



## Research on Prediction of Shanghai's Population Development From 2008 to 2050

Kepei Men

College of Mathematics & Physics

Nanjing University of Information Science & Technology

No.219 Ningliu Road 210044, China

E-mail: menkp@yahoo.com.cn

Ping Zhou

College of Mathematics & Physics

Nanjing University of Information Science & Technology

No.219 Ningliu Road 210044, China

E-mail: zhoujas929@sina.com

Jing Liu

College of Mathematics & Physics

Nanjing University of Information Science & Technology

No.219 Ningliu Road 210044, China

E-mail: mi\_lu\_er@sina.com

Liangyu Jiang

College of Mathematics & Physics

Nanjing University of Information Science & Technology

No.219 Ningliu Road 210044, China

E-mail: elanjiang1110@163.com

### Abstract

Based on the latest statistic data of Shanghai's population during 1952-2007, a grey model with the pretreatment of data is established for prediction and researching on the Shanghai's future household registration population up to the year 2050. This model's feasibility is illustrated compared with other grey models. The result shows that at the end of 2008, 2010, 2020, 2030, 2040 and 2050 Shanghai's population will reach 13.868, 14.047, 14.974, 15.962, 17.016, 18.138 million, respectively.

**Keywords:** Shanghai's population, Grey model, Population prediction

### 1. Summary of the development of Shanghai's population

Shanghai, the largest city in China, is the economic, financial, technology and communication centre of this country. It is not only one of the most important industrial bases, the major ports and trade centers, but also the flagship of the Yangtze delta city group. As the famous international metropolis, Shanghai's development is having an effect on the whole national economy and pushing it forward. As is well known, Shanghai has a large population and the highest population density in China. Based on the statistics data, by the end of 2007, Shanghai's resident population is 18,580,000, which is 429,200 more than last year, and its household registration population is 13,788,600, which is 2.6 times of the number in the early period after the founding of New China and 1% of national total population(NBS 2007). According to the household registration population, the population density is 2175 people / km<sup>2</sup>, Shanghai has 1.04% population but only 0.06% land area of the whole country. Therefore, the research on prediction of Shanghai's population development not only has the important and realistic meaning to the native social and economic development, but also is pertinent and typical for the whole country.

At the beginning of 1950s, Shanghai's total population increased at a high speed. In the continuous 7 years from 1952 to 1958, the population natural growth rate has been keeping above 30%os, and the highest record once reached to 45.6%os. The main reason includes: Shanghai gets into the first five-year plan in 1953, the city needs large quantities of labor

force for some important projects. And because the population migration wasn't limited strictly in policy, the population in Shanghai grew fast in short term. By the end of 1950s, the natural growth rate of population in Shanghai City decreased sharply and kept gradually. It was at a period of low increase. Been influenced by the natural disaster, the natural fluctuation of population sunk to a low ebb, and then rallied slowly. Except 1969, population natural growth rates were annually below 10‰ from 1966 to 1980. Because of the baby boom in 1950s and the surplus period, the population natural growth rate had a certain raise in the early 1980s. But because the family planning policy was carried out strictly in 1970, it could still maintain at a low level. The population natural growth rate was basic stable and the total population rose slowly. It appears a negative population growth rate in Shanghai since 1993. This is inseparable with the result of the family planning policy. According to the statistics, from the starting of family planning policy to the beginning of 2005, the total population in Shanghai is born 7,000,000 less, nearing the half of total population in Shanghai, and it makes a great contribution for national population development and control. Based on the latest statistics data, this paper researches on the Shanghai's future household registration population development by building up the improved grey model dealt with the pretreatment of data, in order to provide reliable evidence for the social and economic sustainable development.

## 2. Prediction of Shanghai's population

### 2.1 Establishment of grey model

Supposing the original population time sequence is

$$x^{(0)} = [x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)]$$

and the time sequence  $x^{(0)}$  through 1-AGO (accumulating generation operator) to get a newly-generated sequence  $x^{(1)}$ , namely,

$$x^{(1)} = (x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n))$$

in which

$$x^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i) \quad (k = 1, 2, \dots, n)$$

with which to make a general GM (1, 1) prediction model, namely, grey differential equation (image equation)

