Various Comparative Studies of Constantly Applied Methods in Assisting the Demography Research of China

ABSTRACT

Demography research is of great importance due to it can affect various aspects of a country. We mainly focus on using various comparative studies to examine and predict the population quantity, female size, and urban population size of China. We initially compared the effectiveness of several interpolation algorithms to add more samples and then we constructed two more comparative studies of the Malthus model, Logistic model, GM (1, 1) model, BP neural network, and Leslie model to assist the demography research of China. Especially, three novel metaheuristic algorithms were applied to the Logistic model. Compared results verified that metaheuristics can assist the Logistic model to obtain far more accurate parameters and also revealed that using the Logistic model combined with a novel metaheuristic to predict the population quantity and urban population size and the BP neural network to estimate the female size can acquire the most ideal results.

Keywords: Demography research; Comparative studies; Applications of metaheuristics; Prediction-model .

https://doi.org/10.32388/K0EAYP

INTRODUCTION

The 7th census of China shows the population quantity in 2020 is 1.41178 billion (male occupies 51.24%), which increased 72.06 million comparing the year of 2010, the growth rate is 5.38%. The census brings the demography research back to people's attention. Demography derives from Greek etymologically, which means "description of people". Basically, it examines and predicts the size, structure, and variation of human population, which has become one of the most important issues in the world due to its direct influence of economy, policy, culture, education, and environment and determination of exploring and exploiting natural resources. Practically, mathematical modeling strategies were implemented by mathematicians to observe and analyze collected datasets and make predictions in order to assist the demography research. Models can mainly be classified into two categories, one is differential equation-based (e.g. Malthus growth model, Logistic growth model, and GM (1, 1) growth model) and the other is non-differential equation-based (e.g. BP neural network and Leslie growth model). Differential equation-based models are widely established in many fields due to their simplicity and capability of interpreting the general rule. However, comparing with artificial intelligence-based BP neural network or matrix calculation-based Leslie growth model, these models can yield unsuitable results when data contain unconventional features.

The population growth is affected by many factors including the size, birthrate, death rate, and sex proportion, particularly, many special factors can influence the growth even more intensively, such as national policy, education, economic environment, and human interactions. In this article, we mainly focus on estimating the population quantity, female size, and urban population size of China in the next ten years using various comparative studies. We will compare the capability of several interpolation techniques including Kriging, Inverse Distance To a Power (IDTP), and Radial Basis Function (RBF) at first and determine the most suitable method to enlarge the size of originally collected data sequences. After the interpolation, another comparative study will be lunched using the Malthus model, Logistic model, GM (1, 1) model, and BP neural network, which aims to select the best prediction model of different predicting objectives. Besides, three novel metaheuristic algorithms including Barnacles Mating Optimizer (BMO), Sine Cosine Algorithm (SCA), and Hunger Games Search (HGS) will be applied to the Logistic model to realize the acquisition of its parameters. Moreover, predictions generated from the selected models will also be compared with the Leslie model to generate intensive comparisons and forge further conclusions. At the end of this article, we will simulate and observe the effect of different census period using resampled datasets of different resampling intervals and provide relevant conclusions and suggestions.

METHODOLOGY

As we mentioned above, we consider to apply constantly-used prediction models including the Malthus model, Logistic model, GM (1, 1) model, BP neural network, and Leslie model to assist the demography study of China. At this section, preconditions and assumptions of these models will be explained initially, and then basic methodologies and mechanisms will be introduced in detail. Among these prediction models, involved parameters of the first three models based on differential equation will be acquired utilizing the linear regression analysis or the least-square method. Particularly, three novel metaheuristic algorithms will be introduced and applied to the Logistic model to assist its parameters acquisition.

Assumptions

Successful mathematical modelling needs detailed preconditions and rigorous assumptions, thus we assume:

1. No natural disasters, wars, and other occasional events happen in the next ten years;

2. People can only live 100 years ultimately;

3. Migration of people only happens between cities and rural areas;

4. The fertility rate of women at all ages is a fixed number and the sex proportion of the newborns is 105 (This assumption is mainly defined for the Leslie growth model).

5. Collected data are objective and rigorous.

Malthus growth model

In 1798, Thomas R. Malthus assumed the population growth rate r (r can be positive or negative) is a fixed number and proposed the famous Malthus growth model (By, 2007). We denote $x(t)$ as the population size at the t-th year, with the fixed growth rate r, equation (1) can be obtained easily if we want to acquire how many people increased or decreased between $[t, t+\Delta t]$.

$$x(t+\Delta t)-x(t)=\mathrm{r}x(t)\Delta t \tag{1}$$

as $\Delta t \to 0$, the differential form of equation (1) can be written like

$$\begin{cases} \dfrac{dx(t)}{dt} = \mathrm{r}x(t) \\ x(t)\big|_{t=0} = x_0 \end{cases} \tag{2}$$

the solution of equation (2) is

$$x(t)=x_0 \exp(\mathrm{r}t) \tag{3}$$

equation (3) is the famous Malthus model. As we can see from equation (3), the growth rate r is the key of the Malthus model. We can calculate the logarithm result of the left and the right part of equation (3) simultaneously to obtain it.

$$\begin{aligned} &\ln x(t)=\ln x_0 + \mathrm{r}t \\ &\ln x(t)=\mathrm{r}t + \ln x_0 \to y = kx + b, k = r, b = \ln x_0. \end{aligned} \tag{4}$$

If the rule of population growth of a specie satisfies the Malthus model that means the population grows exponentially with time. Hence, this model is also called the exponential model and used to describe the growth of population of a newly emerged specie frequently. However, as population size increases, the environment's ability to support the population decreases, which means food availability decreases, habitat place shrinks, and birth rate declines while death rate increases. Fixed growth rate no longer suits this circumstance.

Logistic growth model

If we carefully analyze the long-term feature of the population growth, a fact can be acquired that the growth will never increase in an exponential manner. Instead, the growth rate tend to decrease gradually after it reached a certain value. The main reason is environmental constraints can limit the expandision of population. Based on this idea, the Dutch mathematician Verhulst assumed the growth rate should be a variable and decrease as the population expands. Therefore, Verhulst proposed the Logistic model in 1838. This model explicitly incorporates the idea of environment limitations and yields more reasonable results. Implementation of it can be given as follows:

$$\frac{dx(t)}{dt} = r(x)x(t),\ x(0) = x_0 \tag{5}$$

where r can be defined like

$$r(x) = \mathrm{r}_0 - sx\ (\mathrm{r}_0 > 0, s > 0) \tag{6}$$

r_0 in equation (6) is the initial growth rate when x=0, s is a damping factor and it equals r_0/x_m. x_m is the up limit of the population growth, which means the population size will cease to expand when x=x_m, it relates to the capacity of the environment and natural resources.

$$r(x) = \mathrm{r}_0(1 - \frac{x}{x_m}) \tag{7}$$

thus we substitute equation (7) into equation (5) and we can get equation (8).

$$\frac{dx(t)}{dt} = \mathrm{r}_0 x(t)(1 - \frac{x(t)}{x_m}),\ x(0) = x_0 \tag{8}$$

equation (8) is the Logistic growth model and its solution can be written like

$$x(t) = \frac{x_m}{1 + (\frac{x_m}{x_0} - 1)e^{-\mathrm{r}_0 t}} \tag{9}$$

Just like the Malthus model, parameters of the Logistic model also need to be determined. However, equation (9) cannot be linearized directly just like we mentioned before, therefore, linear regression analysis theory cannot be applied directly. We consider the following operations for equation (8).

$$\frac{dx(t)}{dt}/x(t) = \mathrm{r}_0 - sx(t),\ x(0) = x_0 \rightarrow y = kx + b,$$
$$y = \frac{dx(t)}{dt}/x(t),\ k = -s,\ s = \frac{\mathrm{r}_0}{x_m},\ b = \mathrm{r}_0. \tag{10}$$

It can be seen that with the assistance of equation (10), we can still use linear regression to estimate the parameters. The data of y can be obtained easily using numerical differentiation method.

GM (1, 1) growth model

Models that we've already introduced are simple and classic, however, they deeply rely on the simple variation tendency of the observed data, which can yield bad results when data vary in an unconventional way. The gray prediction model also based on differential equation is proposed and applied by scholars (Fang et al., 2013; Huang et al., 2013; Zhao et al., 2015) to alter the problem. The idea of its construction is to use some linear or non-linear processions to enhance the regularity of the collected data, and then forging a differential equation based on this regularity to make predictions. Sometimes, the irregular or unconventional nature of the collected data can be suppressed effectively, but othertimes, this method will suffer from the same problem that the Malthus model or the Logistic model has.

GM (1, 1) is the most commonly-used model in the gray prediction theory. We suppose the collected data are

$$x^{(0)} = (x^{(0)}(1), x^{(0)}(2), \ldots, x^{(0)}(n)) \tag{15}$$

and we should test whether the obtained data are suitable for GM (1, 1) or not at first, therefore, we can calculate a factor λ using equation (16).

$$\lambda(k) = \frac{x^{(0)}(k-1)}{x^{(0)}(k)}, k = 2,3,...,n \quad (16)$$

If λ belongs to the permitted range $\Theta = (e^{-\frac{2}{n+1}}, e^{\frac{2}{n+2}})$, which means the obtained data are qualified and suitable for processing. Nevertheless, some transforms are needed to modify the data, the easiest way is to add a constant.

$$y^{(0)}(k) = x^{(0)}(k) + \text{constant}, k = 1,2,...,n \quad (17)$$

thus the modified data $y^{(0)} = (y^{(0)}(1), y^{(0)}(2), ..., y^{(0)}(n))$ can satisfy equation (18).

$$\lambda_y = \frac{y^{(0)}(k-1)}{y^{(0)}(k)} \in \Theta, k = 2,3,...,n \quad (18)$$

After the test, we assume the qualified data sequence is $x^{(0)}$, and we can use equation (19) to generate a new sequence.

$$x^{(1)} = (x^{(1)}(1), x^{(1)}(2), ..., x^{(1)}(n)) = (x^{(1)}(1), x^{(1)}(1) + x^{(0)}(2), ..., x^{(1)}(n-1) + x^{(0)}(n)) \quad (19)$$

where $x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i)(k = 1,2,...,n)$. With the newly generated sequence, we can calculate the average between two neighbouring values and make another sequence $z^{(1)} = (z^{(1)}(2), z^{(1)}(3), ..., z^{(1)}(n))$.

$$z^{(1)}(k) = 0.5x^{(1)}(k) + 0.5x^{(1)}(k-1), k = 2,3,...,n \quad (20)$$

Therefore, the so called gray differential equation (equation (21)) and the white differential equation (equation (22)) can be formed.

$$x^{(0)}(k) + cz^{(1)}(k) = v, k = 2,3,...,n \quad (21)$$

$$\frac{dx^{(1)}}{dt} + cx^{(1)}(t) = v \quad (22)$$

If we denote $u=(c,v)^T, Y=(x^{(0)}(2),x^{(0)}(3),...,x^{(0)}(n))^T$, $B=\begin{pmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{pmatrix}$,

hence, according to the least-square theorem, we can minimize equation (23) to acquire the needed parameters.

$$J(\hat{u})=(Y-B\hat{u})^T(Y-B\hat{u}) \tag{23}$$

So that $\hat{u}=(c,v)^T=(B^TB)^{-1}B^TY$ and equation (21) can be solved

$$x^{(1)}(k+1)=(x^{(0)}(1)-\frac{v}{c})e^{-ck}+\frac{v}{c}, k=0,1,...,n-1,... \tag{24}$$

With the relationship of $\hat{x}^{(0)}(k+1)=\hat{x}^{(1)}(k+1)-\hat{x}^{(1)}(k), k=1,2,...,n-1,...$, the predicted result can be easily acquired. Additionally, we can still use linear regression to process just like we mentioned in equation (10):

$$\frac{dx^{(1)}}{dt}=-cx^{(1)}(t)+v \rightarrow y=kx+b, k=-c, b=v. \tag{25}$$

BP neural network

Stochastic or quasi-stochastic or irregular system (the variation features of the system are not very obvious) cannot be effectively described by differential equation-based models. However, artificial intelligence-based BP neural network can directly discover the potential relationships from a large amount of unclear and random data. Therefore, it can build its own prediction model and make estimations.

