Application of fuzzy time series to population forecasting

A.M. ABBASOV, M.H. MAMEDOVA

ABSTRACT

The potential of fuzzy logic application in simulating of demographic processes by the example of population forecasting task has been investigated. The particularities of population as dynamical system functioning under the condition of uncertainty have been examined and fuzzy statement problem has been suggested. The strategy of population forecasting using the method of fuzzy time series model has been proposed.

The simulations on retrospective evaluation of population are carried out and on the base of the results of these simulations the conclusion avocet the effectiveness of utilization of fuzzy model for demographic forecasting has been model.

1 INTRODUCTION

The development stage of society, characterized by the implementation of market mechanism into all the spheres of human activity, sets greater requirements upon the perspective calculation of forecasted total population. The demographic forecasting, by being an integral part of social-economic development, enables us to assess the anticipated total population, economically active population, the size of different age cohorts etc. These factors should be taken into consideration in formulating scientifically feasible social-economic development policy and preparing the complex practical measures aimed at the implementation of this policy.

Presently, there are different ways of dealing with the problem of forecasting total population and a great number of methods for handling it. Among them can be cited statistical, adaptive, imitation models, smooth dynamic series, auto regression etc. The development of information technologies and software resources has opened up new opportunities for modeling demographic processes and handling forecasting problems. Researches carried out over the recent years prove that the application of traditional analyzing methods and modeling of population growth process on the basis of processing numeric/quantitative data don't produce the desired results and even involve considerable risks and errors. One of the main causes for this undesirable circumstance stems from the fact that a great many forecasting models are not sufficiently efficacious owing to the incompatibility of highly accurate quantitative methods of classical mathematical apparatus with the great complexity of population growth process. The other cause, in our opinion, is that these methods aimed at the mathematical analysis of accurately determined systems are not capable to encompass certain characteristics of the research sphere.

Thus, population is a large dynamic (economic, social, ecological) system irrespective of a specific territory and definite group. A distinguishing feature of this system consists of its functioning under indefinite, uncertain conditions due to a multitude of causes. These uncertainties are, first of all, associated with: a) the impossibility of identifying all the factors that determine the development dynamics of human population; b) variability and inconstancy of boundaries of many indicators used in demographic analysis and considerable variations in the values of some indicators; c) the lack of comprehensive aprior information pertaining to demographic processes associated with the data source problems and impossibility of recording all the demographic events; To this list of causes can be added data incompleteness, uncertainties involved while collecting some indicators from various sources: results of population census and demographic researches, current registration of population movement etc. Discrepancies between official and unofficial data, expert evaluations can cause further difficulties.

Population growth is a multi-factored and time-dependent process. But it is not possible to consciously influence this process by varying certain parameters and observing changes in others. Various uncontrollable factors (wars, inter-ethnic conflicts, natural disasters, ecological factors etc.) significantly affect population growth. If systematic statistical material with regard to the concerned problem is lacking, empirical data becomes the sole information source.

Thus, keeping in mind that the demographic data are incomplete and accuracy of some or all the available data are questionable for any of several reasons, a demographic analysis based on this incomplete, inaccurate information bears a special significance. The vagueness, inaccuracy, incompleteness, fuzziness of the data on the demographic events and processes and the prevailing evaluation methods excluding this data characteristics necessitate taking a new approach to the analysis and evaluation of demographic situations, particularly, population growth forecasting.

Within the context of the above-mentioned arguments, exploring the possibilities of the application of the fuzzy sets theory or the apparatus of sets, going by the name of fuzzy logic, to modeling demographic processes bears a special interest.

2 THE FUZZY SETS THEORY AND FORECASTING

The fuzzy sets theory can be defined as a mathematical formalism that enables us to eliminate indefiniteness and deal with incomplete, inaccurate information of both qualitative and quantitative by nature. The fuzzy sets theory, advanced by L.Zadeh [1], one of the well-known representatives of modern applied mathematics, by excluding any definite description of the task offers such a solution scheme of the problem that a subjective reasoning and evaluation plays a principal role in evaluating indefinite, unclear fact. Thus anyone, encountering indefinite, incomplete information/data, can form some conclusion, if even in a rough way, by passing through his/her reasoning all these realities. The use of fuzzy verbal notions in every-day speech (much, more, little, small, many, a number of etc.) enables us to give a qualitative description of the problem which must be tackled and take account of its indefinite nature as well as attain the description of the factors that can't be described qualitatively.