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

dispensing the above equation to establish grey differential equation

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

in which

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

where  $a$  and  $b$  are coefficients to be determined, called a development coefficient and grey actor, respectively.  $a$  has its valid limits (-2, 1) and is obtained through the following expression via the least squares method

$$\hat{a} = (a, b)^T = (B^T B)^{-1} \cdot B^T \cdot Y_{n-1}$$

in which

$$Y_n = \begin{pmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{pmatrix}, \quad B = \begin{pmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{pmatrix}$$

The solution is the time response function in the form

$$\begin{cases} \hat{x}^{(1)}(k+1) = [x^{(1)}(m)d - \frac{b}{a}] \cdot e^{-a(k+1-m)} + \frac{b}{a} \\ \hat{x}^{(0)}(k+1) = (1 - e^a)[x^{(1)}(m)d - \frac{b}{a}] \cdot e^{-a(k+1-m)} \end{cases}$$

Not all raw data are used for establishing a gray model and different dimension (or length) produces different values of  $a$

and  $\hat{b}$ , leading to different predictions that constitute a prognostic grey interval. According to grey modeling theory, one-step forecast precision is above 98% and two to five step forecast precisions are above 97% as  $|a| \leq 0.3$ , while one to two step forecast precisions are above 90% and ten-step forecast precision is above 80% as  $0.3 < -a \leq 0.5$  (Liu et al., 1999). Thus, to improve the prediction accuracy we have to select a grey model of a suitable number of dimensions (Liu et al. 2004, 1999).

The grey system theories have two basic principles: new information priority and small-sampled and poor-information. Based on these two principles, we take  $m = 1, 2, \dots, n$  in turn to construct prediction equation and compute prediction error. We choose the value  $m$  from the models which has the smallest average error and then we establish the best grey optimal model (Zhang, 2002). When  $m = 1$ , the model we establish is general-type GM. If we apply one-order difference operation to the raw data, and then establish general grey model, this model we call: grey increment model (Men et al. 2004, 2005). In order to improve the prediction accuracy, this paper will first carry on preparing a processing to the original data before making use of above-mentioned models to carry on an estimate. Generally, there are two methods to prepare a processing, one kind opens the data square, another one turns data logarithm to make the data smooth and be suitable for grey model. After getting estimate values, we have to restore the data. This paper will make use of above methods to build up models respectively and carry on more analysis to the results we get.

## 2.2 Prediction and analysis of the future population of Shanghai

### 2.2.1 General grey prediction models and test

According to the data of table 1, this paper adopts nonlinear least squares fitting method, building up following GM models respectively.

#### (1) General-type GM (1, 1)

To select a suitable model, we single out 4-10 dimensions short series to construct models of general-type GM (1, 1) and carry on experimental prediction to the data of 2007 Shanghai City population. Test shows that the 5-dimension model gives the closest result, so we have general-type GM (1, 1) (2002-2006):

$$\hat{x}(k+1) = 209107.548391 \exp(0.0064k) - 207773.318391$$

#### (2) Grey increment model

To select a suitable model, we single out 4-10 dimensions short series to construct models of grey increment model and carry on experimental prediction to the fact value of 2007 Shanghai City population. Test shows that the 6-dimension model gives the closest result, thus we have grey increment model (2001-2006):

$$\hat{x}(k+1) = 386.945335 \exp(0.020115k) - 381.435335$$

#### (3) Grey optimal model

To select a suitable model, we single out 4-10 dimensions short series to construct models of grey optimal model and carry on experimental prediction to the fact value of 2007 Shanghai City population. Test shows that the 5-dimension model gives the closest result, thus we have grey optimal model (2002-2006):

$$\hat{x}(k+1) = 209107.566547 \exp(0.006400k) - 207773.318391$$

### 2.2.2 Grey model in which data is pretreated and test

#### (1) Turn data logarithm to construct GM (1, 1)

To select a suitable model, we single out 4-10 dimensions short series which are turned logarithm to construct models of grey model and carry on experimental prediction to the fact value of 2006 and 2007 Shanghai City population. Test shows that the 5-dimension model gives the closest result, which is thus taken for use.

$$\text{Model for 2006: } \hat{x}(k+1) = 7870.735399 \exp(0.000914k) - 7863.544616$$

$$\text{Model for 2007: } \hat{x}(k+1) = 8105.660402 \exp(0.000888k) - 8098.464290$$

#### (2) Open the data square

To select a suitable model, we single out 4-10 dimensions short series which is opened square to construct models of grey optimal model and carry on experimental prediction to the fact value of 2006 and 2007 Shanghai City population. Test shows that the 5-dimension model gives the closest result, which is thus taken for use.

$$\text{Model for 2006: } \hat{x}(k+1) = 11573.104381 \exp(0.003142k) - 11536.750001$$

$$\text{Model for 2007: } \hat{x}(k+1) = 11435.194206 \exp(0.003189k) - 11398.764119$$

The predictions of 2006 and 2007 population of Shanghai from the above models are listed in Table 2 and Table 3.