The BP neural network is mainly composed of an input layer, a hidden layer, and an output layer. It has self-learning ability, non-linear mapping ability, and generalization ability so that it can process various complex information in parallel. The implementation process of it is to input the original observed sequence and treat the output as the prediction result. After adjusting the weights of each layer through the backward propagation of estimated errors continuously, the non-linear relationship is established when the error between the output and the actual value is minimized, thus the BP neural network can make predictions. For more specific details, please see references. (Yin et al., 2016; Yin and Chen, 2005; Wang and Shi, 2015; Luo and Huang, 2004)

Leslie growth model

The Leslie model is different from all the models that we've introduced, it predicts using matrix operations. Australian scholar Leslie first proposed a discrete model concerning the female size, birthrate, and mortality rate with a specific matrix mechanism in 1945, which can predict not only the size but the structure of the population. Therefore, it overcomes the shortcomings of previously mentioned models that only concerns the total amount of people. The Leslie model is well acknowledged to demographers and applied to carry out population estimation work widely.

The Leslie model mainly concerns the variation and structure of the female (Chen, 2008). It needs to divide the whole population into 91 groups according to their age (numbered from 0 to 90) at first. Namely, equation (26) can be obtained easily.

$$X_i = (x_i^0, x_i^1, \ldots, x_i^{90})^{\mathrm{T}}, i = 0, 1, \ldots, k. \tag{26}$$

where $x_i^j \, (j = 0, 1, \ldots, 90)$ means the size of the female at the i-th year and j-th group. The offspring of X_0 will be sorted into the first group of X_1, which means

$$x_1^1 = f_1 \cdot x_0^0 + f_2 \cdot x_0^1 + \ldots + f_{90} \cdot x_0^{90} \tag{27}$$

$f_j (j = 0,1,\ldots,90)$ in equation (27) means the female fertility rate ($f = \frac{100}{100+\alpha} \cdot \frac{t}{1000}$, $\alpha \geq 100$ means the sex proportion of the female newborns) of each groups (in ‰). Moreover, people in *j*-th group of X_0 still have $P_j \cdot x_0^j$ can get into the *j*+1-th group of X_1, which means $x_{i+1}^{j+1} = P_j \cdot x_i^j$, P_j is the female immortality rate ($P = 1 - \frac{v}{1000}$, v is the female mortality rate) of the *j*-th group. The growth model can be written in the matrix form:

$$X_1 = \begin{pmatrix} x_1^0 \\ x_1^1 \\ \vdots \\ x_1^{90} \end{pmatrix} = \begin{pmatrix} f_0 & f_1 & \cdots & f_{89} & f_{90} \\ P_0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & P_{89} & 0 \end{pmatrix} \begin{pmatrix} x_0^0 \\ x_0^1 \\ \vdots \\ x_0^{90} \end{pmatrix} \tag{28}$$

We can use a clearer and simpler form like $X_1 = MX_0$ to substitute equation (28). *M* is the famous Leslie matrix:

$$M = \begin{pmatrix} f_0 & f_1 & \cdots & f_{89} & f_{90} \\ P_0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & P_{89} & 0 \end{pmatrix} \tag{29}$$

As P_{90} in equation (29) probably will not be zero and they cannot be sorted into a new group in X_1, thus *M* should be modified like

$$B = \begin{pmatrix} f_0 & f_1 & \cdots & f_{89} & f_{90} \\ P_0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & P_{89} & P_{90} \end{pmatrix} \tag{30}$$

With the definition of equation (28), $X_k = BX_{k-1}$ can be easily generalized, which means the female or the population quantity (if the sex proportion is determined) at the *k*-th year can be easily obtained using modified Leslie matrix multiplications.

Novel metaheuristic algorithms and their application of acquiring the defined parameters of the Logistic model

The Malthus model, Logistic model, and GM (1, 1) model, intrinsically, are the same for they all consider the prediction process as a differential equation modeling problem (especially, first-order differential equation modeling problem). Certainly, it can be further transformed into an optimization problem and solved by gradient-based methods (e.g. Newton method and the Conjugate Gradient method) or metaheuristic algorithms (e.g. PSO, SA, and GA). Transforming the prediction process into an optimization problem, that is, to determine the optimal parameters of selected models so that the difference between the generated data and the collected data can be minimized, therefore, further estimations can be obtained more accurately. In this article, we mainly consider to use recently proposed metaheuristic algorithms to predict the parameters of the Logistic model due to its solution has a clear form and parameters cannot be acquired conveniently through conventional methods. Involved strategies include BMO, SCA, and HGS. Relevant published articles (Mirjalili, 2016; Yang et al., 2021; Sulaiman et al., 2020) proved that they are more excellent than classical algorithms (e.g. PSO, SA, and GA). Basic information and processes of implementing these algorithms are discussed bellow:

Sine Cosine Algorithm

SCA creates multiple random candidate solutions initially and requires them to fluctuate around the best solution using a mathematical model based on sine and cosine functions to solve optimization problems. Several random and adaptive variables are also integrated to this algorithm to emphasize exploration and exploitation of the search space in different stages of optimizing. Mirjalili (2016) first proposed the following position updating equations as the core of SCA:

$$x_i^{g+1} = x_i^g + r_1 \times \sin(r_2) \times | r_3 P_i^g - x_i^g | \tag{31}$$

$$x_i^{g+1} = x_i^g + r_1 \times \cos(r_2) \times | r_3 P_i^g - x_i^g | \tag{32}$$

$$r_1 = \mathrm{c} - g\frac{\mathrm{c}}{\mathrm{G}}, \quad \text{c=costant}, \; g = 1,2,3,\ldots\mathrm{G}, \; i = 1,2,3,\ldots\mathrm{N}. \tag{33}$$

where x_i^{g+1} is the solution in the *i*-th dimension and *g*-th iteration, G means the total iterations, and N means the number of solutions. Especially, r_2, r_3 are random numbers, which belong to [0, 2π], [0, 2], respectively. P_i^g is the position of the best solution obtained so far in the *i*-th dimension and "||" means the absolute function. Besides, these two equations can be integrated (equation (34)) using a random number r_4, which belongs to [0,1].

$$x_i^{g+1} = \begin{cases} x_i^g + r_1 \times \sin(r_2) \times | r_3 P_i^g - x_i^g |, & r_4 < 0.5; \\ x_i^g + r_1 \times \cos(r_2) \times | r_3 P_i^g - x_i^g |, & r_4 \geq 0.5. \end{cases} \tag{34}$$

Barnacles Mating Optimizer

Barnacles are micro-organisms existed since Jurassic times. It can swim at birth and they will attach themselves to objects in the water and grow shells when they reach their adult stage. Particularly, there are more than 1400 discovered species of barnacles and most of the barnacles are hermaphroditic. An interesting fact about barnacles is that they have long penises, which is about seven to eight times the length of their bodies in order to cope with the changing tides and sedentary lifestyle. A barnacle's mating objects consists of all the neighbors within its penis reach and all its potential competitors for mates, thus variation in penis reach may have an important role in determining mating group size and local mate competition. With this background, BMO is proposed and this algorithm aims at mimicking the mating behaviour of barnacles in nature in order to solve optimization problems. (Sulaiman et al., 2020; Sulaiman et al., 2021)

Initialization:

BMO assumes the search agents are barnacles and the vector of solutions (barnacles) can be expressed as follows:

$$\mathbf{X} = \begin{pmatrix} x_1^1 & \cdots & x_1^N \\ \vdots & \ddots & \vdots \\ x_n^1 & \cdots & x_n^N \end{pmatrix}, x_i^j \in [lb^j, ub^j]. \tag{35}$$

where N is the dimension of control variables and n is the number of solutions or barnacles. ub and lb represent the upper and lower bounds of variables. Just like SCA, $\mathbf{X}$ can be initialized randomly.

Selection:

BMO takes a different approach for the selection of mating compared to other evolutionary algorithms such as GA and DE. Since the selection of mating barnacles is based on the barnacle's penis length pl, following assumptions are made:

(a) Mating objects are randomly selected if they locate within the reach of the penis (the distance to the mating objects is smaller than pl).

(b) Each barnacle can only be fertilized by one barnacle at one time.

(c) Self-mating will not be considered in this article.

(d) The sperm cast process will be launched if the distance to the mating objects at the certain iteration exceeds *pl*.

As we can see from these assumptions, we can notice that the penis length *pl* plays an important role in determining the selecting or mating behavior of barnacles. In the algorithmic language, the penis length *pl* determines the exploitation and exploration process of optimization, which is the essence of every optimizer.

Equation (36) is the mathematical form of the selection of barnacles locating within the reach of the penis:

$$\begin{aligned} barnacle\_d &= \text{randperm}\left(\tilde{n}\right) \\ barnacle\_m &= \text{randperm}\left(\tilde{n}\right) \end{aligned} \tag{36}$$

where *barnacle_d* and *barnacle_m* are the parents and randomly selected to be mated. $\tilde{n}$ ($\tilde{n} \leq n$) is the number of solutions. However, if the location of mating barnacles exceeds the value of *pl* that has been set initially, just like the assumption (d) says, the sperm cast process will be launched and can be expressed as follows:

$$x_i^N = \text{rand} \times x_{barnacle\_m}^N \tag{37}$$

where rand() is a random number between [0, 1]. One thing needs to be noted is the length or location mentioned in this algorithm is operated virtually, which means it cannot be measured in reality.

Reproduction:

The reproduction process of BMO is mainly based on the Hardy-Weinberg principle (Guo and Thompson, 1992; Sulaiman et al., 2020), which emphasizes on the inheritance characteristics or genotype frequencies of barnacles' parents in producing the offspring. The following expression are proposed to produce new solutions from barnacles' parents:

$$x_i^N = px_{barnacle\_d}^N + qx_{barnacle\_m}^N \tag{38}$$

where *barnacle_d* and *barnacle_m* are the parents defined using equation (36), p and q (q=1-p, p belongs to [0, 1]) are two uniformly distributed random number, which represent the percentage of characteristic of the parents that embedded in the next generation.

Summary of implementing BMO:

BMO creates a set of random solutions to initialize **X**. New generation of barnacles can be generated using equations (37)-(38). Generated solutions will be sorted (from the best to the worse) and placed in **X** (from the top to the bottom). Another important technique of BMO to optimize is it will select parts of the top solution of the current generation and insert them into the next generation. Therefore, the solution will converge into the global best gradually as the iteration process continues.

