An Adaptive Statistical Data Processing Algorithm Applied to SD Modeling of Iran's Demographic Transition

Hamed Shakouri G.[1], Shakib Taheri[2], Ayyub Ansarinejad[3], M. Sadegh Shahmohammadi[4]
Department of Industrial Engineering, Faculty of Engineering, University of Tehran, Iran
[1]hshakouri@ut.ac.ir
[2]shakib_taheri@yahoo.com
[3]ansarinejad@gmail.com
[4]sadegh_shahmohammadi@yahoo.com

Abstract:

There are different official estimations about current and future growth rate of Iran's population. Inadequacy and unreliability of data in addition to usage of unsuitable forecasting methods are the main reasons for existence of this variety. To have accurate estimates for year on year growth rate, in this research, a population system dynamics model is implemented. To run the model, total fertility rate and other needful fertility parameters are calculated by processing raw data. In the next step and to resolve the statistical inconsistencies in census data which have been revealed by calculation of survival fractions and death rates, an appropriate adaptive process is proposed and applied to modify the parameters. The result of applying model shows that the next ten-year average growth rate will be about 1.9. Finally, simulation results of three possible scenarios on the fertility factor are obtained that warns on exceeding of population over 100m by 2020.

Keywords: System Dynamics Modeling, Population Dynamics, Uncertain Data Processing, Census Data Correction, Parameter Adaption.

1-Introduction

"Perhaps no single factor is more important for local government planning than the size and composition of a region's population and the way it will change in the future. Changes in Population composition can fundamentally alter the need for public facilities and services." (Klosterman 1990). Demographic forecasting has a long history (De Gans 1999) and has important effects. Population, household's number and related forecasts form the basis of social and economic planning and are fundamental to many other forecasting exercises (Booth 2006). To provide necessary detail in demographic forecasting, the population must be disaggregated by age and sex. To achieve this, the three components of population change (mortality, fertility and migration) must be separately forecasted and appropriately combined (Booth 2006). Population forecasting is thus a highly complex and difficult undertaking (Keyfitz 1985). During the twentieth century, fertility was the most important component in determining population size. The historical booms and busts in childbearing have to a large degree provided the urges for the more recent renewal of interest in demographic forecasting. The uneven and transitional age structure created by past fertility instabilities is the dominant feature of present day demography in the developed world (Booth 2006).

Several useful research and studies have previously done on demographic forecasting (Lutz, Goldstein, and Prinz 1996; Ahlburg 1982). Walonick remarked that many scholars have proposed a variety of ways to categorize forecasting methodologies. And the following classification is a modification of the schema developed by Gordon over two decades ago to forecasting: Genius forecasting, Trend extrapolation, Consensus methods, Simulation methods, Cross-impact, Matrix method, Scenario and Decision trees (Walonick 2006). Booth categorized approaches to forecasting demographic processes to three parts: extrapolation, expectation (individual-level birth expectations or population-level opinions of experts), and theory-based structural modeling involving exogenous variables (Booth 2006). Also there are several methods to forecast population growth. The following methods were reviewed to find out which is most applicable to forecasting the Iran population for the short and long term. These methodologies are:

• Cohort Component Model: Need for more comprehensive and more in-depth projections in population Demographics detail have led to the development of the cohort-component projection technique. Cohort Component Model is a system of demographic accounting in which the population is advanced forward in time through the application of time specific survivorship ratios by age and sex and the derivation of births from time-specific fertility rates of women by age; migration by age and sex can also be incorporated (Preston, Heuveline, and Guillot 2001). This method can be quite accurate when forecasting population up to ten or twenty years into the future where in-migration is not the major growth factor (Van Buskirk and Associates 2004).

• Simple Curve fitting or Extrapolation Model: Simple Curve fitting or Extrapolation Model assumption is that the future will be a continuation of the past. This method is the most common approach in demographic forecasting and widely used (Schmitt 1954; Smith, Tayman, and Swanson 2001). The most commonly-used method of extrapolation is univariate ARIMA modeling (Box, Jenkins, and Reinsel 1976). Simply extended linear curve into the future is not an exact estimate of future growth. For the reason that communities in Iran is in their mid-stage of growth the extrapolation of the growth curve into the future greatly overestimate future growth. Therefore, this model is not the most appropriate one for Iran.

• Exponential Model: An exponential trend is one where the trend is increasing at a constant rate of change each year. This compounding effect of a constant rate of growth can result in astronomical increases in forecasted population in the long term. While this type of trend in growth may exist for a period of five years or even ten years, it cannot sustain itself for longer terms (Van Buskirk and Associates 2004). This model would be misleading for forecasting long-term growth for Iran.

• Gompertz (Sigmoid or Logistic) Model: Many biological populations (including cities) tend to grow at a rate over time that simulates a logistic or Sigmoid Curve. Population growth increases at an increasing rate over time until it reaches an inflection point, then the increase in population growth is at a decreasing rate until it reaches upper growth limit. One of the key variables in this growth equation is its upper growth limit (build-out).As an example, the upper limit for large-scale pre-platted communities such as bounded city can be precisely defined by calculating the total number of housing units that can be built on platted lots, and un-platted lands. The housing units can be translated to population. The Sigmoid Model is a more scientifically sophisticated variation of an extrapolation model and should be more accurate than other methods for forecasting short and long term growth for small zone given the larger role of in-migration and estimated "build-out" population (Van Buskirk and Associates 2004; Capece). But for a country with enough amounts of resource and vast land like Iran, this model is not fit with situations.