Comparison of tests in table 2 and table 3 indicates that if we carry on preparing a processing to the original data firstly,

the model has higher prediction accuracy than other grey models. Furthermore, it is superior to other models in some respects: ① Higher prediction accuracy is maintained in extended and long-range predictions (far exceed other ordinary models); ② No large quantities of data is demanded in collection [four to eight samples can be chosen for model establishing, which is especially fit for the occasion that data are difficult to get.]; ③ It is flexibility and handy in model operation [with small calculation]. Consequently, it is an ideal and economic new tool for population prediction.

According to table 2 and table 3, we prepare a 5-dimension grey model after opening the data square on the data from 2003 to 2007 because this kind of model has the highest accuracy,

$$\hat{x}(k+1) = 11487.621011 \exp(0.003196k) - 11450.990707 \quad (*)$$

Tests show for Model (\*) at  $C = 0.05, P = 1$  we have the mean fitting precision  $q = 99.949\%$ . The effect of prediction is quite satisfactory. Therefore, this model satisfies first-grade accuracy requirement for the use of the extended and long-range prediction of the population of Shanghai, with the predictions shown in Table 4.

### 3. Conclusions and discussions

(1) The accurate estimate of population is the basis of a city's development program. There's an erroneous view that a good model must be complicated. In fact, the simple and practicable model awarding with facts is the best one for prediction. In this paper, the raw data series are dealt with the extraction while building up the prediction model of Shanghai's population, in order to improve the precision of prediction. It's proved that, precision of prediction model raises consumedly after the extraction of the initial series. It is a kind of more viable way to predict. The demonstration proves that this is an effective way with economical and practical facilities on population prediction, because it's agility, convenient, small calculation amount and small sample. It can be predicted in this paper that, household register population in Shanghai will grow smoothly and slowly in the coming 40 years, and annually the population net growth won't over 100,000. While holding the World's Fair in 2010, population in Shanghai will attain 14,047,000 or so; and in 2050, the population is about 18,140,000.

(2) The development of Shanghai is open, and the floating population comes out at the top of the list in major cities. It attracts more and more domestic and international population since 1992, a great deal of labor force come from other parts of country gets into Shanghai for the employment opportunities. Because the population natural growth rate has continuously presented a negative growth for 15 years, the estimate result in this paper points out that, the growth of household register population will be stability. As a result, now and in a long period of time to come, the key factor which will influence the population scale and structure change in Shanghai is ab extra floating population. Ab extra floating population in Shanghai is 1,060,000 in 1988, which has already doubled to 2,510,000 in 1993. The fifth census data shows that, Shanghai's resident population is 16,400,000 in 2000, among which, the ab extra floating population is 3,870,000, 23.6% around. Currently, floating population who live in Shanghai for more than half a year is about 1/4 of the permanent population(Xinhua, 2006). It is predicted that, in 2010 Shanghai's population will be 20,000,000, and then the ab extra floating population will be 6,000,000 or so. The elements which can influence the floating population are: ① The upsurge of infrastructure construction caused by the World Fairs; ② The urbanization strategy of countryside which is needed by the economy development of Yangtze River delta and manufacturing industry belt; ③ The pressure and responsibility which is brought by urbanization of cities in China; ④ The need of labor force in some service industries caused by the reduction of labor force and aging population in Shanghai around 2010. Therefore, we must tightly hold the strategic opportunity of development to make the full use of the human resource in ab extra floating population, in order to strengthen the comprehensive competition ability of Shanghai and promote the great-leap-forward development of Shanghai in the new century.