Hunger Games Search

HGS is designed according to the hunger-driven activities and behavioral choices of animals. It follows a simple concept of "Hunger", which is the most crucial reason for all animals to behave and make decisions. The basic steps of implementing HGS can be described as follows (Yang et al., 2021):

Approaching food:

To express the animals' behavior of approaching food in mathematical language, the following formulas are defined:

$$\overrightarrow{X(t+1)} = \begin{cases} \overrightarrow{X(t+1)} \cdot (1 + \mathrm{randn}(1)), & r_1 < l; \\ \overrightarrow{W_1} \cdot \overrightarrow{X_b} + \vec{R} \cdot \overrightarrow{W_2} \cdot \left| \overrightarrow{X_b} - \overrightarrow{X(t)} \right|, & r_1 > l, r_2 > E; \\ \overrightarrow{W_1} \cdot \overrightarrow{X_b} - \vec{R} \cdot \overrightarrow{W_2} \cdot \left| \overrightarrow{X_b} - \overrightarrow{X(t)} \right|, & r_1 > l, r_2 < E. \end{cases} \tag{39}$$

where $\vec{R}$ is in the range of [-a, a]; r_1 and r_2 are two random numbers, which are in the range of [0, 1]; randn() is a random number satisfying the normal distribution; t means the t-th iteration; $\overrightarrow{W_1}$ and $\overrightarrow{W_2}$ are two important factors representing the weight of hunger; $\overrightarrow{X(t)}$ represents each individual's location; $\overrightarrow{X_b}$ means the acquired best individual's location; l is a predefined parameter, which also belongs to [0, 1]. Especially, the formula of E is as follows:

$$E = \mathrm{sech}\left(\left| F(i) - BF \right| \right), \quad i = 1, 2, 3, ..., n. \tag{40}$$

where $F(i)$ represents the fitness value of each individual; BF is the best fitness obtained in the current iteration; sech() is a hyperbolic function ($\mathrm{sech}(x) = \frac{2}{e^x + e^{-x}}$).

The formula of $\vec{R}$ can be described as follows:

$$\begin{gathered} \vec{R} = 2 \times a \times rand - a \\ a = 2 \times \left(1 - \frac{t}{Max\_iter} \right) \end{gathered} \tag{41}$$

where $rand$ is a random number in the range of [0, 1]; Max_iter stands for the largest number of iterations.

Hunger role:

$\overrightarrow{W_1}$ in equation (39) is as follows:

$$\overrightarrow{W_1(i)} = \begin{cases} hungry(i) \cdot \frac{N}{SHungry} \times r_4, & r_3 < l; \\ 1, & r_3 > l. \end{cases} \tag{42}$$

The formula of $\overrightarrow{W_2}$ in equation (39) is determined as follows:

$$\overrightarrow{W_2(i)} = \{1 - \exp(-|hungry(i) - SHungry|)\} \times r_5 \times 2 \tag{43}$$

where *hungry* represents the hunger of each individual; *N* represents the number of individuals; *SHungry* is the sum of hungry feelings of all individuals (sum(*hungry*)); r_3, r_4, and r_5 are random numbers in the range of [0, 1]. The formula of *hungry*(*i*) is defined as:

$$hungry(i) = \begin{cases} 0, & AllFitness(i) == BF; \\ hungry(i) + H, & AllFitness(i) != BF. \end{cases} \tag{44}$$

where *AllFitness*(*i*) preserves the fitness of each individual at the current iteration. The formula for *H* can be seen as follows:

$$\begin{aligned} TH &= \frac{F(i) - BF}{WF - BF} \times r_6 \times 2 \times (UB - LB) \\ H &= \begin{cases} LH \times (1 + r), & TH < LH; \\ TH, & TH \geq LH. \end{cases} \end{aligned} \tag{45}$$

where r_6 is a random number in the range of [0, 1]; *F*(*i*) represents the fitness value of each individual; *BF* and *WF* is the best and the worst fitness obtained at the current iteration, respectively; *UB* and *LB* indicate the upper and lower bounds of the search space, respectively.

Objective function and stop criterion

As we mentioned earlier, using metaheuristic algorithm to solve the prediction problem will treat involved parameters as search agents (in this case, x_m, x_0, r_0, and constant of the Logistic model are the search agent) and then different optimization techniques will use different mechanical metaphors to guide these search agents to converge. Finally, if the *root mean square error* (*RMSE*) (defined in equation (46)) is minimized or the iteration reaches the predefined threshold, we can terminate the iteration process, acquire the optimal parameters of the Logistic model, and achieve further estimations.

$$RMSE = \sqrt{\sum_{i=1}^{k} \frac{(x_{obs}(i) - x_{pre}(i))^2}{\mathrm{K}}}$$
$$x_{pre}(i) = \frac{x_m}{1 + (\frac{x_m}{x_0} - 1)e^{-r_0 i}} + \text{constant} \tag{46}$$

where x_{obs} is the observed data, x_{pre} is the predicted data that generated from the Logistic model, and K is the size of the data.

RESULTS

Insufficient data-input can lead to bad consequences in many aspects, thus we will first compare the effectiveness of several interpolation algorithms including Kriging, IDTP, and RBF in order to select the best algorithm to add more samples to the originally collected data sequences of the population quantity, female size, and urban population size of China. With the interpolated data sequences, introduced prediction models including the Malthus model, Logistic model, especially, Logistic model modified by three metaheuristics, GM (1, 1) model, and BP neural network will be applied and compared so that the best prediction model of different predicting objectives can be selected. After the selection, selected models will be compared with the Leslie model to generate intensive comparisons and form further conclusions. At the end of this section, we will carefully compare and analyze the effect of different census interval of prediction to select a suitable census period.

Data interpolation

The originally collected census datasets from the National Bureau of Statistics (NBS) (http://www.stats.gov.cn/tjsj/ndsj/2012/indexch.htm.) are shown in Table 1. They are accurately acquired but insufficient. More complete data sequences are displayed in the appendix. However, it is obtained by sampling surveys combined with manual interference, in other words, it is deduced by the NBS so that the published datasets suffer low accuracy. A precondition of successful predictions is the availability of collecting sufficient data, which means involved models need more accurate samples to observe and disclose the statistical nature of predicting objectives. Therefore, related model's parameters can be obtained more effectively and generated predictions can be more precise and yield more reliability. We choose the commonly-used data interpolation techniques including Kriging, IDTP, and RBF (Briggs, 2012; Cordell, 2012; Cressie, 1990; Wright, 2003) to assist our study. Moreover, through a comparative study, we will select the finest method and use it to generate more samples.

The interpolated results (the interval of interpolation is one year) are displayed in Figure 1. The pink circle in Figure 1 represents the actually collected data from the NBS. The red, blue, and dark dash is the processed result of RBF, IDTP, and Kriging, respectively. Especially, the green star-shaped mark represents the data deduced by the NBS. It can be seen that the interpolated results of Kriging and RBF only yield minor differences. Contrarily, IDTP fluctuates between the census data. Hence, if we calculate the Euclidean distance between the interpolated data and the deduced data of the NBS (calculated results are displayed in Figure 2), the one with the smallest distance can be selected as the finest method. As we can see from Figure 2, Kriging yields the best performance of all categories, RBF shows slightly higher distances than Kriging. However, IDTP leaves the largest distances of three predicting objectives, thus we consider to use Kriging to make further processions. Interpolated results of Kriging are given in the appendix.

Evaluating involved prediction models

With the interpolated data sequences, initially, we will test these prediction models and evaluate their performance to obtain the most suitable model of different prediction objectives. There are few things need to be noted before the test:

(1) Data sequences suiting all the requirements of the Leslie model that we collected from the NBS can only be available at 2018. Therefore, the prediction of the matrix-based Leslie model mainly relies on one year data' s contribution, which makes it unsuitable to compare with others directly, thus we consider testing the previously established Malthus model, Logistic model, modified Logistic model using three different metaheuristics, and BP neural network initially and then constructing another comparative study to compare the Leslie model and selected models.

(2) We select 90% of the interpolated data to obtain the parameters of differential equation-based models and train the BP neural network, especially, 10% of the interpolated data are extracted from the selected data to be the verification set of the BP neural network. The rest 10% are used for estimation and comparison.

(3) The number of candidate solutions or search agents and iterations of the optimization process using metaheuristics are set to 500 and 5000, respectively. Therefore, the optimizing capability of involved metaheuristics can be examined extensively.

(4) We calculated the factor λ of the interpolated datasets and we found only very few numbers in the data sequence of urban population exceed the permitted range [0.968256677143911, 1.03225531193291], thus we believe that the GM (1, 1) model can be utilized to process.

Figure 3-10 are the predicted results of different models excluding the Leslie model. Analyzing these figures, from a qualitative point of view, the population quantity and urban population size of China can be estimated more effectively using the Logistic model combined with metaheuristics than others. The most probable reason is that they can describe the variation logic of the data properly and not affected by the approximation of numerical differentiation or the error of regression analysis. The BP neural network yields the best performance as for the estimation test of the female size due to its continuous training, testing, and learning correction modeling mechanism. Estimated results of the GM (1, 1) model combined with the linear regression parameter estimation method are better than the processed results of the Malthus model due to the data sequences vary not in a fixed rate way. Using the least-square theorem to estimate the parameters of the GM (1, 1) model can obtain the worst performance due to its inefficiency of acquiring proper parameters. Table 2 and Table 3 demonstrate the estimated parameters of different models using different parameter acquiring strategies including metaheuristics, linear regression, and least-square theorem.

After the qualitative analysis, three indicators named *error* (Euclidean distance), *Delta_mean* (relative residual), and *C* (variance ratio) are established in equation (47) to quantitatively evaluate the efficiency of these models. Both *Delta_mean* and *C* are dimensionless values. The calculated results are displayed in Table 4.

$$
\begin{aligned}
&error = \left\| x_{obs} - x_{pre} \right\|_2 \\
&delta\_mean = \text{mean}\left(\text{abs}\left(\frac{x_{obs} - x_{pre}}{x_{obs}} \right) \right) \\
&C = \text{std}(x_{obs} - x_{pre}) / \text{std}(x_{obs})
\end{aligned}
\tag{47}
$$

where x_{obs} is the observed data and x_{pre} is the predicted result, mean() and std() are two functions for calculating the average value and standard deviation, respectively. As we can see from equation (47), the most effective model can yield the smallest *error*, *Delta_mean*, and *C*, thus we can use these factors to distinguish involved models. The bolded values in Table 4 are the smallest and they can also verify the interpretations that we made in the qualitative analysis part. Therefore, combining the qualitatively and quantitatively compared results, we select the Logistic model combined with the hunger games search algorithm to predict the population quantity and urban population size and the BP neural network to estimate the female size.

Using selected models to make population-related estimations of China

Figure 11 shows the predicted results of the next ten years of the population quantity and urban population size using the Logistic model combined with HGS and the result of the female size using the BP neural network. Additionally and especially, the predicted results of the population quantity and the female size using the Leslie model are also shown in graph (a) and (b) of Figure 11 to form intensive comparisons. As we mentioned before, the Leslie model mainly uses the female structure related datasets as its input, therefore, using it to predict the population quantity of China needs a factor to deduce the male size, which is called the sex proportion. With the consideration of the publicly published data series several years before, the determined sex ratio in our research is 1.02. Predicted details are given in Table 5 and Table 6. Combining Figure 11 and Table 5-6, we can see that the female size in 2030 predicted by the BP neural network will be 694,268,722. The Logistic model combined with the HGS algorithm estimates that the population quantity and the urban population size of China in 2030 will be 1,456,213,032 and 1,032,793,473, respectively. The population quantity and the female size of China in 2030 predicted by the Leslie model will be 1,473,073,295 and 722,094,752. Therefore, estimated result of the population quantity of China of Logistic_HGS and the Leslie model only yielded mild difference. However, as we can see from the graph (b) in Figure 11 that the predicted result of the Leslie model of the female size is slightly off the trend, which means the result of the BP neural network and the Leslie model yielded a relatively obvious difference. The reason of the obviously occurred difference probably can be explained by the originally input data of the Leslie model. Predictions made by the Leslie model are based on the contribution of insufficient data, thus more consistency of the BP neural network can be achieved using more data. Finally, we determined the population quantity, urban population size, and female size of China in

2030 will be 1,456,213,032, 1,032,793,473, and 694,268,722, respectively.