The advent of fuzzy logic made it possible to tackle a great many problems with fuzzy input data [2]. One of them was a forecasting problem. Many of the structural elements of the latter (input data and interdependence between its components, interval evaluation of indicators and their interdependence, expert evaluations and judgments etc.) are either of a fuzzy nature or, by being in fuzzy relationships, condition the fuzzy description of the problem.

The application of fuzzy logic to the handling of forecasting problems was undertaken by the researches in which the mathematical models of fuzzy time series were described in a fuzzy form for handling the problem with fuzzy input data [3,4]. This approach was developed later by other scientists dealing with the solution of analogous problems [5,6,7]. To tackle the task, the authors proposed a model of fuzzy time series and tried to reduce the average forecasting error by making adequate alterations in the model.

The above-mentioned features of population, functioning under indefinite, uncertain circumstances, condition the fuzziness of input data or "loads" the task onto fuzzy environment. Therefore, from both theoretical and practical standpoints, handling the concerned problem based on fuzzy time series would be more expedient.

Thus, the major purpose of the proposed approach is methodological: 1) putting forth an evaluation method based on fuzzy time series for estimating model parameters; 2) testing the extent to which the model is adequate to reflect the real process, that is to say, computing the method error; 3) conducting the comparative analysis of computation results; 4) revealing the practical and theoretical importance of the model.

3 A BRIEF INFORMATION ON FUZZY TIME SERIES

Time series represents a consecutive series of observation that is conducted by equal time intervals and lies at the root of exploring real processes in economics, meteorology and natural sciences etc.

The analysis of time series of observation consists of the followings: 1) constructing the mathematical model of time series of observation of real processes; 2) model identification or selection of quantitative evaluation/ estimation method for assessing model parameters in order to test the extent to which the model is adequate to reflect the real process; 3) the conversion of identification model into time series through the statistical evaluation of model parameters.

Formally, time series can be defined as a discrete function x(t) whose argument and function values are dependent on discrete time moments as well as argument values, function values at different time intervals.

It is assumed, the time interval $0 \le t \le T$ of process x(t) is observed, that is to say, the parameter t varies along the time interval $[0,T]$ (set R) or assumes any integer belonging to this interval. For every fixed time moment t=s, the value of function, beginning from this moment, is generally determined by the values of function arguments at all the time moments ranging from t=0 to t=s-1, and value of function at all the time moments ranging from t=0 to t=s-2.

3.1 Fuzzy time series

Let us assume that $U=\{u_1, u_2, \ldots, u_n\}$ is a universal time set. The fuzzy set A of universal set U is defined as follows:

$$A=\{(\mu_A(u_1)/ u_1, \mu_A(u_2)/ u_2 \ldots\ldots, \mu_A(u_n)/ u_n)\}.$$

$$A=\{(\mu_A(u_i)/ u_i)\},\ u_i\ [0,1]$$

Where $\mu_A(u_i)$ -membership function, $\mu_A(u_i):U \Rightarrow [0,1]$; $\mu_A(u_i)$ is a degree of belonging of u_i to the set A $\mu_A(u_i) \subseteq [0,1]$, "/" is a division sign.

Let us assume that Y(t) (t=….,0,1,2…),which is a subset of of set R of real numbers, is simultaneously a universal set on which is defined a fuzzy set $\mu_i(t)$, (t=1,2,…), that is to say, the membership function is time-dependent. Let us define a set F(t) arranged out of $\{\mu_i(t), t=1,2,\ldots\}$. More precisely, F(t) is a set of fuzzy sets $F(t) =\{\mu_i(t), t=1,2,\ldots\}$. Then F(t) is a fuzzy time series defined on a universal set Y(t) (t=1,2,3….). It is evident, if F(t) is accepted as a linguistic variable, the fuzzy sets$\{\mu_i(t), t=1,2,\ldots\}$. Out of which we arranged F(t) will assume the possible corresponding values of F(t). Besides, as is evident, F(t) is time-dependent, which means, the function F(t) will assume different values at different time moments.

3.2 Fuzzy time series in demographic foresting.

The intensive changes in demographic processes that are caused by the influence of the social-demographic factors, have rendered the determination of perspective variation in total population one of the most important tasks to be tackled for demographic forecasting. To solve task of forecasting total population, we have introduced a model of fuzzy time series in this article. More precisely, the problem is described as follows: a) for a given time interval, data pertaining to the total population in Azerbaijan or to be more clear, the dynamics and respective variation of total population are available. The point is to find the anticipated total population based on the variations of the previous years.