(3) Population is the foundation of a city's development. When investigating the relationship between population and social development, there is the chief question that how many are the optimum population and the limitation of Shanghai population with the total area 6340.5 km<sup>2</sup>? The latest research of the population institute in Fudan University replies: the limitation of Shanghai population is about 28,000,000-30,000,000, and around 2020 the population will be 22,000,000-24,000,000. Since the middle of 1990's, many research centers and experts study up on the optimum population in Shanghai by various ways. Although the results of research are different, but on the whole, the optimum population of Shanghai should be 18,000,000-20,000,000 in 2010. Shanghai locates in the lower end of Yangtze River delta, natural resources and environment would not restrict the capacity of population. Without consideration of other social factors, the population in Shanghai attaining 20,000,000 wouldn't bring any negative influence to the citizen's living quality(Peng, 2002). However, if there is not large addition of new population, it will weaken the adjustment of the quality and structure of population and the competitive of sustainable development in Shanghai. To consider the composition of population, the negative factors of city development are: ① The structure of labor power in Shanghai is increasingly aging. The relevant personage points out many times that the labor force population in household register population in Shanghai will descend gradually after 2005, and hit the lowest level around 2010, then, the economic activities population of Shanghai will reduce from 8,500,000 to 6,500,000. ② The society evolves from aging of

population to aging of the aged, Shanghai will face double pressures from the aging society and old aged society. ③ The age structure of population is characterized by Multi-Peak-Valley form. ④ The quality of population is relatively weak, which is not correspond to a modern international metropolis.

Zhang Rongzhou (Population, 2003) in Shanghai population information center said that the discussion of the population in Shanghai should not only consider the need of Shanghai's social economic development and capacity of "resources", but also consider it in the urbanization of China.

First, we should have a full understanding of Shanghai's position in the progress of the urbanization of china. Urbanization, it appears to be a huge process of population gathering first, accompanying with the gathering of industries and various resources. Since the industrial revolution, the development of modern city experienced 3 stages: First stage: urbanization; Second stage; especially big urbanization; Third stage: especially big urbanization groups. Today, the development of urbanization and metropolitan area is the common phenomenon in many developed countries. Take an example, Tokyo's population is 11,000,000 in 2000, which is less than Shanghai's population. But it constitutes 8.8% of the Japanese population, much higher than Shanghai. And the population of metropolitan area in Tokyo is 32,000,000, 25.6% of Japanese population. The population of Tokyo is less than Shanghai, but in fact, the degree of its dense population is far above Shanghai.

Second, "Resources" isn't the bottleneck of Shanghai's population development. According to the data of Chinese urban net, a per-capita floor space in Shanghai is 48.3 m<sup>2</sup>, and it is 64.3 m<sup>2</sup> in Tianjin, 70.9 m<sup>2</sup> in Beijing, 76 m<sup>2</sup> in Guangzhou. In the world, it is 195 m<sup>2</sup> in Hamburger, 136 m<sup>2</sup> in Munich, 116 m<sup>2</sup> in London. The experts suggest that the indicator of urban construction land should appropriately rise along with the development of Shanghai economy, and accelerate the process of urbanization. Concern about the problem of water, many researches point it out early that Shanghai isn't in a water shortage, but is in need of clean water. Shanghai should improve its ability of water resources protection and wastewater treatment. Now, Shanghai's resources are conditioned by three main factors: technique, management and devotion ability. "Resources" will not become the bottleneck of Shanghai's population development in the foreseeable future.

(4) To sum up, Shanghai has to keep the social and economy developing continuously, and raise comprehensive ability of competition. It has to build a good population environment, speed up favorable recruiting and population flowing within the next 10 years, steadily carry out and adjust the family planning policy, and program its future development on the strategy of Shanghai metropolis area. It should take the expansion of population in short time as the cost of social and economy development. In the course of the city population's exchange and adjustment, Shanghai will promote its comprehensive strength and competitiveness, and accelerate the social and economic development. Shanghai will be striding toward the modern international metropolis with a completely new shape.