An efficient tool to study the population transition is System dynamics (SD). System dynamics is an approach to understanding the behavior of sophisticated systems over time and firstly introduced by Forrester (Forrester and Wright 1961). The importance of all research in population statistics can be understood by considering concepts of Planning, Employment, Poverty and Wealth, Health factors and Education. There is no doubt that having precise population statistics is very helpful in the fields of better economical planning and social development. The population model is developed using an object-oriented modeling software package called Vensim (Ventana Systems, Inc., 2002). This software is designed for use in collaborative model building projects and has a graphical interface that allows model participants to concentrate their attention on symbolic objects, rather than numbers and equations. The model objects symbolize system elements such as stocks, processes, material and information flows. The software employs coupled systems of finite difference equations as a modeling framework (Cole and Flenley 2008).

In this study a theory-based structural cohort component model, involving both endogenous and exogenous variables (SD model) is implemented to observe Iran's demographic transition. Then a novel combined adaptive process is proposed and applied to modify unreliable data using the SD model. If a reliable and precise set of raw data exists (a complete census statistics plus death and fertility rate of each age group, etc.) then employing a system dynamics model would have lots of profits like: accurate and trustworthy results, ability to study population transitions in different cohorts, evaluate effects of exogenous variables, population policies, and so on. Unfortunately, existing data, which is available on the Statistical Center of Iran (sci.org.ir) is not adequate and/or accurate.

To solve the problem and to obtain reliable data an adaptive statistical data processing algorithm is designed. The algorithm applies the implemented SD model to obtain reliable parameters essential to run the model and achieve valid simulation results.

The paper is organized as follows. Section 2 briefly describes the system dynamics model structure and different variables used in population projection. Section 3 explains mathematical relations used in the SD model. Section 4 discusses on statistical data and explains the lack of accuracy in the official data. Then by processing data we have estimated the parameters related to fertility. Section 5 proposes an adaptive statistical data processing algorithm using the SD Model for data correction. The designed adaptive algorithm is a novel method by which the problem of data inconsistency is resolved. Model validation and some simulation results that can be obtained by the model are represented in Section 6, including population growth rate forecasts for three different scenarios on fertility factor. Section 7 concludes the paper.

2-The Model

An aging chains model of population dynamics is developed according to the well known basic definitions such as state (stock) variables, rate (flow) variables, auxiliary variables, and cohort component (Sterman 2000). The population in Iran is divided to males and females in 5-year age groups, having results of the 10-year census and the population pyramid. For the range of new babies to more than 100 years old people, we have 21 state variables in the aging chain. There are many structural and practical reasons for using cohorts; for example, mortality rate is high in early childhood birth, which reduces for children and young people, and then increases for older, finally reaching one in the aged people (Figure 1).

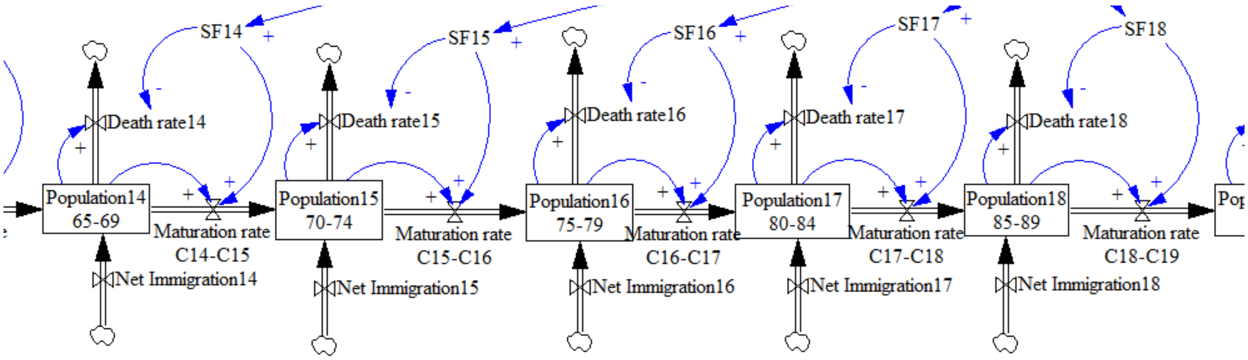

Figure 1: A part of the SD model built by Vensim the Cohorts

Women between the ages of 15 to 50 years have ability to give birth. Birth rate in these years doesn't have a uniform distribution and is affected by various biological, social and economic factors (Figure 2).

These cohorts are required and useful for studying the diversity of population behavior according to different cohorts and also to forecast and then plan for meeting special needs of the different groups.