The following principles are recommended for the solution of this problem:

1) Since the proposed method is applied to demographic forecasting for the first time, for the identification of the model or finding the extent to which it conforms to (reflects) the real process, we should, first of all, give "retrospective forecast" that comprises the followings: a) one of the previous years (t=s) is selected as a forecast year and the total population is calculated for this (t=s) year based on the variations in total population of prior years (t=s-1, s-2, …, s-k); b) the obtained results are compared with the retrospective data (real data of the s-th year) and the subsequent error is estimated. c) the experiment is carried out over a fixed time interval; d) the effectiveness of the method is evaluated based on the value of the subsequent error.
2) If there are obtained positive results, the model should be applied to calculating the anticipated total population.

4 FORECASTING METHODOLOGY.

4.1 In accordance with the description of the problem, the following forecasting methodology is proposed:

1. Definition of universal set U containing the interval between the least and greatest variations in total population.
2. Division of the universal set U into equal-length intervals containing variation values corresponding to different population growth rates.
3. The qualitative description of variation values of total population as a linguistic variable, that's to say, determining the respective values of linguistic variable or the set of fuzzy sets F(t).
4. Fuzzifying the input data or the conversion of numerical values into fuzzy values.

   This operation enables us to reflect the corresponding numerical/qualitative values of qualitative representations of population growth rates in the value of membership function.
5. Selection of parameter W>1, corresponding to the time period prior to the concerned year, calculation of fuzzy relationships matrix $P^w(T)$ and forecasting of population growth in the next year.
6. Defuzzifying the obtained results or conversion of fuzzy values into qualitative values.

On a small scale, the application of the proposed methodology to population forecasting is described in the following example:

The first step. Table 1 gives the dynamics of total population over 1980-2002 years (input data for "retrospective" forecast) and variation in total population between every next and previous year. Variation for the current year is understood to be the difference between the sizes of population in current and previous years. For example, variation for 1990 is equal to 7131900-7021200=1110700. To define a universal set U, first of all, the smallest and greatest variation values must be found over the period [1980,2001], later, to ensure the smoothness of boundaries of the interval, adequate values D_1, D_2 (positive figures are selected. After that, the universal set U can be defined as: U:U=$[V_{min}-D_1, V_{max}+D_2]$, where V_{min}=62800 is the smallest variation (year 2000), V_{max}=115900 is the greatest variation (year 1993), D_1=1800, D_2=1100. Thus, the universal set U will be as follows:

U=[61000, 117000].

Table 1. The dynamics and variation of population growth for the period [1980-2001].