## References

- Liu Sifeng, Dang Yaoguo, and Fang Zhigeng, (2004). *Grey System Theory and Its Application* (3rd edition), Beijing: Science Press.
- Men Kepei, Guan Lingling and Jia Lingyun, (2005). China's future population: predictions and prospects. *Population Review*, 44(1), 1-10.
- Men Kepei and Zeng Wei, (2004). Study on the Prediction of the Population of China over the Next 50 Years. *The Journal of Quantitative & Technical Economics*, (3), 12-17.
- National Bureau of Statistics of China, (2007). *China Statistical Yearbook-2007*. Beijing: China Statistics Press.
- PRC State Council, (2000). *Population and Development for the 21st Century in China*. [Online] Available: <http://news.xinhuanet.com> (Oct18, 2007)
- Peng Xizhe and Ren Yuan,(2002). Population factor and city synthesize a competition ability-aresearch developed to city in Shanghai. *Chinese Population Science*, (6), 33-39.
- Population, (2003). The increase of total population in Shanghai is the result of the development of economic and city. [Online] Available: <http://www.popinfo.gov.cn> (Feb1, 2008)
- Shanghai Statistics Bureau,(1991-2007). *Statistical yearbook of Shanghai*. Beijing: China Statistics Press.
- Shanghai Statistics Bureau, (2008). *Statistical Communique of Shanghai on the 2007 National Economic and Social Development*. [Online] Available: <http://www.stats-sh.gov.cn> (Feb4, 2008)
- Xinhua News Agency, (2006). Shanghai: "floating resident population" is a quarter of resident population. *Jingling Evening Paper*,( Nov19, 2006)
- Zhang Dahai, Jiang Shifang and Shi Kaiquan, (2002). Theoretical Defect of Grey Prediction Formula and Its Improvement. *Systems Engineering-Theory & Practice*, 22(8), 140-142.

Table 1. Statistics of Shanghai's population in 1952-2007 (Unit: 10<sup>4</sup> Persons)

Year	Household Registration Population	Birth		Death		Natural Growth	
		number	birth rate(‰)	number	death rate(‰)	number	natural growth rate (‰)
1952	572.63	21.98	39.1	4.93	8.8	17.05	30.3
1953	615.24	25.61	43.1	5.26	8.9	20.35	34.2
1954	662.71	33.70	52.7	4.55	7.1	29.15	45.6
1955	623.10	27.20	42.3	5.28	8.2	21.92	34.1
1960	1056.30	28.91	27.7	7.16	6.9	21.75	20.8
1965	1093.79	18.58	17.0	6.20	5.7	12.38	11.3
1970	1072.55	15.10	13.9	5.41	5.0	9.69	8.9
1971	1066.82	13.03	12.2	5.6	5.2	7.43	7.0
1972	1064.11	11.54	10.8	5.93	5.6	5.61	5.2
1973	1070.01	10.94	10.2	5.84	5.5	5.10	4.7
1974	1073.78	9.85	9.2	6.25	5.8	3.60	3.4
1975	1076.72	10.14	9.4	6.47	6.0	3.67	3.4
1976	1081.30	11.04	10.2	6.63	6.1	4.41	4.1
1977	1086.47	11.72	10.8	7.05	6.5	4.67	4.3
1978	1098.28	12.36	11.3	6.82	6.2	5.54	5.1
1979	1132.14	13.76	12.3	6.81	6.1	4.41	6.2
1980	1146.52	14.31	12.6	7.39	6.5	6.92	6.1
1981	1162.84	19.38	16.8	7.44	6.4	11.94	10.4
1982	1180.51	21.68	18.5	7.35	6.3	14.33	12.2
1983	1194.01	17.80	15.0	8.19	6.9	9.61	8.1
1984	1204.78	16.38	13.7	7.82	6.5	8.56	7.2
1985	1216.69	15.43	12.7	8.10	6.7	7.33	6.0
1986	1232.33	17.75	14.5	7.93	6.5	9.81	8.0
1987	1249.51	19.02	15.3	8.27	6.7	10.75	8.6
1988	1262.42	16.53	13.2	8.47	6.8	8.06	6.4
1989	1276.45	15.91	12.5	8.43	6.6	7.48	5.9
1990	1283.35	13.12	10.2	8.63	6.7	4.49	3.5
1991	1287.20	10.08	7.8	8.56	6.7	1.52	1.1
1992	1289.37	9.37	7.3	9.10	7.1	0.27	0.2
1993	1294.74	8.40	6.5	9.40	7.3	-1.00	-0.8
1994	1298.81	7.63	5.9	9.42	7.3	-1.79	-1.4
1995	1301.37	7.11	5.47	9.79	7.53	-2.68	-2.06
1996	1304.43	6.79	5.21	9.77	7.50	-2.98	-2.29
1997	1305.46	6.42	4.92	9.57	7.33	-3.15	-2.41
1998	1306.58	6.17	4.73	10.13	7.75	-3.96	-3.03
1999	1313.12	6.56	5.01	9.54	7.28	-2.98	-2.27
2000	1321.63	6.95	5.27	9.45	7.17	-2.50	-1.90
2001	1327.14	5.76	4.34	9.34	7.05	-3.58	-2.71
2002	1334.23	6.20	4.66	9.67	7.27	-3.47	-2.61
2003	1341.77	5.73	4.28	10.07	7.52	-4.34	-3.24
2004	1352.39	8.09	6.00	9.65	7.16	-1.56	-1.16
2005	1360.26	8.52	6.08	10.23	7.54	-1.98	-1.46
2006	1368.08	8.12	5.95	9.80	7.19	-1.68	-1.24
2007	1378.86	10.08	7.34	10.22	7.44	-0.14	-0.10