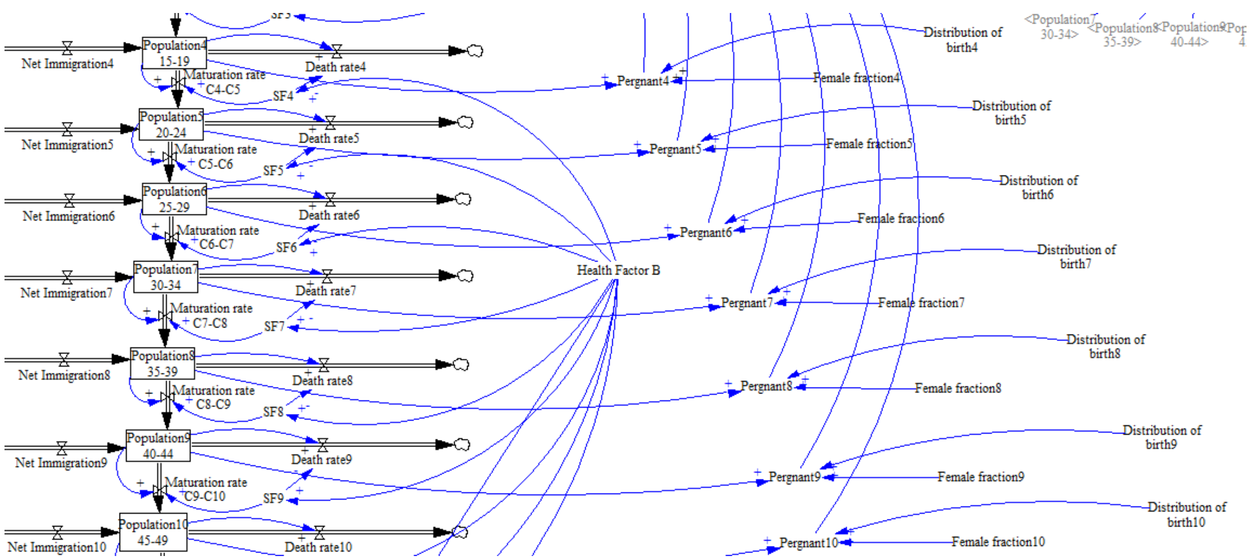

Figure 2: A part of the SD model; distribution of women in childbearing cohort and the Health Factor as an exogenous variable.

Considering phenomenon such as increasing population growth ratio in 1980's (so called baby boom) and consequently effects of this event, on need to educational facilities, employment, social welfare and retirement, with a delay emphasize importance of the cohort model. Moreover, impact of birth-giving age groups on the dynamics of the system by repeating this event and rebuilding next generation, which leads to future growth of the population, is the main object of building such a model.

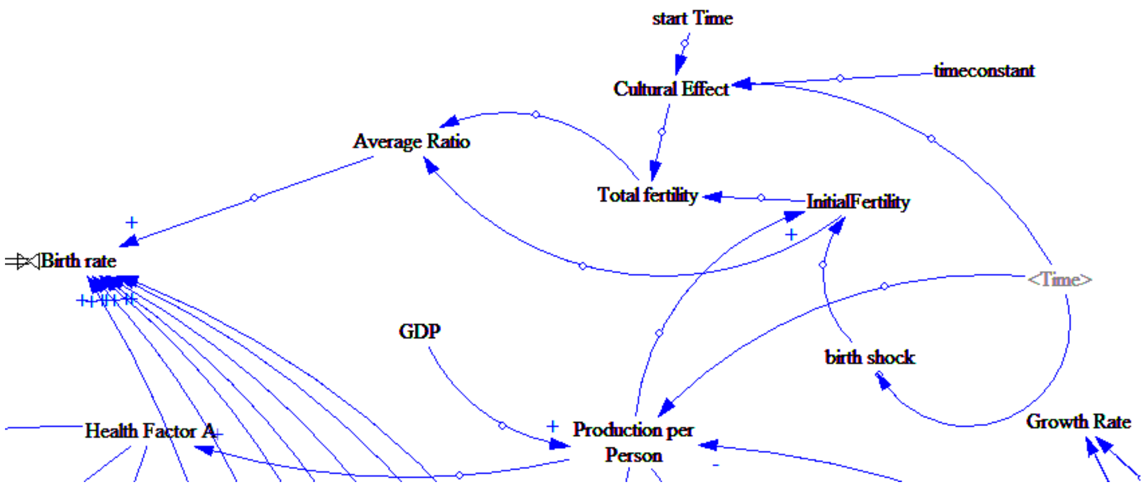

Figure 3: A part of the SD model; birth rate, growth rate exogenous variables like: policy, GDP, culture and health.

As mentioned, in this article the proposed model is used to forecast population growth rate; however, for further researches this model can be utilized in several fields such as labor market, housing market, education, social welfare and health. In addition, the interaction of each mentioned field with socio-economic factors like GDP, oil price, recession, inflation, per capita income and other political and cultural factors and/or health and population control policies can be separately considered. This section will not cover the benefits and reason of using this specific model or the possibility of construction of different casual diagram in different perspectives. In this model and for long or short perspectives, the abovementioned factors are considered as exogenous variables (Figure 3). We can run different scenarios within the current system boundary by changing some of these variables.

3- Mathematical relations

According to the pyramid of age groups known in demographics, states variables are used to demonstrate different age-sex groups and these variables increase or reduce with their related flow variables. For the first cohort, newborn babies up to 4 years old (5 years minus 1 day) are considered as infant cohort.

$$P_{s,1}(t) = P_{s,1}(t_0) + \int_{t_0}^{t} (B_s(t) + I_{s,1}(t) - D_{s,1}(t) - M_{s,1}(t))dt$$

Here, $P_{s,1}(t)$ is the population in the first cohort, $B_s(t)$ is the birth rate, $I_{s,1}(t)$ is net immigration to this group that migration out will be deducted from migration to country. $M_{s,1}(t)$ is the rate of growth which shows the number of people going from the first Cohort to the second Cohort. Birth rate is the sums of children are born during the reproductive years of women.

$$P_{s,i}(t) = P_{s,i}(t_0) + \int_{t_0}^{t} (M_{s,i-1}(t) + I_{s,i}(t) - D_{s,i}(t) - M_{s,i}(t))dt \qquad (for\ i \in \{2, \dots, 20\})$$

$$P_{s,21}(t) = P_{s,21}(t_0) + \int_{t_0}^{t} (M_{s,20}(t) + I_{s,21}(t) - D_{s,21}(t))dt$$

$P_{s,i}(t_0)$Is initial population in moment of t_0 and cohort *i* and $P_{s,i}(t)$ is population in cohort *i* and $M_{s,i-1}(t)$ is the rate of population entered from previous cohort. also $M_{s,i}(t)$ is the rate of growing up and going from cohort *i* to cohort *i+1* and $D_{s,i}(t)$ is mortality rate in the cohort of *i* and the amount of it is calculated using the survival factor of cohorts, $I_{s,i}(t)$ is the net migration to the current group which described above. Other variables except $I_{s,i}(t)$ are endogenous variables. Index s, in all variables is used to separation of gender to male and female groups.