Years	Total population (thousand person)	Variation (thousand person)	Fuzzification of variations
1980	6 114.3		
1981	6 206.7	92.4	A^{81}=(0.12/u_1), (0.21/u_2), (0.43/u_3), (0.90/u_4), (0.83/u_5), (0.39/u_6), (0.19/u_7)
1982	6 308.8	102.1	A^{82}=(0.07/u_1), (0.10/u_2), (0.18/u_3), (0.36/u_4), (0.77/u_5), (0.94/u_6), (0.47/u_7)
1983	6 406.3	97.5	A^{83}=(0.09/u_1), (0.14/u_2), (0.27/u_3), (0.58/u_4), (1/u_5), (0.64/u_6), (0.29/u_7)
1984	6 513.3	107.0	A^{84}=(0.05/u_1), (0.08/u_2), (0.13/u_3), (0.24/u_4), (0.5/u_5), (0.96/u_6), (0.74/u_7)
1985	6 622.4	109.1	A^{85}=(0.05/u_1), (0.7/u_2), (0.11/u_3), (0.20/u_4), (0.41/u_5), (0.86/u_6), (0.87/u_7)
1986	6 717.9	95.5	A^{86}=(0.10/u_1), (0.16/u_2), (0.32/u_3), (0.70/u_4), (1/u_5), (0.53/u_6), (0.25/u_7)
1987	6 822.7	104.8	A^{87}=(0.06/u_1), (0.09/u_2), (0.15/u_3), (0.29/u_4), (0.62/u_5), (1/u_6), (0.60/u_7)
1988	6 928.0	105.3	A^{88}=(0.06/u_1), (0.09/u_2), (0.14/u_3), (0.27/u_4), (0.59/u_5), (1/u_6), (0.63/u_7)
1989	7 021.2	93.2	A^{89}=(0.11/u_1), (0.20/u_2), (0.40/u_3), (0.85/u_4), (0.87/u_5), (0.42/u_6), (0.26/u_7)
1990	7 131.9	110.7	A^{90}=(0.05/u_1), (0.07/u_2), (0.10/u_3), (0.18/u_4), (0.35/u_5), (0.75/u_6), (0.95/u_7)
1991	7 218.5	86.6	A^{91}=(0.18/u_1), (0.35/u_2), (0.76/u_3), (0.95/u_4), (0.48/u_5), (0.23/u_6), (0.13/u_7)
1992	7 324.1	105.6	A^{92}=(0.06/u_1), (0.09/u_2), (0.14/u_3), (0.27/u_4), (0.57/u_5), (1/u_6), (0.65/u_7)
1993	7 440.0	115.9	A^{93}=(0.04/u_1), (0.05/u_2), (0.08/u_3), (0.12/u_4), (0.22/u_5), (0.46/u_6), (0.93/u_7)
1994	7 549.6	109.6	A^{94}=(0.05/u_1), (0.07/u_2), (0.11/u_3), (0.19/u_4), (0.39/u_5), (0.83/u_6), (0.90/u_7)
1995	7 643.5	93.9	A^{95}=(0.11/u_1), (0.19/u_2), (0.38/u_3), (0.81/u_4), (0.91/u_5), (0.45/u_6), (0.22/u_7)
1996	7 726.2	82.7	A^{96}=(0.24/u_1), (0.52/u_2), (0.97/u_3), (0.72/u_4), (0.33/u_5), (0.17/u_6), (0.10/u_7)
1997	7 799.8	73.6	A^{97}=(0.57/u_1), (1/u_2), (0.65/u_3), (0.30/u_4), (0.15/u_5), (0.09/u_6), (0.06/u_7)
1998	7 879.7	79.9	A^{98}=(0.31/u_1), (0.68/u_2), (0.99/u_3), (0.55/u_4), (0.25/u_5), (0.14/u_6), (0.08/u_7)
1999	7 953.4	73.7	A^{99}=(0.57/u_1), (1/u_2), (0.65/u_3), (0.30/u_4), (0.16/u_5), (0.09/u_6), (0.06/u_7)
2000	8 016.2	62.8	A^{00}=(0.95/u_1), (0.49/u_2), (0.23/u_3), (0.13/u_4), (0.08/u_5), (0.05/u_6), (0.04/u_7)
2001	8 081.0	64.8	A^{01}=(1/u_1), (0.60/u_2), (0.28/u_3), (0.15/u_4), (0.09/u_5), (0.06/u_6), (0.04/u_7)

The second step. The universal set U must be divided into several equal intervals. In our case, this set U is divided into seven eqall-length intervals: u_1=[61000,69 000], u_2=[69000,77000], u_3=[77000,85000], u_4=[85000,93000], u_5=[93000,101000], u_6=[101000,109000], u_7=[109000,117000].

If we take into account the fact that forecasting with fuzzy time series exhibits the least average error, it's necessary to find the middle points of the intervals: u_m^1=65000, u_m^2=73000, u_m^3=81000, u_m^4=89000, u_m^5=97000,

$$u_m^6=105000,\ u_m^7=113000. \qquad (1)$$

The third step. Fuzzy sets are defined on the universal set U. In this case “ the variation in total population” is a linguistic variable that assumes the following linguistic values: A_1=(very low level population growth (VLLPG)); A_2=(low level population growth (LLPG)); A_3=(changeless population growth (CPG)); A_4= (moderate population growth (MPG)); A_5=(normal-level population growth (NLPG)); A_6= (high-level population growth (HLPG)); A_7=(very high-level population growth (VHLPG)). To every linguistic value here corresponds a fuzzy variable which, according to a certain rule is assigned against a corresponding fuzzy set determining the meaning of this variable.

For example, the linguistic value “ very-low-level population growth” is given by the fuzzy variable <VLLPG, [61000, 69000], A_1>, where A_1 is a fuzzy set defined on the domain [61000, 69000] of the universal set U. See example (3).