Observing the effect of different census interval and select a suitable one

Census interval is an important factor for the demography research, it determines the frequency of collecting the population-related information. Theoretically, better understandings of the size, structure and variation tendency of the population will be obtained with smaller census intervals, however, it is often unrealistic due to various limitations. Hence, it's worthy to determine a suitable census interval. We consider the selection of an applicable census interval as a sensitivity problem, which means the interpolated datasets will be resampled using a fixed interval (e.g. 15 years, 10 years, and 5 years), thus involved parameters of selected models can be acquired using the resampled data sequences and further predictions and comparisons can be achieved to analyze the effect of different census periods. Here, we mainly consider using the Logistic model combined with the HGS algorithm to assist the analysis due to it can obtain higher performance when dealing with the population-related estimations of China. Predicted results generated from Logistic_HGS utilizing different sets of resampled data are presented in Figure 12. As Figure 12 shows, the red dash in graph (a), (b), and (c) means the originally interpolated data. The black dash means using the originally interpolated data to estimate model parameters and then make predictions utilizing the interval of 1 year. The blue circle, green star, and yellow square means using the resampled datasets of different sampling intervals (5, 10, and 15 years, respectively) to estimate model parameters and then make predictions using the same intervals. If we compare the calculated results in Figure 12 from the qualitative perspective, minor differences of the estimated results generated from various resampled datasets in graph (a) and (c) can be observed. However, graph (b) in Figure 12 displays a contradictory case, that is, the estimated result yields irrelevant variation tendency as the model parameters are acquired using a sampling interval of 15 years. Figure 13 demonstrates the calculated Euclidean distances between the estimated results based on different sampling intervals including 5, 10, and 15 years and the estimated results based on the originally interpolated data. The accumulated results displayed in the fourth column of Figure 13 mean to sum all the calculated distances of different predicting objectives of different sampling intervals.

As Figure 13 depicts, the best performance was yielded by the sampling interval of 5 years, which is practically reasonable. The estimated result of the female size using the sampling interval of 15 years shows a relatively big distinction comparing with others, which verifies the qualitative interpretations in Figure 12 quantitatively. Besides, the accumulated results produced by the sampling interval of 10 years are far more close to the case of 5 years than the sampling interval of 15 years. Therefore, combining the descriptions of Figure 12 and Figure 13, the census period of China can still be set to 10 years due to the relevant resampled datasets can assist the prediction model to generate almost the same effect of more densely resampled datasets and make estimations following the variation tendency of the observed data faithfully. Moreover, all the descriptions mentioned here also proved the robustness of the Logistic model combined with the HGS algorithm.

DISCUSSION

Demography research is of great importance due to related information and policies will have deep impacts on various aspects of a country. The process of the traditional demography research can be divided into two stages, one is to collect relevant statistical data, and the other is to make reasonable interpretations and estimations based on the collected datasets using proper mathematical models. Therefore, the discussion part of this article will mainly focus on decoding the two stages.

As we mentioned already, successful predictions need the support of sufficient data, which means prediction models need more accurate samples to observe the statistical feature of predicting indexes, thus related parameters of prediction models can be obtained effectively and further generated predictions can be more credible. But in our case, collected census data from the NBS are accurate but insufficient. Hence, we made an initial comparative study about comparing the effectiveness of various interpolation techniques including Kriging, IDTP, and RBF and finally selected Kriging, which yielded the best performance, to generate more reasonable samples so that the size of the collected data sequences can be enlarged. Actually, this comparative study can be more representative and yield more general meanings if we compare more interpolation algorithms. However, we only determined three algorithms to process considering they are commonly-used and the interpolation of datasets is not the theme of this article. Another thing needs to be noted is that as scholars and researchers obtain more coherent and deep understandings about the demography research, various prediction models have been proposed. In general, most of these models are different due to they are based on diverse assumptions, yet they all contain the same problem, that is, unsuitably or unreasonably selected parameters of these prediction models will leave unacceptable estimations. Therefore, with the interpolated datasets, we used another comparative studies of the Malthus model, Logistic model, GM (1, 1) model, BP neural network, and Leslie model to observe their effectiveness of assisting the demography study of China using different parameters estimation strategies. The linear regression analysis or the least-square method was utilized to acquire the parameters of the first three models, especially, three novel metaheuristic algorithms including BMO, SCA, and HGS were applied to the Logistic model to assist its acquisition of parameters. Compared results concluded that metaheuristics can assist the prediction model to obtain more accurate parameters, thus more effective prediction can be guaranteed, which gave us a thought that this way of processing can be well transformed to solve other parameter acquisition problems, such as the optimal designs of engineering (Sulaiman et al., 2020) and the inversion problems of geophysics (Moro et al., 2007). After this comparison, the BP neural network and the Logistic model combined with the HGS algorithm were chosen to make final predictions and predicted results were compared with the Leslie model intensively. Compared results of one of the predicting objectives yielded an obvious irrelevance. We believed that the irrelevance can be explained by the insufficiency of the originally input data of the Leslie model. Hence, situations might be altered if we can collect more effective data to construct the Leslie matrix.

The final thing that we want to discuss is that we carefully analyzed the effect of different census interval of the demography study at the last part of this article by considering the selection of a feasible census interval as a sensitivity problem. We used various resampled datasets to predict required parameters of selected models and then make estimations so that the effect of different census periods can be simulated. Generated analysis presented a fact that the census period of China can still be set to 10 years due to relevant resampled datasets can assist the prediction model to generate similar effects of more densely resampled datasets and make faithful predictions following the variation tendency of the observed data sequences.

CONCLUSION

This article aimed to estimate the population quantity, female size, and urban population size of China in the next ten years using various comparative studies. We first compared the efficiency of several interpolation algorithms including Kriging, IDTP, and RBF to enlarge the size of originally collected datasets of China. Compared results demonstrated that high performance interpolation was achieved by Kriging. Hence, with the interpolated datasets, we used a comparative study of the Malthus model, Logistic model, GM (1, 1) model, and BP neural network to select the best prediction model of different predicting objectives. Moreover, three novel metaheuristic algorithms were applied to the Logistic model to realize the acquisition of its parameters. Through a comprehensive evaluation of the capability of these models, we finally determined to use the BP neural network to predict the female size of China and the Logistic model combined with the HGS algorithm to predict the population quantity and urban population of China in the next ten years. Predictions generated from the two models are also compared with the Leslie model, where the Leslie model was acknowledged to be an effective prediction model for its ability of considering the structure information. Compared results demonstrated that the capability of estimating the population quantity of China of Logistic_HGS and the Leslie model only yielded mild difference. However, the BP neural network and the Leslie model's ability of predicting the female size of China yielded a relatively obvious difference due to a possible reason of the Leslie matrix was ill-constructed using insufficient datasets. Finally, we determined that the population quantity and the urban population size of China in 2030 predicted using Logistic_HGS will be 1,456,213,032 and 1,032,793,473 and the female size of China predicted using the BP neural network in 2030 will be 694,268,722. Another fact about the predicted results of this article is that the growth rate of the population quantity and female size are decreasing, yet the size of urban population is increasing. Therefore, as the intensification of the aging of the population and the increase of the urban population size, it is of great necessity to implement the three-child policy. The last part of this article extracted the interpolated data sequences at the interval of 5, 10, and 15 years respectively to simulate and compare the effect of different census period of demography study, which presented a conclusion that the census period of China can still be set to 10 years due to the relevant resampled datasets can assist the prediction model to generate almost the same effect of more densely resampled datasets and make estimations following the variation tendency of the data sequences consistently.

REFERENCES

Briggs, I. C. . "Machine contouring using minimum curvature." Geophysics 39.1(2012).

By, E. . "Malthus: 'An Essay on the Principle of Population'." Cambridge University Press (2007).

Cordell, L. . "A scattered equivalent-source method for interpolation and gridding of potential-field data in three dimensions." Geophysics 57.4(2012):629.

Cressie, N. . "The origins of kriging." Mathematical Geology 22.3(1990):239-252.

Caballero, and Armando. "Developments in the prediction of effective population size. " Heredity 73.6(1994):657-679.

Chen, W. . "The Application of Leslie Revised Model in Population Prediction." World Sci-Tech R & D (2008).

Fei, L. , and W. Feng . "Zhengzhou Population Prediction Base on the Gray Model." Henan Science (2012).

Fang, L. , et al. "Demand prediction of Heilongjiang professional health technicians: Based on gray dynamic model." International Conference on Management Science & Engineering IEEE, 2013.

Guo, S W, , and E. A. Thompson . "Performing the exact test of Hardy-Weinberg proportion for multiple alleles. " Biometrics 48.2(1992):361-372.

Huang, Songqian, et al. "Application of gray metabolic GM (1, 1) model in prediction of annual total yields of Chinese aquatic products." *Asian Agricultural Research* 5.3 (2013): 21.

Jia, N. , H. U. Hong-Ping , and Y. P. Bai . "Population prediction based on BP artificial neural network." Journal of Shandong University of Technology 10.9(2011):3297-3301.

Liu, Gousheng , and J. Yu . "Gray correlation analysis and prediction models of living refuse generation in Shanghai city." Waste Management 27.3(2007):345-351.

Lingshuang, Zhang, and Wang Fengxiao. "Combination Gray Forecast Model and Its Application in Population Prediction [J]." *Henan Science* 8 (2009).

Luo, R. G. , and M. M. Huang . "Study on population prediction of Yangtze Basin based on BP neural network." Journal of Wuhan University of Technology (2004).

Mirjalili, S. , G. G. Wang , and L. Coelho . "Binary optimization using hybrid particle swarm optimization and gravitational search algorithm." Neural Computing & Applications 25.6(2014):1423-1435.

Mirjalili, S. . "SCA: A Sine Cosine Algorithm for Solving Optimization Problems." Knowledge-Based Systems 96(2016).

Mirjalili, S. , A. Lewis , and S. Mostaghim . "Confidence measure: A novel metric for robust meta-heuristic optimisation algorithms." Information Sciences 317(2015):114-142.

Malthus, Thomas . "An Essay on the Principle of Population." Norton.

Moro, G. D. , M. Pipan , and P. Gabrielli . "Rayleigh wave dispersion curve inversion via genetic algorithms and Marginal Posterior Probability Density estimation." Journal of Applied Geophysics 61.1(2007):39-55.

MD Buhmann. "New Developments in the Theory of Radial Basis Function Interpolation." International Workshop 1993.

Nn, A , and B. Mf . "Black gold falls, black plague arise - An Opec crude oil price forecast using a gray prediction model - ScienceDirect." Upstream Oil and Gas Technology 5.

Ren, Chao , et al. "Optimal parameters selection for BP neural network based on particle swarm optimization: A case study of wind speed forecasting." Knowledge-Based Systems 56.jan.(2014):226-239.

Statistics of the People’s Republic of China.China Statistical Yearbook.http://www.stats.gov.cn/tjsj/ndsj/2012/indexch.htm.(in Chinese)

Sulaiman, M. H. , et al. An Application of Barnacles Mating Optimizer Algorithm for Combined Economic and Emission Dispatch Solution. 2021.