$$B_{\mathrm{s}} = 7\mathrm{S}_{\mathrm{s}}(\frac{\mathrm{TF}}{(\mathrm{CY}_{\mathrm{f}} - \mathrm{CY}_{\mathrm{i}} + 1)}) \sum_{a=\mathrm{CY}_{\mathrm{i}}}^{\mathrm{CY}_{\mathrm{f}}} w(a) P_{\mathrm{fem}}(a)$$

$$\sum_{a=\mathrm{CY}_{\mathrm{i}}}^{\mathrm{CY}_{\mathrm{f}}} w(a) = 1$$

In this formulation $P_{fem}(a)$ is the population of women in cohort of a, and *TF* (total fertility) is the total number of children per women during childbearing years. First and last years of fertility is calculated and multiplier 7 shows number of cohorts considered for the time of fertility. Consequently $7(\frac{TF}{(CY_f - CY_i + 1)})$ is the average baby's birth per each woman, during the childbearing years. $w(a)$ is a coefficient of weight that dependent to each cohort, actually shows births than occur in any of the childbearing years. Weight coefficients depend on two factors, including biological factors such as maturity and nutrition and social economic factors such as the role of women in the community, age of marriage and education. Coefficient of gender is the ratio of each gender in all births and is generally close to 0.5. This ratio in all societies and over time is not constant and the preferences of people and using technology could be change this ratio a little.

$$Exit\ Rate(i)_{\mathrm{s}} = DELAYI\big(\mathrm{M}_{\mathrm{s}}(\mathrm{i}-1) + \mathrm{I}_{\mathrm{s}}(\mathrm{i}), YPC(i)\big)$$

$$\cong {P_{\mathrm{s}}(\mathrm{i})}\big/{YPC(i)}$$

Auxiliary exit rate variable represents that each population which entered a state variable finally will be out with a maximum first degree of delay which is composed of two parts, first transfer to the next cohort and the second one, is death.

$$\mathrm{M}_{\mathrm{s}}(\mathrm{i}) = Exit\ Rate(i)_{\mathrm{s}} * SF_s(i)$$

$$\mathrm{D}_{\mathrm{s}}(\mathrm{i}) = Exit\ Rate(i)_{\mathrm{s}} * (1 - SF_s(i))$$

$SF_s(i)$ or survival fraction shows the coefficient of population which transfer to the next cohort. Survival fraction can be calculated through different method such as life table, survival distributions for the population. (Keyfitz and Caswell 2005; Rosner 2006; Lee 1992) Describe the mathematics of life tables and survival analysis in discrete and continuous time.

By using survival factor, probability of death per year, or annual death rate of each cohort, is calculated through equation below:

$$FDR_{\mathrm{s}}(\mathrm{i}) = -\ln(SF_s(i) / YPC(i))$$
$$SF_s(i) = \exp(-FDR_{\mathrm{s}}(\mathrm{i}) * YPC(i))$$

These equations are obtained due to geometric growth and multiplying survival factor for each year in the remaining population of previous year. $YPC(i)$ is the range of ages in each cohort i and in our model the range of ages in each cohort except the ages above 100 years, is 5 years. Till here the structure of our model is described. In long-term perspective other auxiliary and exogenous variables in model are considered. Such as cultural effects (Auxiliary variable) which continuously and over time reduces the fertility rate and health factors that considered in three categories (Health Factor A, B, C). These three regulatory variables utilized as a result of policy making in the model and potentially can affect survival fractions and death rate in different cohorts.

4-Discussion on Statistic Data

Iran's dynamic population model after reviewing relationships among variables was constructed. In the next step by entering data and parameters into the model and utilizing Vensim, simulation process was done. It is important to mention that unfortunately in Iran statistics data and population census, even 10 years interval population and housing census data, is not consistence and reliable. One of the clear mistakes that happen in general census is reporting the number of infants and babies lower than actual numbers. The actual number of infants and babies will report in the next ten years general census. The elderly and single households also will be counted lower than reality. Because of various reasons, including the lack of people cooperation and lack of efficient and accurate statistics system, also statistics about dead people, is not accurate.

Next section will show that available $SF(i)'$s which are derived from census data is not precise and reliable consequently an adaptive statistical data processing algorithm combined with SD model, will apply to modify $SF(i)'$s parameters.

Also fertility parameters are not clearly available. To achieve reliable statistics about fertility data processing methods used. $w(a)'s$ or weighted coefficient of each woman childbearing in specific cohort, which is described above, calculated according to statistics published by the United Nations population division and the impact of two biologic and social-economical factors. First column (in the left) of each cohort in Figure 4 is derived from United Nation Population Division and shows the distribution of birth by mothers' age.

Second, third and forth columns of each cohort in Fig. 1 show the result of $w(a)$ calculated by using data obtained from United Nation 2008 and Abbasi (Abbasi-Shavazi and McDonald 2005). These processed results are shown in table 1.

