilized for birth population prediction, along with the integration of the entropy weight method and grey relational analysis method to analyze the impact of four factors on the birth population. The basic idea of grey relational analysis is to determine the degree of similarity between the geometric shapes of reference data and several comparison data, reflecting the correlation between curves [16]. By calculating the grey relational degree using the model formula and utilizing the calculated grey relational degree, we can determine the extent to which each indicator affects the birth population, providing an accurate and objective understanding of the magnitude of the impact of each indicator on the birth population. The data in this study is sourced from the (Chengdu Statistical Yearbook) and the (China Statistical Yearbook).

2. The Gray Verhulst Mathematical Model

Verhulst model is mainly used to describe the process with a saturated state, that is, the S-shaped process, which is often used in population prediction, biological growth, reproduction prediction, etc. The basic principles and calculation methods of Verhulst model are summarized as follows:

ISSN 2616-5805 Vol. 4, Issue 4: 45-50, DOI: 10.25236/AJMS.2023.040407

2.1 Model Establishment

Set raw data series,

$$x^{(0)}=\left(x^{(0)}(1),x^{(0)}(2),\ldots,x^{(0)}(n)\right),$$

First cumulative generated sequence,

$$x^{(1)}=\left(x^{(1)}(1),x^{(1)}(2),\ldots,x^{(1)}(n)\right),$$

Generates a sequence for the mean value of $x^{(1)}$,

$$z^{(1)}=\left(z^{(1)}(2),z^{(1)}(3),\ldots,z^{(1)}(n)\right),$$

To build the whitening equation of the gray Verhulst model,

$$\frac{dx^{(1)}(t)}{dt}+ax^{(1)}(t)=b\left[x^{(1)}(t)\right]^2 , \tag{1}$$

Where a is the development coefficient, b is the grey action, t is time.

2.2 Model Solving

Assume $u=\left[a,b\right]^T$ is a sequence of parameters, and

$$B=\begin{bmatrix} -z^{(1)}(2) & \left(z^{(1)}(2)\right)^2 \\ -z^{(1)}(3) & \left(z^{(1)}(3)\right)^2 \\ \vdots & \vdots \\ -z^{(1)}(n) & \left(z^{(1)}(n)\right)^2 \end{bmatrix}, Y=\begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix},$$

Where u satisfied $\widehat{u}=\left[\widehat{a},\widehat{b}\right]^T=(B^TB)^{-1}B^TY$,

The solution of equation (1) is

$$x^{(1)}(t)=\frac{\widehat{a}x^{(0)}(1)}{\widehat{b}x^{(0)}(1)+\left[\widehat{a}-\widehat{b}x^{(0)}(1)\right]e^{\widehat{a}t}},$$

The time response sequence of the grey Verhulst model is

$$\widehat{x}^{(1)}(k+1)=\frac{\widehat{a}x^{(0)}(1)}{\widehat{b}x^{(0)}(1)+\left[\widehat{a}-\widehat{b}x^{(0)}(1)\right]e^{\widehat{a}k}}.$$

2.3 Model Testing

Pass the residual test, make the residual as $\varepsilon(k)$,

$$\varepsilon(k)=\frac{x^{(0)}(k)-\widehat{x}^{(0)}(k)}{x^{(0)}(k)}, k=1,2,\ldots,n,$$

If $\varepsilon(k)<0.2$, can be considered to meet the general requirements; If $\varepsilon(k)<0.1$, it is considered to meet the higher requirements.

Birth population prediction based on grey Verhulst prediction model, taking Chengdu as an example.

3. Birth Population Prediction Based on Grey Verhulst Prediction Model, Taking Chengdu as an Example

3.1 Birth Population Status of Chengdu

Based on the relevant data published in Chengdu Statistical Yearbook, the birth population of Chengdu from 2006 to 2022 is sorted out, as shown in Table 1.

It can be observed that from the deregulation of the one-child policy in 2013 to the full

ISSN 2616-5805 Vol. 4, Issue 4: 45-50, DOI: 10.25236/AJMS.2023.040407

implementation of the two-child policy in 2016, there was a birth boom. In 2016, there were 173,600 newborns, an increase of 100,000 over a span of 10 years from 2006. In 2017, there were 196,800 newborns, an increase of 23,200 over 2016, while in 2018, there was a decrease of 14,000 compared to 2017. However, in 2019, the number of newborns peaked at 200,700, showing an increase of 17,900 from 2018, until it started to decline in 2020.