Sulaiman, M. H. , et al. "Barnacles Mating Optimizer: A new bio-inspired algorithm for solving engineering optimization problems." Engineering Applications of Artificial Intelligence 87(2020):103330-.

Sun, W. G. , and E. A. Thompson . "The Hardy-Weinberg law plays an important role in the field of population genetics and often serves as a basis for genetic inference. Because of its importance. . ".

Silan, H. E. , H. Sun , and D. O. Science . "Population Prediction Research in Yunnan Province Based on GreyPrediction and BP Neural Network." Computer & Digital Engineering (2016).

Wright, G. B. . Radial basis function interpolation: Numerical and analytical developments.. Diss. University of Colorado at Boulder. 2003.

Wang, X. , and W. Shi . "Prediction on total power of agricultural machinery in Liaoning Province based on BP neural network." Journal of Chinese Agricultural Mechanization (2015).

Xing, L. Y. , et al. "BP neural network in the application of Chinese population forecast." Basic Sciences Journal of Textile Universities (2010).

Xiaolong, Bi . "BP Neural Network Based Population Forecast." Journal of Wuhan University of Technology(Transportation Science & Engineering) (2007).

Yang, Y. , et al. "Hunger Games Search: Visions, Conception, Implementation, Deep Analysis, Perspectives, and Towards Performance Shifts." Expert Systems with Applications (2021):114864.

Yin, C. , and L. Chen . "Study and Application of Population Forecast Model Based on BP Neural Network." Population Journal (2005).

Yin, et al. "The forest fire prediction in JiangXi province based on PSO-GA-BP neural network." AER ADV ENG RES 2016,81.-(2016):1288-1292.

Zhang, H. , W. Chao , and X. Zhang . "A Prediction Model of China Population Growth." Computer & Information Science 1.2(2009).

Zhao, M. , et al. "Prediction and Analysis of WEEE in China Based on the Gray Model." Procedia Environmental Sciences 31(2016):925-934.

Zhao, Mingyue , et al. "The GRAY PREDICTION GM (1, 1) MODEL IN TRAFFIC FORECAST APPLICATION." (2015).

Zhang, X. . "Prediction Modeling of the Population of Preschool Children Based on Logistic Model." 2017 9th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA) IEEE, 2017.

Zhang, L. , and F Wang. "Combination Gray Forecast Model and Its Application in Population Prediction." Henan Science (2009).

LIST OF FIGURES

Figure 12. Predicted results generated from Logistic_HGS utilizing different sets of resampled data

Figure 13. Calculated Euclidean distances between the estimated results based on different sampling intervals including 5, 10, and 15 years and the estimated results based on the originally interpolated data.

FIGURES

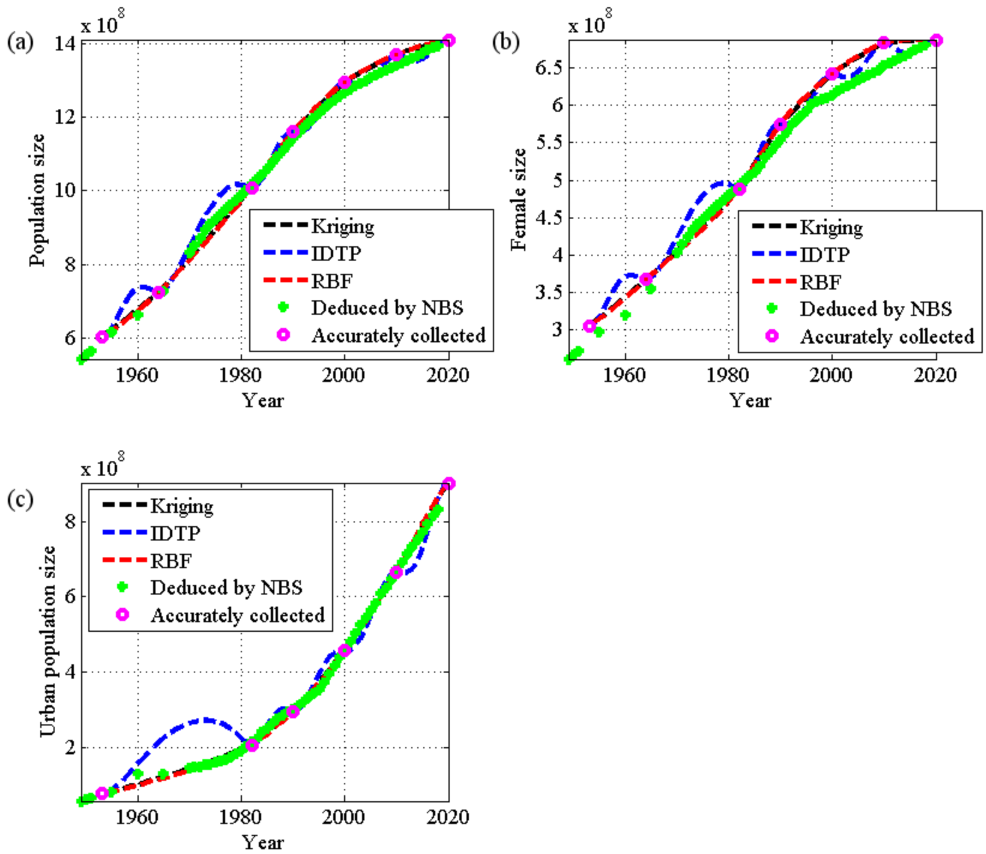

Figure 1. Interpolated results of selected algorithms and the data deduced by the NBS

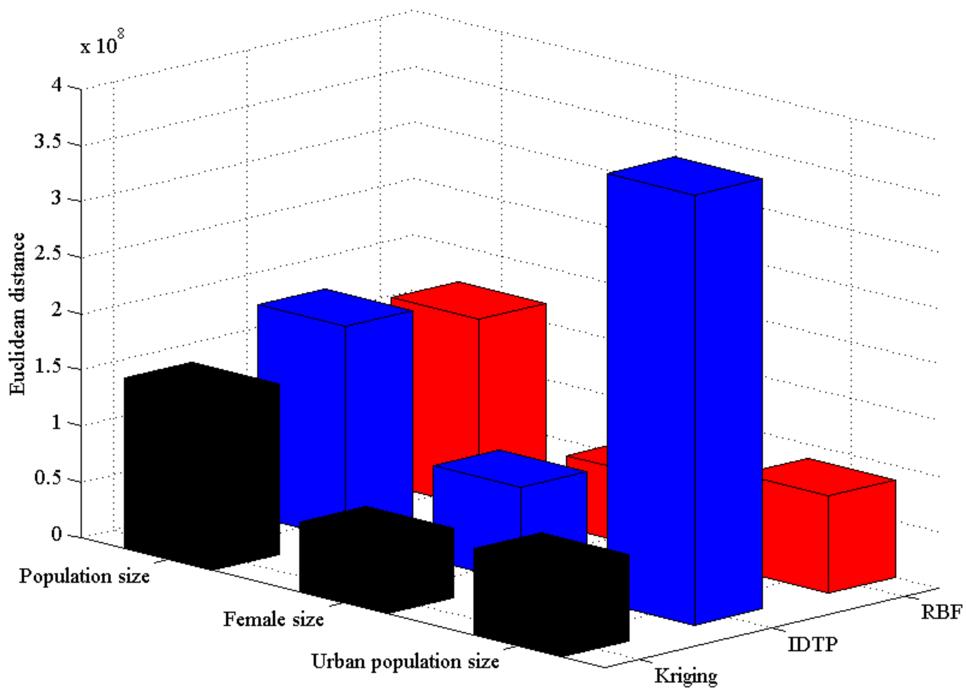

Figure 2. Calculated Euclidean distances between the interpolated data of different

categories generated from different algorithms and the deduced data of the NBS

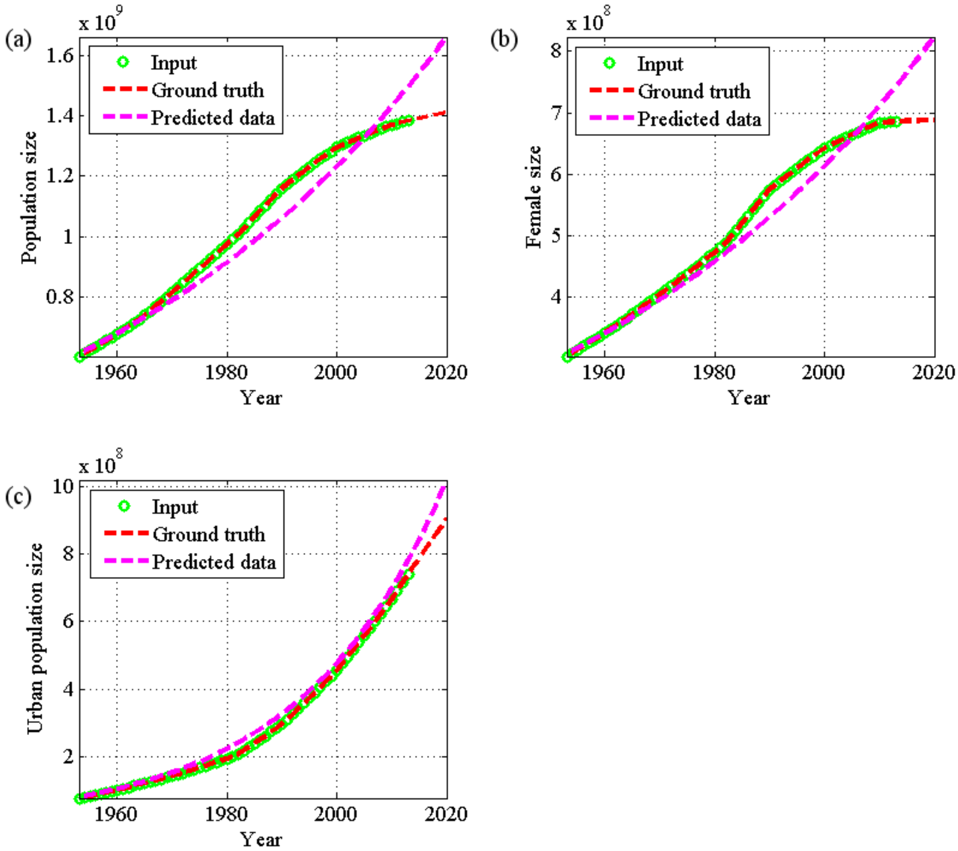

Figure 3. Estimated results of the Malthus model

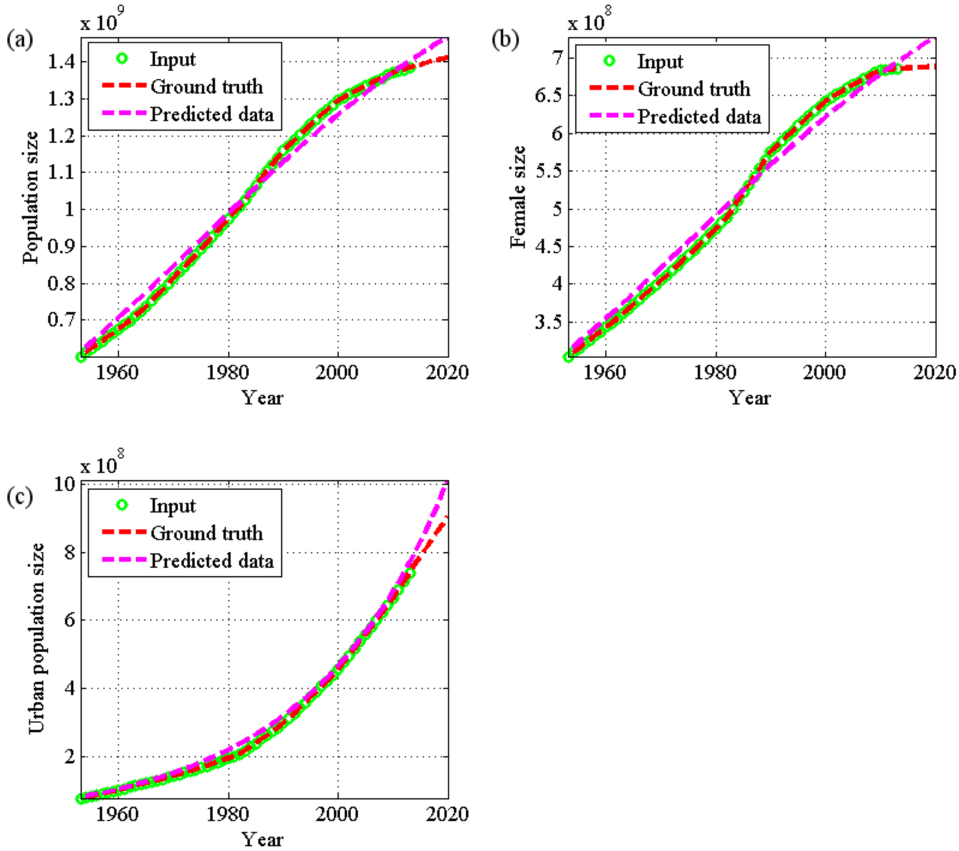

Figure 4. Estimated results of the Logistic model (using linear regression to predict its parameters)