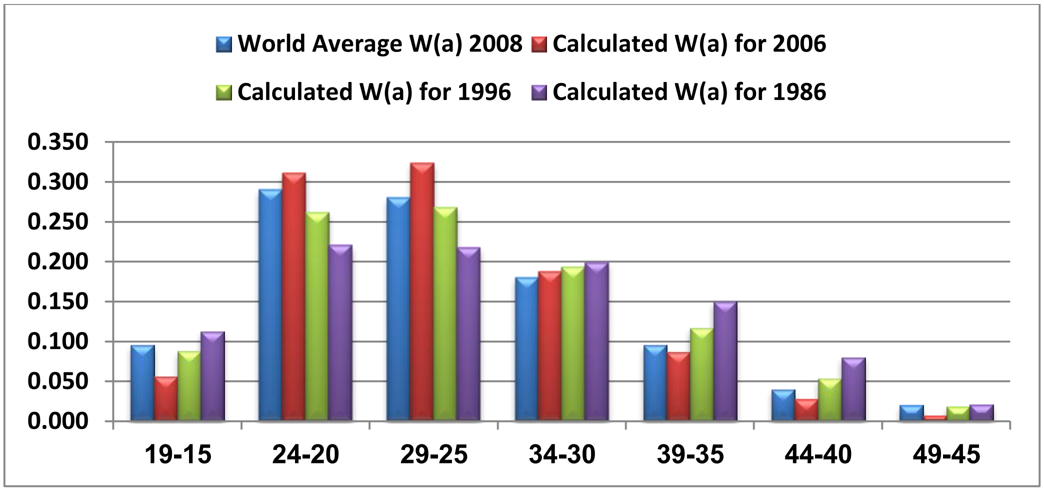

Figure 4: Iran Data Processing Result for $w(a)$.

Table 1: The processed results for Fertility Ratio and $w(a)$

Cohort	Women Population (thousand Person)			Fertility Ratio (per woman)			*W(a)*			
	2006	1996	1986	2006	1996	1986	World Data	2006	1996	1986
15-19	4283.9	3535.7	2531.8	0.02339	0.05554	0.13853	0.095	0.056	0.088	0.113
20-24	4499.6	2655.5	2090.1	0.13045	0.16542	0.27146	0.29	0.311	0.262	0.221
25-29	3564.8	2343.3	1812.7	0.13575	0.16928	0.26766	0.28	0.324	0.268	0.218
30-34	2715.6	1967.3	1446.5	0.07862	0.12227	0.24443	0.18	0.187	0.193	0.199
35-39	2409.6	1754.2	1073.4	0.03637	0.07389	0.18415	0.095	0.087	0.117	0.15
40-44	2007.5	1381	821.6	0.01172	0.03378	0.09835	0.04	0.028	0.053	0.08
45-49	1730.3	1022.9	766.2	0.0031	0.01182	0.02539	0.02	0.007	0.019	0.021
Total Fertility				2.1	3.2	6.1				
Total Population	70495.8	60055.5	494450							

As previously mentioned to calculate the death rate per each cohort *(D)*, survival factor (*SF)* is used. Review of Population and Housing Statistics Center of Iran Census; reveal that there are problems in process and results of census and these problems should be modified.

5-The proposed adaptive algorithm for data correction

These problems in census are described by an example. Let's consider first 0 to 4 years cohort, in the census of 1996. This cohort should be transferred to 10 to 14 years cohort in census of 2006. If we assume that all people in the first cohort during these 10 years, transferred to 10 to 14 years cohort, or in other word if we assume that during these years population didn't change (increase or reduce), then population ratio of first census to second on in 2006, must be exactly equal to 1. Obviously this assumption is not true because in reality during these 10 years numbers of people have died also due to migration numbers of population have changed (increase or reduce). Especially when we notice that mortality rate in early birth is significant. So the survival ratio, without considering of net migration rate, must be between 0 and 1.

The other important point is that the population in the third cohort in 2006, consists of people remained alive during 10 years ago or the first cohorts in 1996 plus immigrants entered the country, minus migrants from this cohort. According to the statistics and information (Iran Yearbook of Statistics in 2007), if we assume that all immigrants during these 10 years add up top third cohort in 2006, the number of immigrants will be equal to 66,421. But according to census of 2006, third cohort in this year is 545,570 people more than first cohort in 1996 census data. This result means that survival fraction is greater than 1 and this is clearly not correct and rational. Probably this result is a consequent of low counting people in 1996 census or high counting people in census of 2006.

Table 2: Inconsistent SF_i Related to Cohorts

i	Cohort	SF2006/1996	SF96/91	SF96/86	SF91/86
1	0-4	1.088523	1.041831	1.003964	0.998965
2	5_9	1.028875	1.005005	0.945475	1.002822
3	10_14	0.992373	0.942815	0.884587	1.000949
4	15-19	1.015375	0.883748	0.906967	0.952825
5	20-24	1.063491	0.951871	0.949053	0.955065
6	25-29	1.045012	0.993705	0.977954	0.959456
7	30-34	1.02741	1.019279	0.960417	0.979059
8	35-39	0.986276	0.980959	0.950798	0.96234
9	40-44	0.979849	0.988006	0.923718	0.953262
10	45-49	0.937876	0.969008	0.862072	0.99068
11	50-54	0.957735	0.870183	0.864872	0.902384
12	55-59	0.876217	0.95843	0.804617	0.974318
13	60-64	0.809372	0.825826	0.714575	0.747132
14	65-69	0.644871	0.956424	0.634577	0.810684
15	70-74	0.525759	0.782768	0.42825	0.543807
16	75-79	0.350601	0.787503	0.276003	0.667117
17	80-84	0.268847	0.456868		0.569051
18	85-89	0.207178			

It is notable that this conclusion was regardless of migration abroad, according this fact, problems in survival fraction *(SF)* Statistics from census will appear clearly. To obtain net migration rates per cohort, accurate statistics is needed, but according to table of page 114 in Iran Yearbook of Statistics 2007, net migration rate of the population is about zero. In Table 2 shown SF from 1986 census up to 2006 census, because of SF> 1, and similar problems such as irrational trends of data and behavioral inconsistency in SF in compare to global statistics of SF, non-acceptable and unreliable data is specified.