The fuzzy set $A_1, A_2, \ldots, A_7$ is defined on the universal set U by the following formula (2):

$$\mu_{A_i}(u_i) = \frac{1}{1+\left[C\cdot\left(U-u_m^i\right)\right]^2} \qquad (2)$$

where U-variation taken from Table1, u_m^i is the middle point of the corresponding interval in (1); C is a constant. C is chosen in such a way that it ensures the conversion of definite quantitative values into fuzzy values or their belonging to the interval. (In our case C=0.0001); $A_i =(\mu_{Ai}(u_i)/u_i)$, $u_i \in U$, $\mu_{Ai}(u_i)\in[0,1]$ is a fuzzy set

If the value of the variable U in formula (2) is accepted as the middle point of the corresponding interval, the fuzzy set A_i (i=1,...,7) will be defined as follows:

$$\begin{aligned}
A_1&=\{(1/u_1),(0.61/u_2),(0.27/u_3),(0.15/u_4),(0.10/u_5),(0.06/u_6),(0.04/u_7)\}\\
A_2&=\{(0,61/u_1),(1/u_2),(0.61/u_3),(0.27/u_4),(0.15/u_5),(0.10/u_6),(0.06/u_7)\}\\
A_3&=\{(0,27/u_1),(0.61/u_2),(1/u_3),(0.61/u_4),(0.27/u_5),(0.15/u_6),(0.10/u_7)\}\\
A_4&=\{(0,15/u_1),(0.27/u_2),(0.61/u_3),(1/u_4),(0.61/u_5),(0.27/u_6),(0.15/u_7)\}\\
A_5&=\{(0,10/u_1),(0.15/u_2),(0.27/u_3),(0.61/u_4),(1/u_5),(0.61/u_6),(0.27/u_7)\}\\
A_6&=\{(0,06/u_1),(0.10/u_2),(0.15/u_3),(0.27/u_4),(0.61/u_5),(1/u_6),(0.61/u_7)\}\\
A_7&=\{(0,04/u_1),(0.06/u_2),(0.10/u_3),(0.15/u_4),(0.27/u_5),(0.61/u_6),(1/u_7)\}
\end{aligned} \qquad (3)$$

An exemplary growth of the continuous membership functions of fuzzy sets Ai depicting the values of the linguistic variable “variation in total population “ is shown in figure 1.

The fourth step. This step consists of the fuzzification of the variation calculated at the first step. This time, if Bи, Bи ∈y_j is a variation for the i-th year, then membership function for μ(yь) is calculated by means of formula (2) by holding valid the equality У=Bи, that’s to say, by separating the interval, to which belongs the considered variation, from the universal set U. The results of fuzzification for all the given years are shown in Table 1.

Here, A^{mn}-t=mn is a fuzzy set of the corresponding variation for the year t=mn where 1981<t≤2001 (for the sake of brevity, the last two digits of the years are shown in Table 1).

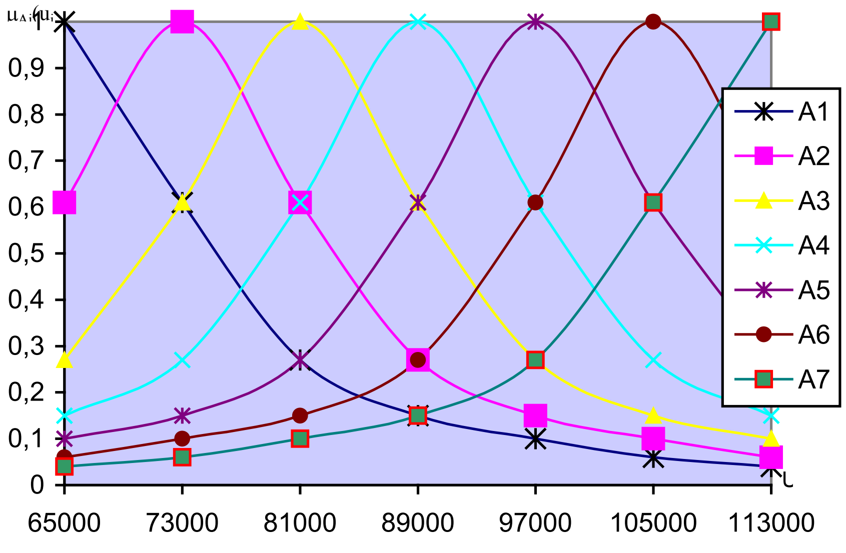

Figure 1. Membership function of values of fuzzy sets of linguistic variable ”variation in total population ”.