Table 1: Birth population in Chengdu from 2006 to 2022

Year	2006	2007	2008	2009	2010	2011
Birth population	74200	92500	10000	91500	97900	105200
Year	2012	2013	2014	2015	2016	2017
Birth population	116900	106700	121200	141100	173600	196800
Year	2018	2019	2020	2021	2022	
Birth population	182800	200700	182900	154800	150000	

3.2 Establishment and Solution of Birth Population Prediction Model in Chengdu

The original data sequence is established from the birth population data of 2013-2022, then the first cumulative sequence is as follows:

$$x^{(1)}=(10.67, 22.79, 36.9, 54.26, \ldots, 161.06),$$

The mean generating sequence: $z^{(1)}=(16.73, \ 29.84, \ 45.58, \ldots, 153.56)$.

And,

$$B=\begin{bmatrix} -z^{(1)}(2) & \left(z^{(1)}(2)\right)^2 \\ -z^{(1)}(3) & \left(z^{(1)}(3)\right)^2 \\ \vdots & \vdots \\ -z^{(1)}(10) & \left(z^{(1)}(10)\right)^2 \end{bmatrix}, Y=\begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(10) \end{bmatrix},$$

Calculated by Matlab,

$$\hat{u}=\left[\hat{a}, \hat{b}\right]^T=(B^TB)^{-1}B^TY=\begin{bmatrix} -0.4610 \\ -0.0025 \end{bmatrix},$$

The Verhulst model is

$$\frac{dx^{(1)}(t)}{dt}-0.461x^{(1)}(t)=-0.0025\left[x^{(1)}(t)\right]^2,$$

Its time response is

$$\hat{x}^{(1)}(k+1)=\frac{-0.461x^{(0)}(1)}{-0.0025x^{(0)}(1)+\left[-0.461+0.0025x^{(0)}(1)\right]e^{\hat{a}k}},$$

Let $k=0, 1, \ldots 15$, the predicted value of $x^{(0)}$ can be obtained. The error analysis data is shown in Table 2, and it can be observed from the relative error values that the accuracy of the model is relatively high, indicating its suitability for prediction purposes.

Table 2: Original year data, predicted value and Verhulst model error

Year	Raw data	Predicted value	Residual	Relative error
2018	18.28	18.817	0.537	0.02937
2019	20.07	21.0087	0.9387	0.0467
2020	18.29	21.1749	2.8849	0.1577
2021	15.48	19.2526	3.7726	0.2437
2022	15	15.9322	0.9322	0.0621

Table 3: Predicted birth population of Chengdu from 2023 to 2028

Year	2023	2024	2025	2026	2027	2028
Births (10,000)	12.19	11.77	10.05	8.041	7.65	7.21

The predicted birth population of Chengdu from 2023 to 2028 is shown in Table 3. It can be

ISSN 2616-5805 Vol. 4, Issue 4: 45-50, DOI: 10.25236/AJMS.2023.040407

analyzed that the birth population of Chengdu will show a downward trend in the future.

4. Analysis of Influencing Factors of Birth Population in Chengdu Based on Grey Correlation Analysis

4.1 Gray Correlation Coefficient Analysis of Birth Population Number Evaluation Index

Grey relational analysis is a widely used multivariate statistical method where the degree of correlation between various factors in the system and the main variable is measured based on the similarity of the development trends among these factors. Its advantages include strong objectivity and high tolerance. In this study, the birth rate in Chengdu City from 2013 to 2022 is selected as the mother sequence, and influencing factors are chosen as the feature sequences to establish a grey relational analysis. After reviewing extensive literature, the selected indicators for the influencing factors are per capita disposable income, average proportion of elderly population, number of kindergartens, and number of healthcare institutions. Among these, the birth rate measures the level of births, per capita disposable income measures the economic level, the average proportion of elderly population measures the burden of elderly care, the number of kindergartens measures the level of education, and the number of healthcare institutions measures the level of healthcare. The data is sourced from the "Chengdu Statistical Yearbook," and the specific calculation process is as follows:

Step 1. Determine the mother sequence and feature sequence,

The mother sequence is $X_0=\left(X_0(1), X_0(2), \ldots, X_0(10)\right)^T$,

The feature sequence is $X_i=\left(X_i(1), X_i(2), \ldots X_i(10)\right)^T, i=1,2,3,4$,

Step 2. Preprocessing of the data is performed by normalizing and dimensionless scaling;

Step 3. Calculate the grey correlation coefficient,

$$\gamma_i(k)=\frac{a+\rho b}{\left|x_0(k)-x_i(k)\right|+\rho b}, i=1,2,3,4 \quad k=1,2,\ldots,10, \tag{2}$$

$\gamma_i(k)$ is the correlation coefficient between feature sequence X_i and mother sequence X_0 on the k index, where $\rho \in [0,1]$ is the resolution coefficient. A larger resolution coefficient ρ indicates a higher resolution, while a smaller resolution coefficient ρ indicates a lower resolution. In this study, the resolution coefficient is taken as $\rho=0.5$;

Where, $a=\min_{1\le i\le m}\min_{1\le k\le n}\left|X_0(k)-X_i(k)\right|$ and $b=\max_{1\le i\le m}\max_{1\le k\le n}\left|X_0(k)-X_i(k)\right|$ respectively represent the minimum difference between the two poles and the maximum difference between the two poles.

According to equation (2), the grey correlation coefficient affecting the evaluation index of the number of births was calculated using SPSS, as shown in Table 4.

Table 4: Results of grey correlation coefficient of birth population number evaluation index

Year	Number of hospitals and health centers X_1	Per capita disposable income in urban areas X_2	Number of kindergartens X_3	Average proportion of elderly population X_4
2013	0.898	0.815	0.968	0.642
2014	0.923	0.754	0.807	0.773
2015	0.709	0.598	0.637	0.924
...	...	...	...	...
2022	0.410	0.337	0.405	0.520

4.2 Entropy Weight Analysis of Birth Population Number Evaluation Index

The entropy weight method does not need to consider the relative importance of each index factor when calculating the weight, and can obtain a relatively objective index weight. Therefore, this paper adopts the entropy weight method to calculate the weight of each evaluation index. The main calculation steps are as follows:

ISSN 2616-5805 Vol. 4, Issue 4: 45-50, DOI: 10.25236/AJMS.2023.040407

Step 1. Non-dimensional processing. Calculate the weight of the $k(k=1,2,\ldots,10)$ sample value under the $i(i=1,2,3,4)$ evaluation indicator,

$$P_{ik}=\frac{\gamma_i(k)}{\sum_{k=1}^{10}\gamma_i(k)},i=1,2,3,4 \tag{3}$$

Step 2. Calculate the information entropy of each evaluation index,

$$e_i=-\frac{1}{\ln 10}\left(\sum_{k=1}^{10}P_{ik}\ln P_{ik}\right),i=1,2,3,4 \tag{4}$$

Step 3. Calculate the weight coefficient of each evaluation index,

$$\omega_i=\frac{1-e_i}{\sum_{i=1}^{4}\left(1-e_i\right)},i=1,2,3,4 \tag{5}$$

The entropy weights of the evaluation indicators for the number of births are calculated using SPSS based on Equations (3), (4), and (5), as shown in Table 5.

Table 5: Results of calculating weights using the entropy method

Evaluation index	Information entropy e_i	Information utility d_i	Weight (%) ω_i
Number of hospitals and health centersX1	0.931	0.069	23.352
Urban per capita disposable incomeX2	0.929	0.071	23.965
Number of kindergartensX3	0.917	0.083	28.012
Average proportion of elderly populationX4	0.927	0.073	24.671

4.3 Calculating Grey Weighted Correlation

$$\gamma_i=\sum_{k=1}^{10}\omega_i\gamma_i(k),i=1,2,3,4 \tag{6}$$

According to Equation (6), the grey relational degree between the mother sequence and the feature sequence is calculated using SPSS, and the grey relational degrees are sorted as shown in Table 6.