To modify the model parameters, a system dynamics model is used and in an adaptive decision-making process the following three criteria is considered:

I. SFi's must be logically and behaviourally consistence with the global statistics

II. Transfer a cohort in a census to the corresponding cohort in another census

III. Total numbers of deaths

Our adaptive algorithm first reads inputs. Then checks the satisfaction of CriterionI, which explain that SF_i's must be logically and behaviorally consistence with the global statistics. If this condition satisfied then goes to next level else modifies SF_i's by using world statistics then goes up. Next step is Inputting Data and Parameters like SF_i's to SD Model of population and Simulating. Then calculate the error that explains below:

$$\text{Error } SF_i = P_i(Simulate) - P_i(Statistical\ Data)$$

$$RE_i = {}^{\text{Error } SF_i}/_{P_i(Statistical\ Data)}$$

Then checks the second criterion by result of SD model, if absolute value of RE_i was greater than $\boldsymbol{\alpha}$ then substitutes $SF_i(new) = SF_i(old) -$ Relative error of SF_i, and goes to first. Else in next step calculates the relative error of total death (RETD) from the data and the simulation results and checks if *RETD* was greater than $\boldsymbol{\beta}$ then checks error sign, if it was positive only selects SF_i's which are related to positive errors and otherwise only selects SF_i's which are related to negative errors. Then utilizes selected SF_i's and for below calculations and goes to first step:

$$w_{SF_i} = P_i(Simulate) - P_i(Statistical\ Data)$$

$$SF_i(new) = SF_i(old) - \frac{w_{SF_i}}{\sum_{\text{Selected}} w_{SF_i}} \times (\text{RETD})$$

The above process continues from first step until all of SF_i's satisfy all criteria simultaneously. Finally the modified parameters and corrected data succeeded.

Following Flowchart shows the adjustment of parameters in the adaptive process (Figure 5):

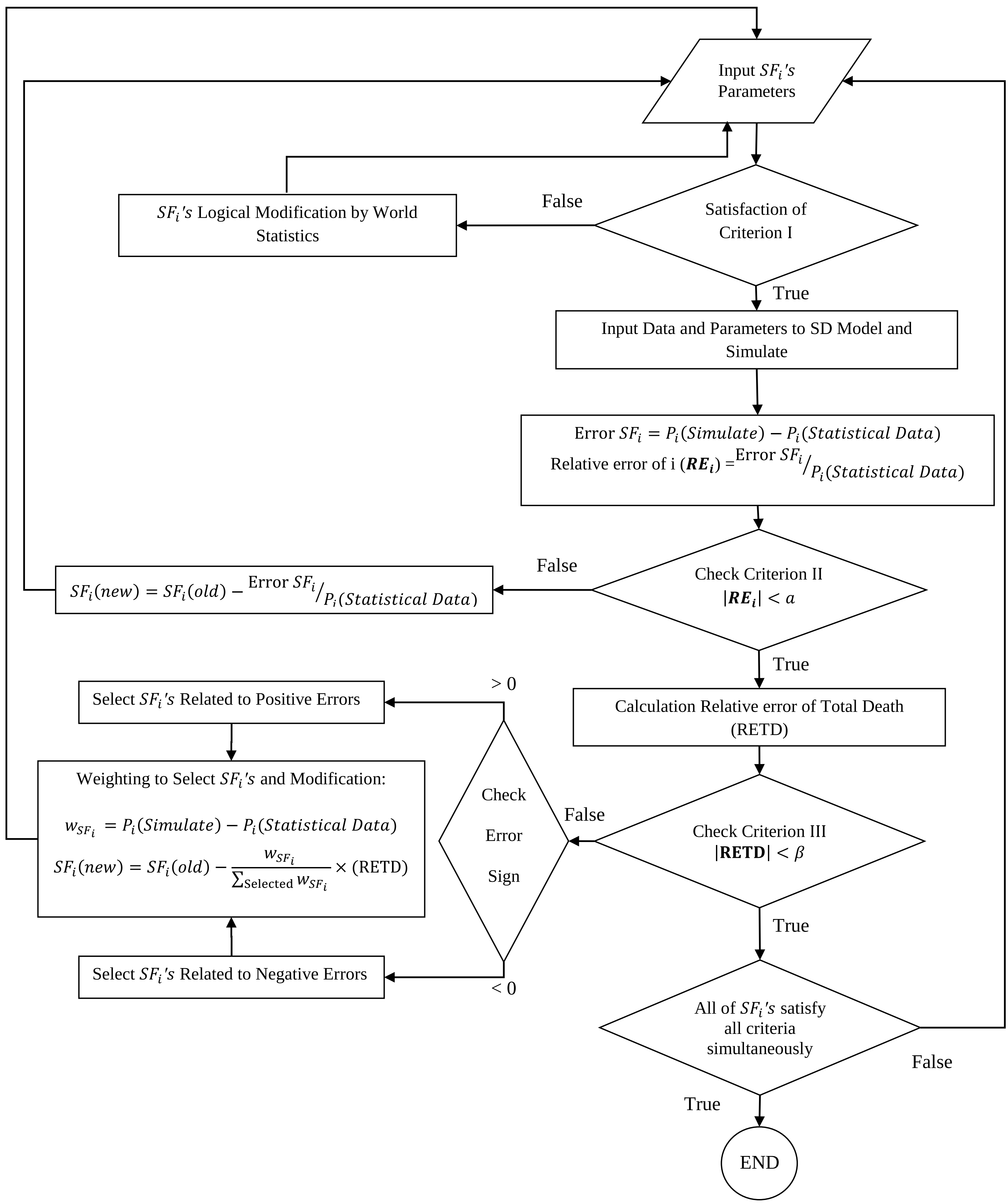

Figure 5: Adaptive Statistical Data Processing Algorithm Using SD Model

Table 3: Adjusted Parameters of SF_i's Resulted from Adaptive Statistical Data Processing Algorithm

Cohort	0-4	5_9	10_14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54
SF_i	0.9402	0.982642	0.998371	0.999648	0.999819	0.999649	0.99839	0.996808	0.99507	0.985694	0.966559