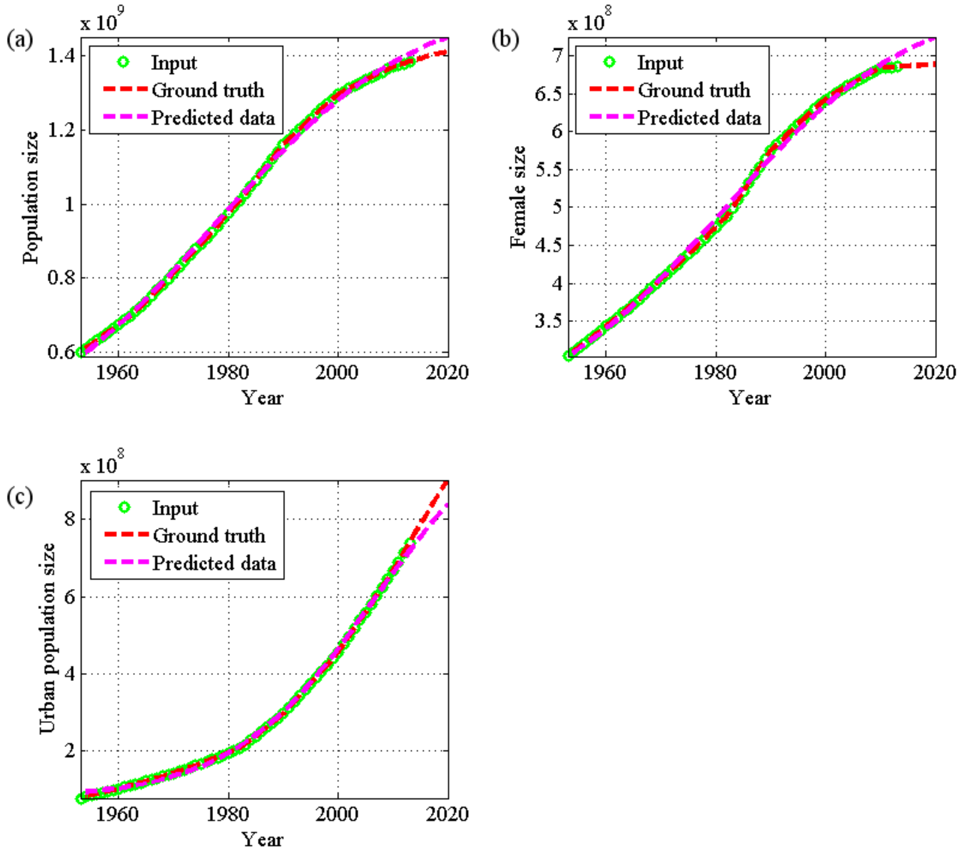

Figure 5. Estimated result of the Logistic model (using BMO to predict its parameters)

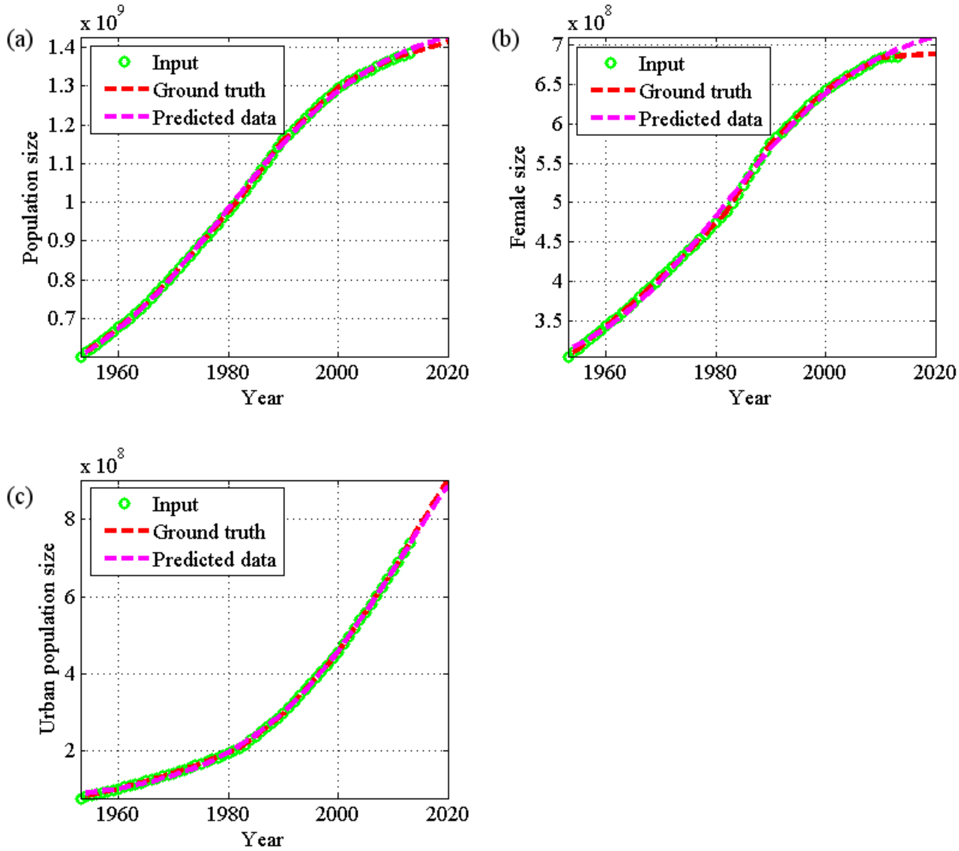

Figure 6. Estimated results of the Logistic model (using SCA to predict its parameters)

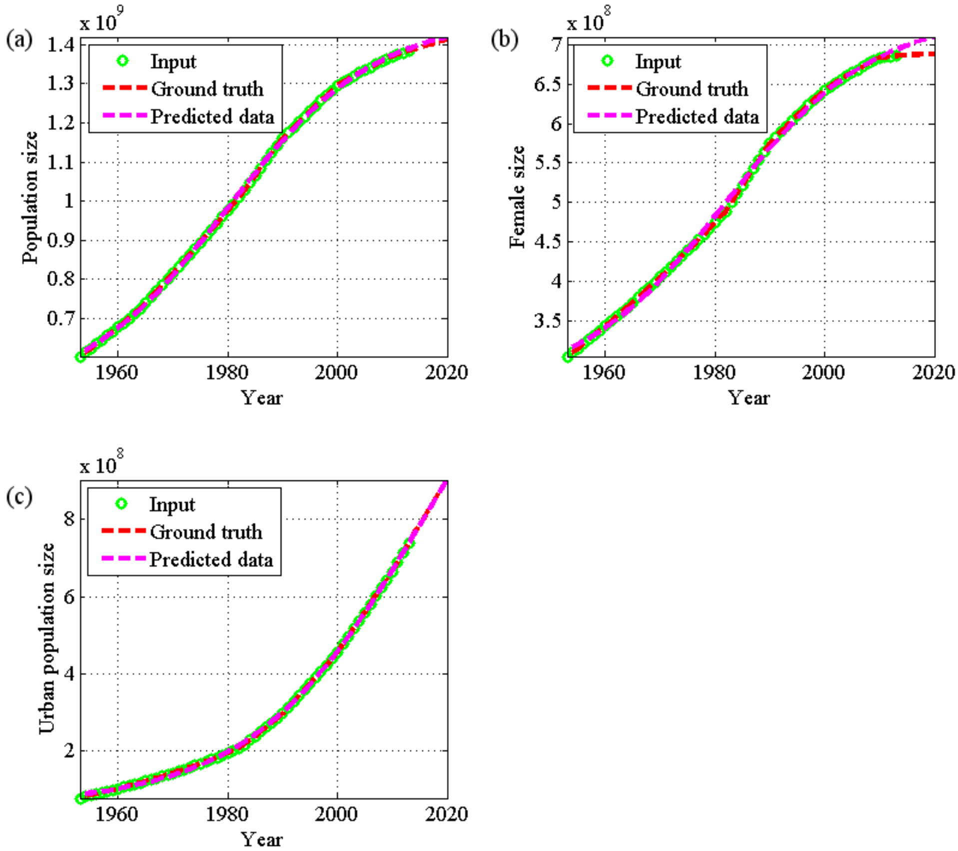

Figure 7. Estimated results of the Logistic model (using HGS to predict its parameters)

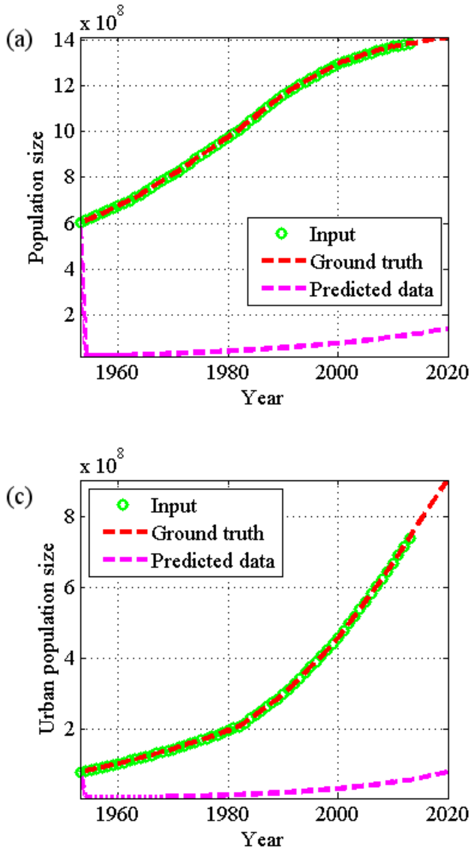

Figure 8. Estimated results of the GM (1,1) model (using least-square theorem to predict its parameters)