Table 6: Results of grey weighted correlation degree

Evaluation items	Relevance	Ranking
Number of kindergartens	0.748	1
Average proportion of elderly population	0.709	2
Number of hospitals and health centers	0.702	3
Urban per capita disposable income	0.631	4

Combining the results of the correlation coefficients, weighted processing is performed, resulting in the final correlation degree values. Using these correlation degree values, the four evaluation objects are ranked. The correlation degree values range from 0 to 1, where a higher value indicates a stronger correlation with the birth rate, implying a greater impact. From the table above, it can be seen that among the four influencing factors, the number of kindergartens has the highest impact (correlation degree: 0.748), followed by the average proportion of elderly population (correlation degree: 0.709).

References

[1] Yanru Li, Forecast and Influence of Birth Population Development Trend in Shanghai, Economic

ISSN 2616-5805 Vol. 4, Issue 4: 45-50, DOI: 10.25236/AJMS.2023.040407

Research Guide, pp. 72-76, 2022.
[2] J. L. Deng, Estimate and Decision of Grey System, Huazhong University of Science and Technology Press, Wuhan, China, 2002.
[3] B. Zeng, M. Tong, and X. Ma, "A new-structure grey Verhulst model: development and performance comparison," Applied Mathematical Modelling, vol. 81, no. 81, pp. 522–537, 2020.
[4] H. M. Duan, X. P. Xiao, and Q. Z. Xiao, "An inertia grey discrete model and its application in short-term traffic flow prediction and state determination, " Neural Computing and Applications, vol. 32, no. 12, pp. 8617–8633, 2020.
[5] X. Xiao, H. Duan, and J. Wen, "A novel car-following inertia gray model and its application in forecasting short-term traffic flow, " Applied Mathematical Modelling, vol. 87, pp. 546–570, 2020.
[6] Z. -X. Wang, D. -D. Li, and H. -H. Zheng, "Model comparison of GM (1, 1) and DGM (1, 1) based on Monte-Carlo simulation, " Physica A: Statistical Mechanics and Its Applications, vol. 542, Article ID 123341, 2020.
[7] H. M. Duan and X. P. Xiao, "A multimode dynamic short-term traffic flow grey prediction model of high-dimension tensors," Complexity, vol. 2019, Article ID 9162163, 18 pages, 2019.
[8] Y. Chen, W. Lifeng, L. Lianyi, and Z. Kai, "Fractional Hausdorff grey model and its properties, " Chaos, Solitons & Fractals, vol. 138, Article ID 109915, 2020.
[9] H. M. Duan, G. Y. Lei, and K. L. Shao, "Forecasting crude oil consumption in China using a grey prediction model with an optimal fractional-order accumulating operator, " Complexity, vol. 2018, Article ID 3869619, 12 pages, 2018.
[10] X. P. Xiao and H. M. Duan, "A new grey model for traffic flow mechanics," Engineering Applications of Artificial Intelligence, vol. 88, Article ID 103350, 2020.
[11] B. Zeng, H. Duan, and Y. Zhou, "A new multivariable grey prediction model with structure compatibility, " Applied Mathematical Modelling, vol. 75, no. 75, pp. 385–397, 2019.
[12] Q. Z. Xiao, M. Y. Shan, M. Y. Gao et al., "Parameter optimization for Bernoulli model on biomass energy consumption prediction," Applied Soft Computing Journal, vol. 95, Article ID 106538, 2020.
[13] S. Mao, Y. Kang, Y. Zhang, X. Xiao, and H. Zhu, "Fractional grey model based on non-singular exponential kernel and its application in the prediction of electronic waste precious metal content, " ISA Transactions, 2020.
[14] S. H. Mao, M. Zhu, X. P. Wang, and X. P. Xiao, "Grey Lotka Volterra model for the competition and cooperation between third party online payment systems and online banking in China, " Applied Soft Computing, vol. 95, Article ID 106501, 2020.
[15] H. Gao, T. Yao, and X. Kang, "Population forecast of Anhui province based on the GM (1, 1) model," Grey Systems: Theory and Application, vol. 7, no. 1, pp. 19–30, 2017.
[16] Jieli Qi, Wenlong Xue. Research on Influence Factors of the Tourism Revenue of Hainan International Tourism Island Based on Grey Relational Analysis. Journal of Hainan Radio & TV University, 2015.