Cohort	55-59	60-64	65-69	70-74	75-79	80-84	85-89	90-94	95-99	100-
SF_i	0.901765	0.867682	0.852311	0.79965	0.670123	0.603504	0.568197	0.5605	0.550449	0.549309

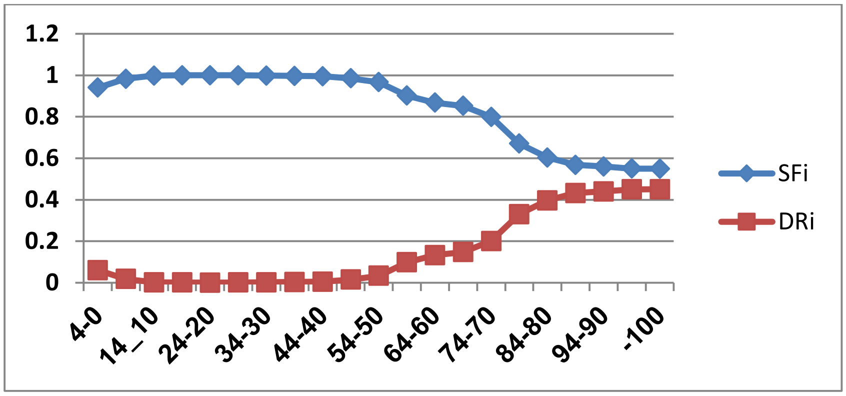

Figure 6: SF_i and DR_i for Each Cohort Resulted from Adaptive Process

The above shows the results of proposed adoptive algorithm.

6-Model Validation and Simulation Results:

In order to check reliability of model and evaluate the accuracy of outputs assume that we are in 1996 and by using data of that time simulate the model and observe population behavior up to 2006 which its data obtained from census.

Table 4: Model Validation Results

	Simulation Results	Census Of 2006	Error Percentage
Iran Population	70845700	70495872	0.496 %
Deaths in 2006	410541	408566	0.483 %
Average Annual Growth Rate 1996-2006	1.666	1.615	3.1 %

Results and low percentage of errors show the validity and accuracy of the model (Table 4).
Answering to this question: "Will be Iran's population growth rate in a few coming years descending?"
The results of the simulation model to predict annual population growth rate in Iran from 2006 to 2016 is shown below:

Table 5: Average Growth Rate and Forecast

10 years Census Intervals	Average Growth Rate
1956-66	3.127%
1966-76	2.714%
1976-86	3.905%
1986-96	1.963%
1996-2006	1.612%
2006-2016 (Simulated)	1.813%

The result shows that the ten-year average growth rate from 2006 to 2016 will be equal to 1.896. As a result of simulation, in the same pattern of Iran population behavior for future, the statement of "Iran population growth rate in near future years, not only will not descending but also has an obvious ascending trend" would be confirmed. This confirm by noticing to the baby boom in 1980's and large population which reaching to the childbearing years seems quite logical. Here the simulation results from 2006 to 2015 for growth rate are shown:

2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
1.812	1.845	1.866	1.872	1.866	1.848	1.819	1.782	1.737	1.686

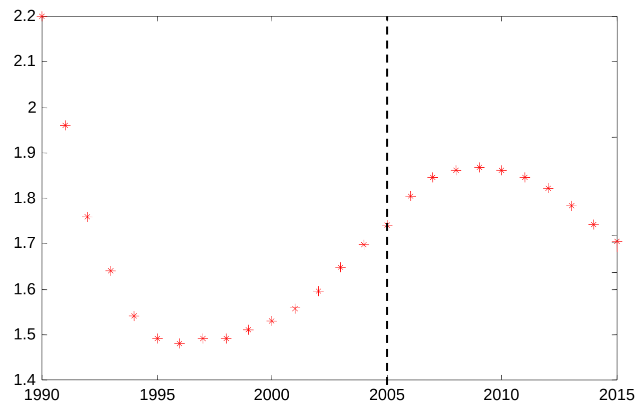

Figure 7: The year-on-year growth rate obtained by the model

To utilize this model in long term three scenarios for Iran Population is considered:

- First Scenario: Continuing current situation like what resulted from calculation of table 1.
- Second Scenario: Decreasing the total fertility (*TF*) to 1.8 based on forecasts from experts at *Population Studies and Research Centre in Asia and the Pacific*.
- Third Scenario: Incentive policies of Iranian government to increase the total fertility (*TF*), which is optimistically assumed to reach 3.2.