The fifth step. We must select a basis w (1<w<*l*, where *l* is the number of years, prior to the current year included in experimental evaluation). Resting on the basis W or the past years, we calculate a fuzzy relationship matrix $R^w(t)$ by means of which is given a forecast. For this purpose, after the selection of w, we establish an operation matrix i×j $O^w(t)$ (here i is the number of rows, which conforms to the sequence of years t-2, t-3,..., t-w, j is the number of columns conforming to the number of variation intervals) and a criteria matrix 1×j K(t) (a row matrix corresponding to fuzzy variation in total population for the year t-1). For example, by assuming that w=7, we can define the operation matrix 6×7 O^7(т) (which is the matrix of fuzzy variations in total population over the years t-2, t-3, t-4, t-5, t-6, t-7) and the criteria matrix 1×7 K(т) (which is the fuzzy variation matrix for the year t-1). Thus for w=7, the previous 8 years' data are utilized (the total population of the (t-8) year must be known to find variation of the (t-7) year).

At last, for example, in order to forecast the total population for 1990, the operation matrix O^7(т) will be established as follows:

O^7(1990)=

fuzzy variation in total population for the 1983-rd year	A^{83}
fuzzy variation in total population for the 1984-th year	A^{84}
fuzzy variation in total population for the 1985-th year	A^{85}
fuzzy variation in total population for the 1986-th year	A^{86}
fuzzy variation in total population for the 1987-th year	A^{87}
fuzzy variation in total population for the 1988-th year	A^{88}

O^7(1990)=

VLLPG	LLPG	CPG	MPG	NLPG	HLPG	VHLPG
0.09	0.14	0.27	0.58	1	0.64	0.29
0.05	0.08	0.13	0.24	0.50	0.96	0.74
0.05	0.07	0.11	0.20	0.41	0.86	0.87
0.10	0.16	0.32	0.70	1	0.53	0.25
0.06	0.09	0.15	0.29	0.62	1	0.60
0.06	0.09	0.14	0.27	0.59	1	0.63

K(1990) =[fuzzy variation in total population for the 1989-th year]-[A^{89}],

That is to say

K(1990)=

VLLPG	LLPG	CPG	MPG	NLPG	HLPG	VHLPG
0.11	0.20	0.40	0.85	0.87	0.42	0.26

According to the method, the relationship matrix R(t) is calculated at the next step

$$R(t)[i,j]= O^{w}(t)\,[i,j] \cap K(t)\,[j],$$

or

$$R(t)=O^{W}(t) \otimes K(t) = \begin{bmatrix} R_{11} & R_{12} & \dots & R_{1j} \\ R_{21} & R_{22} & \dots & R_{2j} \\ \dots & \dots & \dots & \dots \\ R_{i1} & R_{i2} & \dots & R_{ij} \end{bmatrix}$$

Where O^{w}(т) is an operation matrix; P(т) is matrix of fuzzy sets, ⊗-is an operation min(∩).

Later there is defined the forecasted value F(t) for the t year in a fuzzy form as follows.

$F(t)=[Max(R_{11},R_{21}, \dots ,R_{i1})\ Max(R_{12},R_{22}, \dots ,R_{i2}) \dots Max(R_{1j},R_{2j},\dots, R_{ij})]$

In our case 1≤ и ≤ 6, 1≤ъ≤ 7

R(1990)=

VLLPG	LLPG	CPG	MPG	NLPG	HLPG	VHLPG
0.09	0.14	0.27	0.58	0.87	0.42	0.26
0.05	0.08	0.13	0.24	0.5	0.42	0.26
0.05	0.07	0.11	0.20	0.41	0.42	0.26
0.10	0.16	0.32	0.70	0.87	0.42	0.25
0.06	0.09	0.15	0.29	0.62	0.42	0.26
0.06	0.09	0.14	0.27	0.59	0.42	0.26

Finally, the results obtained from population forecast for the year 1990 will be as follows.

F(1990)=

VLLPG	LLPG	CPG	MPG	NLPG	HLPG	VHLPG
0.10	0.16	0.32	0.70	0.87	0.42	0.26

Forecasting results for other years are calculated in an analogous manner.

The sixth step. To fuzzify the obtained results of the 5-th step, the following formula is proposed,

$$V(t) = \frac{\sum_{i=1}^{7} \mu_t(u_i) \cdot u_m^i}{\sum_{i=1}^{7} \mu_t(u_i)}$$

Where $\mu_т(у_и)$ is the calculated value of membership function for the forecast year t, u^i_m are the middle points of intervals.

For example, after calculating F(1990) =93300, that is to say, anticipated population growth for year 1990 equals to 93300 persons. In orders to estimate the forecasted total population for year 1990, we must add the calculated population grown to the total population for the year 1989. In other words

N(1990)=7 021 200+93 300=7 114 500

5 RESULTS OF EXPERIMENTAL CALCULATIONS

For evaluating the effectiveness of the proposed methods' application to handling demographic forecasting problems, the total population has been calculated over a certain time period.