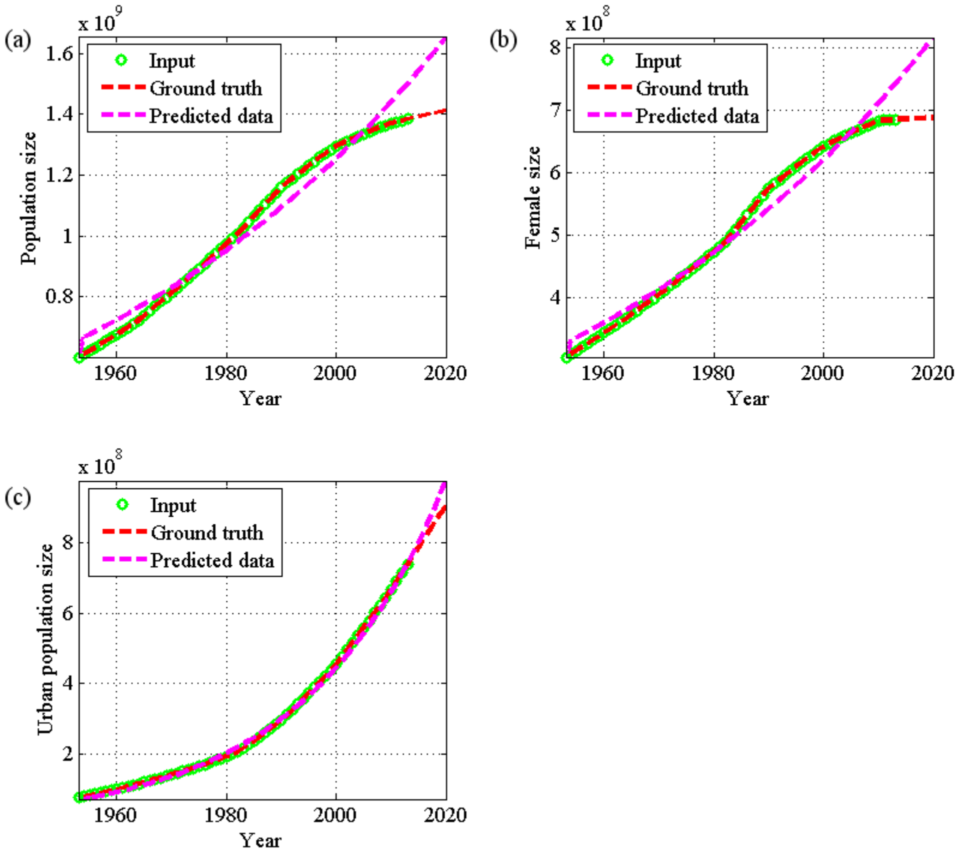

Figure 9. Estimated results of the GM (1,1) model (using linear regression to predict its parameters)

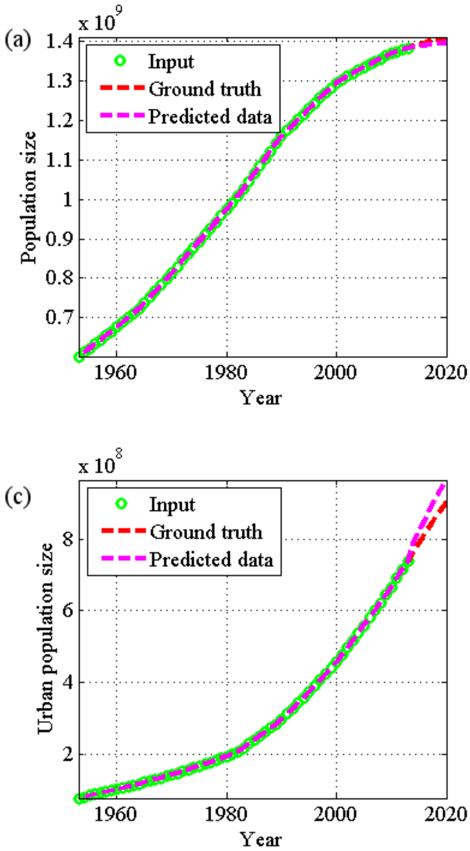

Figure 10. Estimated results of the BP neural network

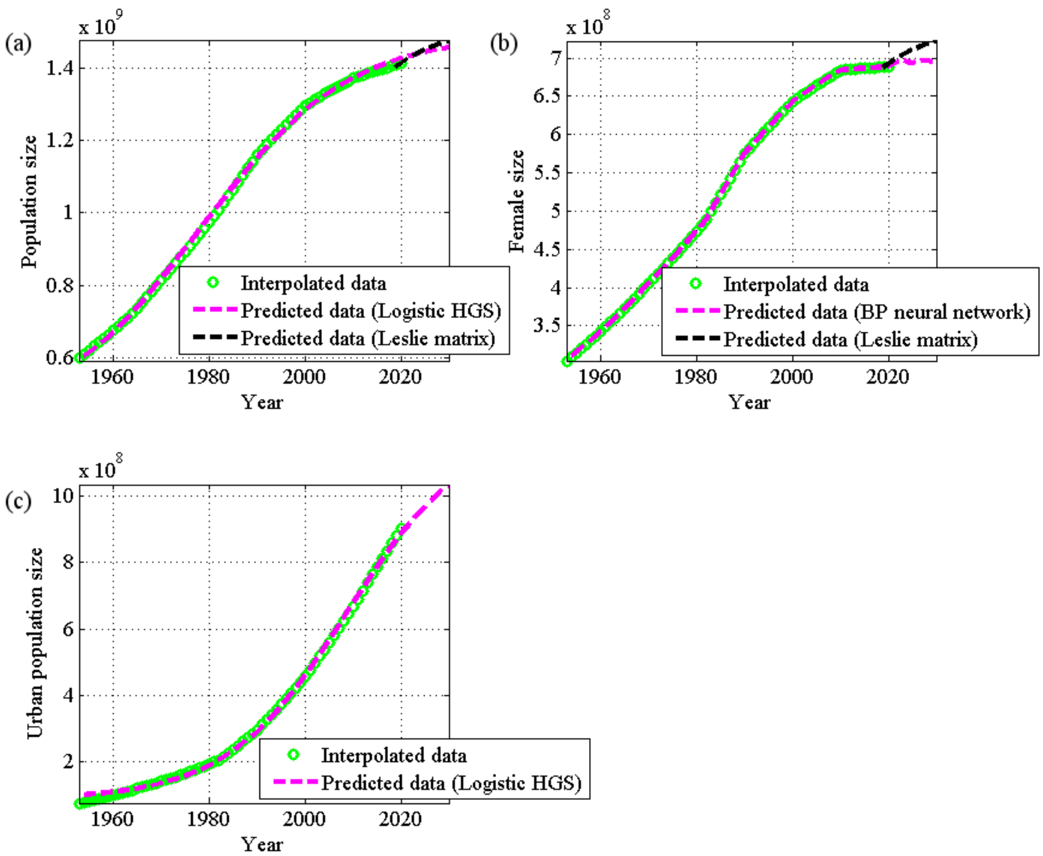

Figure 11. Predicted results of the population quantity, urban population size, and female size using selected models

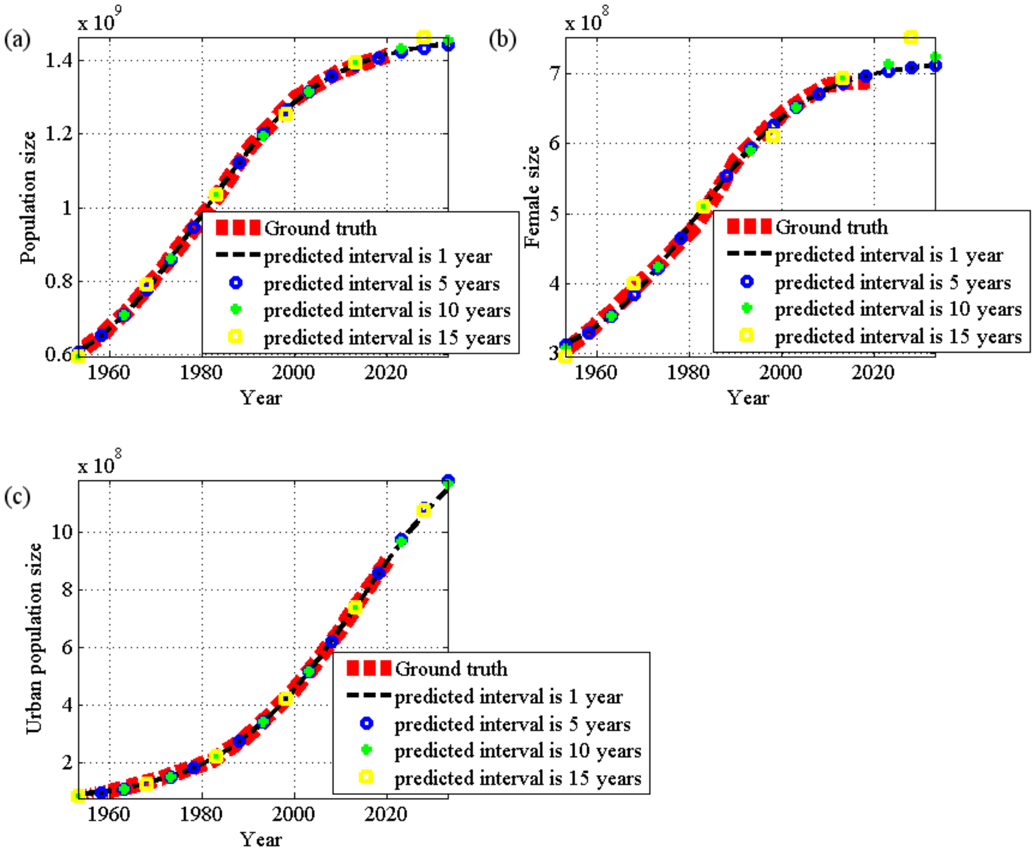

Figure 12. Predicted results generated from Logistic_HGS utilizing different sets of resampled data

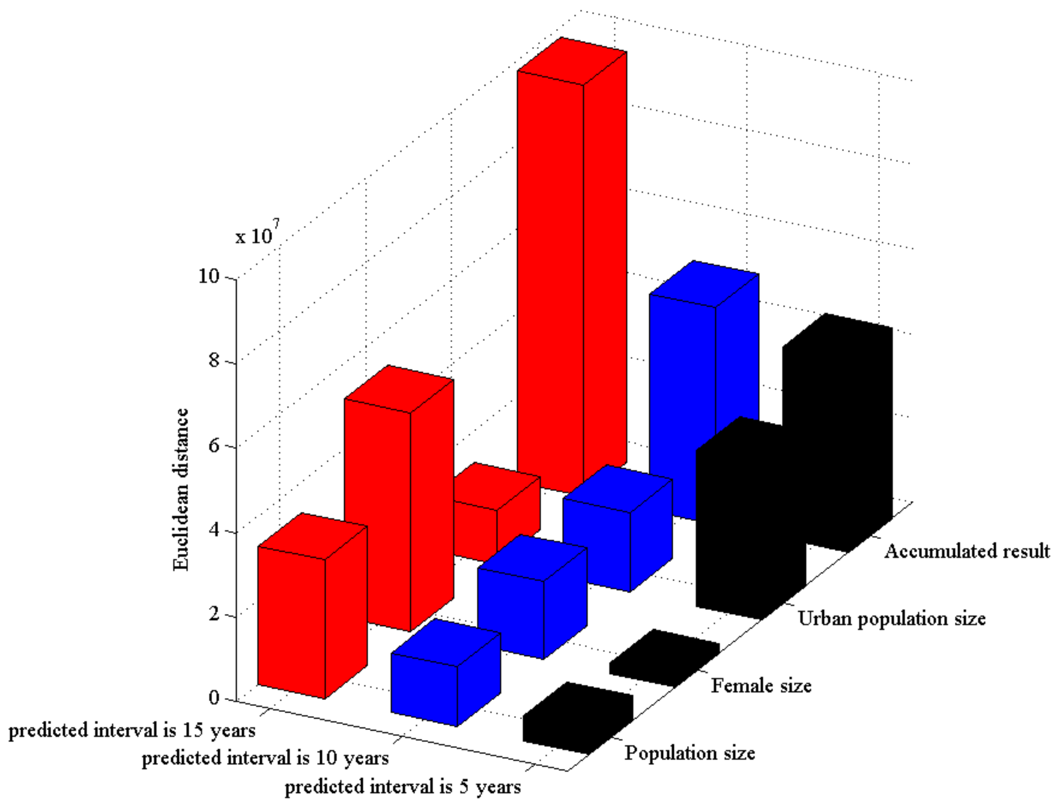

Figure 13. Calculated Euclidean distances between the estimated results based on different sampling intervals including 5, 10, and 15 years and the estimated results based on the originally interpolated data