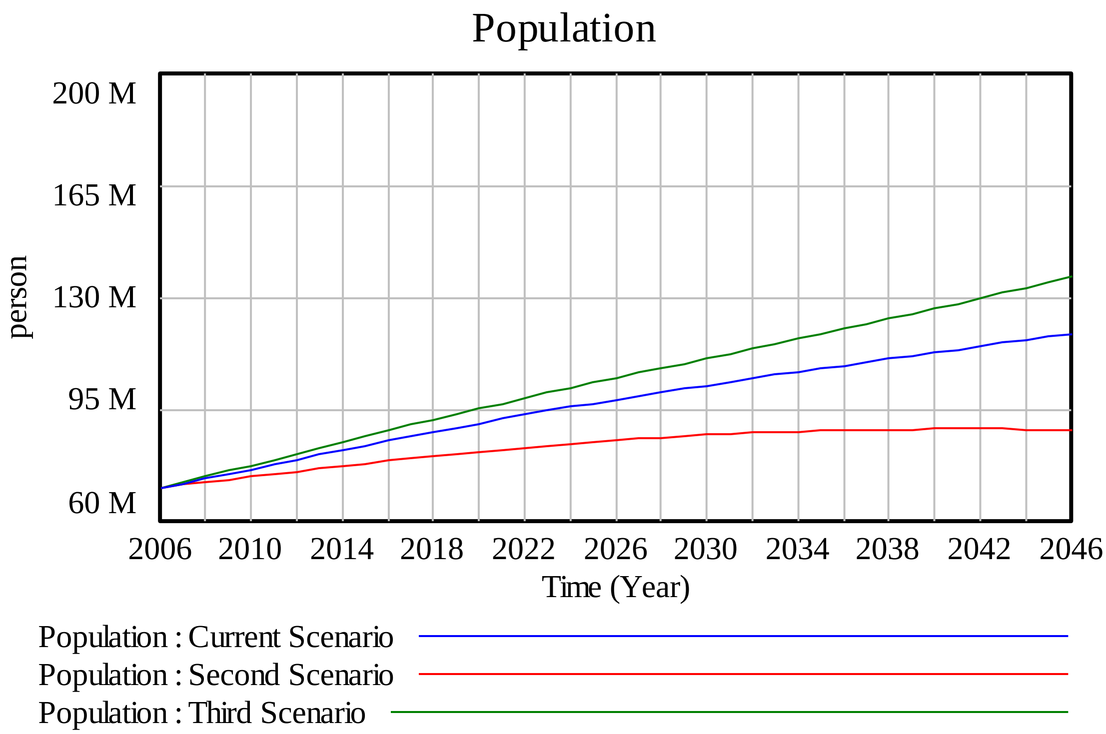

Figure 8: Three Scenarios for Iran Population

7-Conclusion

The aim of this research was to calculate current and future Iran population growth rate about which there are several different official estimations. The main reasons for existence of this variety are unsuitable forecasting method and insufficient and unreliable data. SD has been identified as a very powerful device for dealing with complexities at the level of the content of social and economic systems such as population transition by modeling and simulating them dynamically. In this study a novel combined adaptive process designed for modification of unreliable data with theory-based structural modeling involving endogenous and exogenous variables (SD model) and cohort component model was used to observe Iran demographic transition.

After literature review, the system dynamics model structure and different variables presented for population projection were described. Then the mathematical relations used in SD model to forecast the demographic components were explained. To deal with lack of precise data, these troubles and statistic data discussed then by processing data reached to useful parameters about fertility. In section 5 adaptive statistical data processing algorithm using SD model for data correction was proposed. This adaptive algorithm was a valuable method for solving the problem of data's inconsistencies. Parameters resulted from this adaptive method, were used in simulation and made a very accurate result. Finally model validation was observed and presented part of simulation results like forecasting population growth rate shows that ten-years average growth rate from 2006 to 2016 will increase and would be equal to 1.896.Then three different probable scenarios were simulated for Iran population. In this research, despite of some previous reports, we robustly conclude that Iran population growth rate in near future not only will not be descending but also has an obvious ascending trend. This fact can affect main planning of country and authorities should notice this result carefully.

References

Abbasi-Shavazi, M. J, and P. McDonald. 2005. *National and provincial-level fertility trends in Iran, 1972-2000*. Working Paper in Demography.

Ahlburg, D. A. 1982. How accurate are the US Bureau of the Census projections of total live births. *Journal of Forecasting* 1, no. 4: 365–74.

Booth, H. 2006. Demographic forecasting: 1980 to 2005 in review. *International Journal of Forecasting* 22, no. 3: 547–581.

Box, G. E.P, G. M Jenkins, and G. C Reinsel. 1976. *Time series analysis: forecasting and control*. Holden-day San Francisco.

Capece, J. Population Growth and Water Demand Model For Port LaBelle, Florida.

Cole, A, and J Flenley. 2008. Modelling human population change on Easter Island far-from-equilibrium. *Quaternary International* 184, no. 1 (6): 150-165. doi:10.1016/j.quaint.2007.09.019.

De Gans, H. 1999. *Population forecasting 1895-1945: the transition to modernity*. Kluwer Academic Pub.

Forrester, J. W, and J. Wright. 1961. *Industrial dynamics*. MIT press Cambridge, MA.

Keyfitz, N. 1985. Demography in the twenty-first century: the uses of forecasting. In *International Population Conference, Florence*, 1:59–81.

Keyfitz, N., and H. Caswell. 2005. *Applied mathematical demography*. Springer Verlag.

Klosterman, R. E. 1990. *Community analysis and planning techniques*. Rowman & Littlefield Publishers.

Lee, Elisa. 1992. *Statistical methods for survival data analysis*. 2nd ed. New York: Wiley.

Lutz, W., J. R Goldstein, and C. Prinz. 1996. Alternative approaches to population projection. *The future population of the world. What can we assume today*: 14–44.

Preston, S. H, P. Heuveline, and M. Guillot. 2001. *Demography: Measuring and modeling population processes*. Blackwell Pub.

Rosner, Bernard. 2006. *Fundamentals of biostatistics*. 6th ed. Belmont CA: Thomson-Brooks/Cole.

Schmitt, R. C. 1954. A Method of Projecting the Population of Census Tracts. *Journal of the American Planning Association* 20, no. 2: 102–102.

Smith, S. K, J. Tayman, and D. A Swanson. 2001. *State and local population projections: Methodology and analysis*. Plenum Pub Corp.

Sterman, John. 2000. *Business dynamics : systems thinking and modeling for a complex world*. Boston [u.a.]: Irwin/McGraw-Hill.

Van Buskirk, Ryffel, and Inc Associates. 2004. Population Model to Forecast Population Growth of Lehigh Acres Over Time to Build-out. *Wayne E. Daltry, FAICP Smart Growth Director Office of the Lee County Manager* (April).

Walonick, D. S. 2006. An overview of forecasting methodology. *Statpac, Inc*.