Experiments were conducted at two retrospective and perspective stages.

Table 2.

Results obtained from the retrospective analysis of population forecasting

Years	Actual		Forecasted		Error (%)	Average error (%)
	Total population (thousand person)	Variation (thousand person)	Total population (thousand person)	Variation (thousand person)		
1988	6 928.0	105.3	6 926.7	104.0	0.02	
1989	7 021.2	93.2	7 028.0	100.0	0.10	
1990	7 131.9	110.7	7 114.5	93.3	0.25	
1991	7 218.5	86.6	7 234.9	103.0	0.23	
1992	7 324.1	105.6	7 308.5	90.0	0.22	
1993	7 440.0	115.9	7 425.1	101.0	0.20	
1994	7 549.6	109.6	7 544.3	104.3	0.07	0.13
1995	7 643.5	93.9	7 647.9	98.3	0.06	
1996	7 726.2	82.7	7 736.5	93.0	0.13	
1997	7 799.8	73.6	7 812.0	85.8	0.03	
1998	7 879.7	79.9	7 884.0	84.2	0.05	
1999	7 953.4	73.7	7 962.6	82.3	0.11	
2000	8 016.2	62.8	8 034.4	81.0	0.23	
2001	8 081.0	64.8	8 093.4	77.2	0.15	

1. At the retrospective stage, the time interval [1988,2001] was selected as an experimental base. It is clear, this period's statistical data relating to total population is available.

The essence of the conducted experiment consists of the followings; a) the dynamics of total population for the examined period is considered to be unknown; b) with the aid of the proposed methodology, total population size was forecasted for every year selected from the time interval [1988,2001] based on the changes in population growth rates of the previous years. c) in order to test the model's degree of adequacy, the observed (real) dynamics of total population over the time interval [1988,2001] was compared with the corresponding results of model application and the consequent model error computed .

The error of the proposed method is computed by the following formula

$$\delta(t) = \frac{V^t_{obsev.} - V^t_{forec.}}{N^t_{obsev.}} \cdot 100\%$$

Where $V^t_{obsev.}$ is the variation in total population for the t-th year; $V^t_{forec.}$ is the variation in total population for the t-th year; $N^t_{obsev.}$ is the observed total population for the t-th year, $1988 \le t \le 2001$.

Table 2 displays the observed total population sizes and variations, results of forecasting calculations, the respective error values and the mean (average) error value. For every year ($t=s_i$) selected from the time interval[1988,2001], the forecast calculation is made by taking into account the previous 8 years' total population sizes (s_i-s_j=8, $1988 \le s_i \le 2001$, $1980 \le s_i \le 1993$).

The comparative analysis of the observed and forecasted data and the consequent error of the approximated method have confirmed the high efficacy of the model and made us believe that its application to demographic forecasting would serve our purpose.

Figure 2, depicting graphically the dynamics of the observed (actual) and forecasted total population sizes, displays the two data's remarkable closeness that, in its turn, necessitate continuing the researches conducted in this direction.

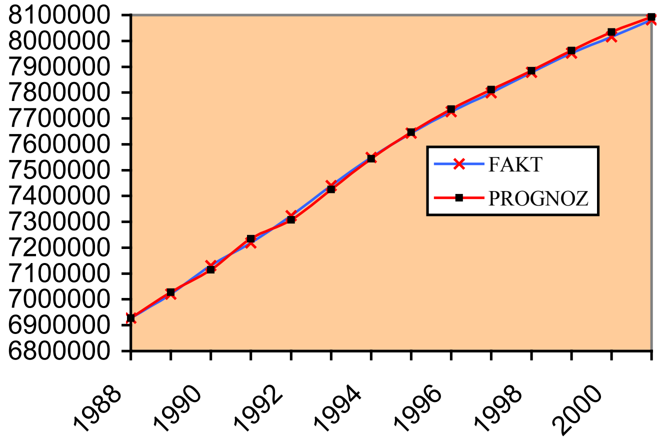

Figure 2: Results of retrospective analysis of total population forecasting.