Note: The above data are taken from the "Shanghai Statistical Yearbook" (1991-2007) and "Statistical Communique of shanghai on the 2007 National Economic and Social Development".

Table 2. Comparison of various models-based predictions for Shanghai’s population for 2006 (Units: 10<sup>4</sup> persons)

Model	Statistics	Prediction	Residual difference	Relative error (%)	Predict accuracy(%)
General-type GM	1368.1	1369.5	1.4	0.101	99.899
Grey increment model	1368.1	1370.5	2.4	0.175	99.825
Grey optimal model	1368.1	1369.5	1.4	0.101	99.899
Turn data logarithm	1368.1	1369.5	1.4	0.102	99.898
Open the data square	1368.1	1368.7	0.6	0.044	99.956

Table 3. Comparison of various models-based predictions for Shanghai’s population for 2007 (Units: 10<sup>4</sup> persons)

Model	Statistics	Prediction	Residual difference	Relative error (%)	Predict accuracy(%)
General-type GM	1378.9	1377.5	1.4	0.102	99.898
Grey increment model	1378.9	1376.8	2.1	0.152	99.848
Grey optimal model	1378.9	1377.5	1.4	0.102	99.898
Turn data logarithm	1378.9	1377.5	1.4	0.102	99.898
Open the data square	1378.9	1377.5	1.4	0.102	99.898

Table 4. The predictions of Shanghai’s population for 2008-2050 (Units: 10<sup>4</sup> persons)

Year	Total	Year	Total
2008	1386.8	2018	1478.4
2009	1395.7	2019	1487.8
2010	1404.7	2020	1497.4
2011	1413.7	2025	1546.0
2012	1422.8	2030	1596.2
2013	1431.9	2035	1648.0
2014	1441.1	2040	1701.6
2015	1450.3	2045	1756.8
2016	1459.6	2050	1813.8
2017	1469		

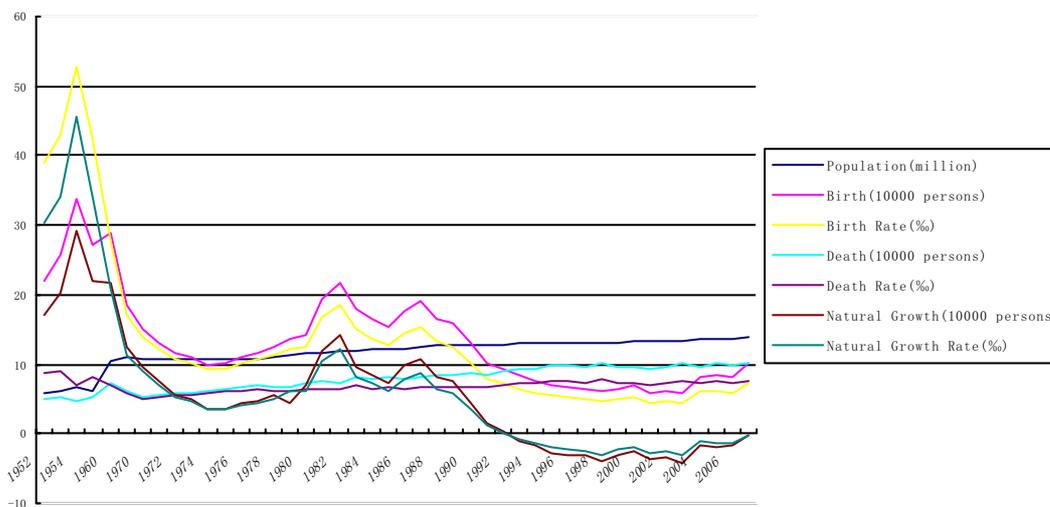


Figure 1. Natural change of Shanghai’s population from 1952 to 2007