LIST OF TABLES

TABLES

Table 1. Census datasets

Year	Population size	Female size	Urban population size
1953	601938035	304384517	77257282
1964	723070269	366553268	/
1982	1008175288	488741919	206588582
1990	1160017381	575067459	296512111
2000	1295990000	642440000	455940000
2010	1370536875	683684303	665570000
2020	1411780000	688440044	901990000

Table 2. Calculated parameters of the Logistic model using different metaheuristics

		Population size	Female size	Urban population size
Logistic_BMO	x_m	1226835737.05643	579317403.161534	1035840248.31665
	r_0	0.0543339457827691	0.0561453673294585	0.0767799294200552
	x_0	242412283.229701	94852543.0965095	15425779.4715406
	constant	336371653.654472	204005240.959116	75074909.7215017
Logistic_SCA	x_m	1022647020.60174	472382504.450131	1310053615.29322
	r_0	0.0692474235894750	0.0745933379393771	0.0681185995863809

	x_0	133272226.862572	45451660.5950816	20947709.2913278
	constant	461841115.228147	264067524.958084	66920910.6420349
Logistic_HGS	x_m	977422324.941170	469940839.592002	1477279984.04099
	r_0	0.0737735354994236	0.0751717999038542	0.0641605688236286
	x_0	111320650.440045	44437651.6383523	24495279.4338506
	constant	490354799.423473	265479718.537339	61206912.0057208

Terminology: Logistic_BMO, Logistic_SCA, and Logistic_HGS are the abbreviated name of the Logistic model combined with BMO, SCA, and HGS, respectively.

Table 3. Estimated parameters of the Malthus model, Logistic model and GM (1, 1) model using linear regression analysis or least-square theorem

	Logistic_linear		GM (1, 1)	
	x_m	r_0	c	v
Population size	1810507418.46249	1.74764308712604e-11	-0.0305549758340000	7.82409867659179e-13
Female size	918802914.115348	3.26470591207721e-11	-0.0305326516835398	1.57755780273090e-12
Urban population size	-13474211921.3205	-2.73472715794872e-12	-0.0462183111036132	4.67928128451738e-12
	GM (1, 1)_linear		Malthus_linear	
	c	v	r	

Population size	-0.0137457026443384	653959823.274615	0.0149392262127867	
Female size	-0.0136701011676947	325121242.516460	0.0146291117288232	
Urban population size	-0.0392284240109878	68753724.5225459	0.0379353117099909	

Terminology: Logistic_linear, GM (1, 1)_linear, and Malthus_linear mean that involved parameters of the Logistic model, GM (1, 1) model, and Malthus model were acquired using linear regression analysis. Especially, GM (1, 1) simply means that we utilized the least-square theorem to obtain relevant parameters of the GM (1, 1) model.

Table 4. Evaluating the prediction results of all categories using different algorithms

	Malthus_linear			Logistic_linear		
	error	*Delta_mean*	*C*	*error*	*Delta_mean*	*C*
Population size	683251154.313277	0.0520006536637692	0.205620926343877	235806314.333430	0.0264774412266669	0.0449531261347272
Female size	331497423.052205	0.0454167229169267	0.224697884410014	135786990.181518	0.0295529019382066	0.0524019250009606
Urban population size	275449902.632817	0.0730920986096183	0.0903410184834129	226474440.981463	0.0553438480582697	0.0829926353667494
	Logistic_BMO			Logistic_SCA		
	error	*Delta_mean*	*C*	*error*	*Delta_mean*	*C*

Population size	116476068.747824	0.0104273843451292	0.0311789235578591	52868111.1658166	0.00451137182391106	0.0144563248101736
Female size	90336920.0603068	0.0132596726714775	0.0602167993086624	57573710.3964426	0.0100533485548502	0.0355079121973712
Urban population size	133359184.053517	0.0326412264299198	0.0535824245451196	45610960.2234501	0.0217006842193835	0.0152595734951213
	Logistic_HGS			GM (1,1)		
	error	*Delta_mean*	*C*	*error*	*Delta_mean*	*C*
Population size	40529082.8735916	**0.00401528543355469**	**0.00962019641384119**	8451615646.75465	0.933421511238396	0.971515373420492
Female size	56874816.2293786	0.0100183229668748	0.0349603530925506	4178882754.37571	0.932429804569333	0.971456345034450
Urban population size	**30174627.3969463**	**0.0187059064292188**	**0.00909491610330343**	3248493586.84607	0.923530284439944	0.923355473774097
	GM (1,1)_linear			BP neural network		
	error	*Delta_mean*	*C*	*error*	*Delta_mean*	*C*
Population size	617970790.857473	0.0496816480937559	0.189956929316286	**22372188.3154398**	0.00447923080889657	0.522365868166679
Female size	307671074.4	0.043503525	0.2126103	**10057728.9**	**0.00478956**	1.53326043

	13999	2203313	60312147	**732318**	**656146885**	846315
Urban population size	131991800.363803	0.0379954623446879	0.0481885517611223	126470334.078822	0.0478095835499668	0.360693664960009

Table 5. Predictions of the population quantity, urban population size, and female size using selected models

Year	Population size	Female size	Urban population size
2021	1428648050	691354468	902685058
2022	1432458738	694825169	919910450
2023	1436060121	695674394	936452939
2024	1439462257	694051409	952301185
2025	1442674886	693181734	967449153
2026	1445707420	694630333	981895722
2027	1448568934	696549009	995644269
2028	1451268167	696556729	1008702219
2029	1453813514	694989072	1021080578
2030	**1456213032**	**694268722**	**1032793473**

Table 6. Predicted results of the Leslie model

Year	Population size	Female size
2018	1391546195	682130487
2019	1401749330	687132024

2020	1411276227	691802072
2021	1420142712	696148388
2022	1428311787	700152836
2023	1435826055	703836301
2024	1442747369	707229102
2025	1449070514	710328683
2026	1454852970	713163220
2027	1460088953	715729879
2028	1464841848	718059729
2029	1469183152	720187819
2030	**1473073295**	**722094752**

APPENDIX

Table 7. Publicly published datasets of the NBS (in 10,000)

Year	Population size	Female size	Urban population size
1949	54167	28145	5765
1950	55196	28669	6169
1951	56300	29231	6632
1955	61465	31809	8285
1960	66207	34283	13073
1965	72538	37128	13045
1970	82992	42686	14424
1971	85229	43819	14711
1972	87177	44813	14935
1973	89211	45876	15345
1974	90859	46727	15595
1975	92420	47564	16030
1976	93717	48257	16341
1977	94974	48908	16669
1978	96259	49567	17245
1979	97542	50192	18495
1980	98705	50785	19140

1981	100072	51519	20171
1982	101654	52352	21480
1983	103008	53152	22274
1984	104357	53848	24017
1985	105851	54725	25094
1986	107507	55581	26366
1987	109300	56290	27674
1988	111026	57201	28661
1989	112704	58099	29540
1990	114333	58904	30195
1991	115823	59466	31203
1992	117171	59811	32175
1993	118517	60472	33173
1994	119850	61246	34169
1995	121121	61808	35174
1996	122389	62200	37304
1997	123626	63131	39449
1998	124761	63940	41608
1999	125786	64692	43748
2000	126743	65437	45906
2001	127627	65672	48064
2002	128453	66115	50212

2003	129227	66556	52376
2004	129988	66976	54283
2005	130756	67375	56212
2006	131448	67728	58288
2007	132129	68048	60633
2008	132802	68357	62403
2009	133450	68647	64512
2010	134091	68748	66978
2011	134735	69068	69079
2012	135404	69395	71182
2013	136072	69728	73111
2014	136782	70079	74916
2015	137462	70414	77116
2016	138271	70815	79298
2017	139008	71137	81347
2018	139538	71351	83137

Table 8. Interpolated results of Kriging

Year	Population size	Female size	Urban population size
1953	601938034	297553517	77257281
1954	611977037	302352507	80603078

1955	622230002	307261093	83987356
1956	632693967	312281232	87413568
1957	643363724	317414396	90885961
1958	654232109	322661547	94409284
1959	665290381	328023107	97988506
1960	676528644	333498913	101628589
1961	687936314	339088157	105334355
1962	699502625	344789305	109110435
1963	711217146	350600005	112961276
1964	723070268	356517001	116891199
1965	738027772	365096588	120904468
1966	753108411	373773086	125005370
1967	768306486	382540420	129198287
1968	783617443	391391720	133487755
1969	799037587	400319467	137878509
1970	814563785	409315630	142375501
1971	830193244	418371804	146983909
1972	845923366	427479318	151709118
1973	861751687	436629320	156556673
1974	877675857	445812845	161532228
1975	893693622	455020874	166641460
1976	909802753	464244396	171889982

1977	926000871	473474485	177283239
1978	942285118	482702408	182826419
1979	958651625	491919762	188524358
1980	975094806	501118664	194381479
1981	991606521	510291949	200401737
1982	1008175287	519433369	206588581
1983	1027253425	527784410	217219825
1984	1046353823	536094275	228022910
1985	1065453455	544359843	238999582
1986	1084526583	552578769	250150992
1987	1103546002	560749174	261477764
1988	1122484464	568869284	272980140
1989	1141316053	576937068	284658199
1990	1160017380	584949921	296512110
1991	1174454359	592154293	311639087
1992	1188724828	599296080	326943332
1993	1202816329	606370232	342426240
1994	1216719827	613371128	358089549
1995	1230429165	620292913	373935142
1996	1243940518	627129907	389964792
1997	1257251933	633877083	406179880
1998	1270363040	640530581	422581119

1999	1283274904	647088226	439168284
2000	1295990001	653550000	455939999
2001	1304182707	657156368	476198028
2002	1312187958	660674469	496634010
2003	1320012623	664111955	517242787
2004	1327664971	667478667	538018055
2005	1335154510	670786087	558952277
2006	1342491725	674046631	580036272
2007	1349687687	677272822	601258375
2008	1356753563	680476364	622603127
2009	1363700057	683667184	644049649
2010	1370536874	686852572	665569999
2011	1375032102	690572448	689951699
2012	1379432797	694291483	714326050
2013	1383744426	698006982	738639994
2014	1387971493	701713610	762833936
2015	1392117998	705404210	786844088
2016	1396187833	709070683	810605309
2017	1400184993	712704767	834054255
2018	1404113544	716298682	857132704
2019	1407977401	719845598	879790817
2020	1411779998	723339956	901989999