Table 3. Forecasted total population over the 2002-2012 period

Years	Forecasted total population (thousand person)	Variation (thousand person)	Fuzzification of variations
2002	8 155.0	74.7	A^{02}=(0.52/u_1),(0.97/u_2),(0.71/u_3),(0.33/u_4),(0.17/u_5),(0.10/u_6), (0.06/u_7)
2003	8 234.5	78.8	A^{03}=(0.34/u_1),(0.75/u_2),(0.95/u_3),(0.49/u_4),(0.29/u_5),(0.13/u_6), (0.08/u_7)
2004	8 316.0	81.5	A^{04}=(0.27/u_1), (0.58/u_2),(1/u_3),(0.64/u_4),(0.30/u_5),(0.15/u_6),(0.09/u_7)
2005	8 399.0	83.0	A^{05}=(0.24/u_1),(0.50/u_2),(0.96/u_3),(0.73/u_4),(0.37/u_5),(0.16/u_6), (0.10/u_7)
2006	8 483.0	84.0	A^{06}=(0.21/u_1),(0.45/u_2),(0.92/u_3),(0.80/u_4),(0.37/u_5),(0.16/u_6), (0.11/u_7)
2007	8 568.0	85.0	A^{07}=(0.20/u_1),(0.41/u_2),(0.86/u_3),(0.86/u_4),(0.41/u_5),(0.20/u_6), (0.11/u_7)
2008	8 654.0	86.0	A^{08}=(0.18/u_1),(0.37/u_2),(0.80/u_3),(0.92/u_4),(0.45/u_5),(0.22/u_6), (0.12/u_7)
2009	8 740.4	86.4	A^{09}=(0.18/u_1),(0.36/u_2),(0.77/u_3),(0.93/u_4),(0.47/u_5),(0.23/u_6), (0.12/u_7)
2010	8 826.3	86.9	A^{10}=(0.18/u_1),(0.34/u_2),(0.74/u_3),(0.95/u_4),(0.50/u_5),(0.23/u_6), (0.13/u_7)
2011	8 914.9	87.6	A^{11}=(0.16/u_1),(0.32/u_2),(0.70/u_3),(0.98/u_4),(0.53/u_5),(0.25/u_6), (0.13/u_7)
2012	9 002.6	87.7	A^{12}=(0.16/u_1),(0.32/u_2),(0.74/u_3),(0.98/u_4),(0.54/u_5),(0.25/u_6), (0.14/u_7)

6 CONCLUSION

The methodology proposed in this article enables us to forecast demographic processes on the basis of fuzzy time series. A peculiar trait of the methodology consists of its capability to forecast the required indicator by utilizing incomplete, fuzzy input data. The described approach, by entering the dynamics of total population until some previous year into an experimental base, helps make forecast calculations for any distant perspective. This, in its turn, allows us to take into account the trend of previous population growth rates and as a result achieve more accurate forecasts.

It should be remarked that the described experiments cover a limited time period. Besides, in order to forecast the concerned indicator for any given year, only the statistics of the previous 8 years was utilized due to the difficulty of all these the calculations being made manually.

The reason, why these mathematical operations were made in a manual way, is associated with the lack of analogous applications of fuzzy time series to demographic forecasting on account of unavailability of appropriate software resources. Therefore, the obtained results and conclusions in this article are of a preliminary nature. The algorithm of the proposed methodology and the development of adequate software resources are intended in the future. The latter would facilitate data processing and enable us to form a final opinion about the expediency of the described method in handling demographic forecasting problems.

As is evident, although exploring the dynamics of total population provides us with its primary, aggregate characteristics, it does not mirror its reproduction process or the structure of population. Therefore, in the future, the range of forecasted population characteristics is intended to be extended by including other population indicators such as age structure, new-borns, the dead, migration etc.

7 REFERENCES

Заде Л.А. Понятие лингвистической переменной и его применение к принятию приближенных решений. М.: Мир, 1976.

Abbasov A.M., Mamedova M.H., Gasimov V.A. Fuzzy Relational Model for Knowledge Processing and Decision Making. –Advances in Mathematics, New-York, 2002, vol.1, pp.191-223.

Q.Song, B.S.Chissom, Fuzzy time series and its models, Fuzzy Sets and Systems 54 (1993).

Q.Song, B.S.Chissom, Forecasting enrollments with fuzzy time series –part II, Fuzzy Sets and Systems 62 (1994).

S.M.Chen, Forecasting enrollments based on fuzzy time series. Fuzzy Sets and Systems 81 (1996).

Proc.7th Internat. Conf. On Information Management, Chungli, Tayouan,Taiwan,ROC, 1996.

М.З.Ахмедов. Новый вариант решения проблемы прогнозирования с помощью нечетких временных рядов. Известия НАН Азербайджана, сер. физ.-техн. и мат. наук, №3, 2001.