Journal of Information Technology Impact
Vol. 8, No. 1, pp. 11-24, 2008

Projectionmanager: A Knowledge Based Cohort Component Model for Population Projection in Nigeria

Olabode Olatubosun[1] Olabode Bola Titilayo[2]
Federal University of Technology, Akure
Nigeria

Abstract

Planning is an activity, which refers to the future. Population Projection provides future estimates of population sizes needed by planners. In Nigeria a number of factors have contributed negatively to having an accurate and reliable demographic data. Successive Nigerian Governments have embarked on Population Census in which huge amount of money was expended. Population projection can be made by examining the influence of demographic components (i.e birth, death and migration) of population as it affects population from time to time so as to determine future population. In order to bring out the implication of current trends in mortality, fertility and migration on population growth, a demographic model has been developed. A knowledge based system was used to implement the models developed on current population size and structure. In this study, a knowledge based system for population projection is developed with focus on the people of a community. The study involves using the secondary data from a population census survey in (Olabode 2005). The system is implemented using Visual Basic 6.0, Microsoft access and Microsoft excel. In the Evaluation of Population by Component method, only one form of age specific fertility and mortality was considered. The results show that, the projected population behaves relatively in a fashion described by the vital rates.

Keywords: Knowledge Based System, Population Projection, Cohort Component Method.

Introduction

The greatest asset of any nation is her citizen. The citizen of a nation is the net value of the birth and death together with migration. Politically, socially and economically the strength of a nation is measured by the size of its population. In (Heisel, 1969), accurate and up to date information on the component and patterns of change of population is a primary requisite for intelligent decision making in any society. It is a fundamental requirement that any nation should estimate her population size, rate of growth, and the components patterns of fertility and mortality. Planning for growth obviously requires the best possible knowledge of the characteristics and prospects of the population being planned for. Population statistics are derived mainly from the following traditional sources: viz. periodic censuses or enumeration, sample surveys or inquiries, vital registration systems and population registers, (Kpendekpo, 1979).

Human populations are subject to change in the number and age structure of their members. The changes operate through the processes of birth, deaths and the balance between immigration and emigration. (Kpendekpo, 1982) iterated that natural population or population models could be described and distinguished from each other by reference to the three types of variables expressing the forces of addition, depletion and the resulting age structure.

Population projections are essentially concerned with future growth. They may be prepared for the total population of nation, their principal geographic subdivisions or specific location within them. The principal characteristics for which projections need to be made are age and sex. Projection may also be made for various social and economic subgroups of the population and for other demographic aggregates. (Henry, 1976).

In most African countries, population projections have become essential items of information for both the government and the private sector. Any development plan should consider in advance, for example, what is going to be the size of the labour force, what will be the housing needs and the demand for schools and for health services. The information on these needs comes from data on the future population size, its age distribution, and its rate of growth. In Nigeria, population projections are particularly important due to current economic and social transformations and changes in population dynamics. To this end there is need to embark on a comprehensive study of demography in Nigeria with particular reference to population projection and come up with an information system that could help in real time information on Nigerian population. One objective of the research work is to formulate a demographic Model for the population growth and projection in Nigeria, with a view to determining how growth structure is affected by random (stochastic) event of death, birth and migration and to implement the model using knowledge based system.

Techniques for Modeling Demographic Data

Demography is the mathematical modeling and statistical investigations of population. It is concerned with the study of three factors, namely, fertility, mortality and migration of the people of a given region. One method of caring out population projection is by cohort component factors method.

Component Factors of Population Structure

There are only three components of change in the total population, namely, births, deaths, and migration. Mortality, that is, death is the permanent disappearance of all evidence of life at any time after birth has taken place. A death can occur only if a live birth has occurred. It is recognized in (Henry *et al*., 1978) that fetal death is employed in present demographic practice to embrace the events called stillbirths, miscarriages, and abortions in medicine or legal usage. The basic data on deaths for mortality studies for the statistically developed areas come from vital statistics registration system and less commonly from national population register system. The analysis of the death statistics from the vital statistics registration system depends on the availability of appropriate population data from a population census or survey or population estimates to be used as bases for computing rates of various kinds.

Life tables provide a description of the most prominent aspects of the state of human mortality. Essentially they illuminate and summarize the mortality experience of a population. In addition to this function, life tables also constitute one of the most important tools in demographic analysis, with wide applications (Kpendekpo, 1982). According to (Kpendekpo, 1982; Vladimir, 2002), the Human Life-Table Database (HLD) is a collection of population life tables for a multitude of countries for many years. Most of the HLD life tables are life tables for national populations, which have been officially published by National Office of statistics. In Nigeria, the Federal Office of Statistics (FOS) is responsible for the creation and maintenance of the HLD.

Population censuses or surveys sometimes provide fertility data to be used independently, either as supplement to vital statistics rate or as substitutes where birth registration does not exist or is inadequate. Birth registration records usually provide information on a few demographic characteristics of the parents of children born in the relevant period; where as census inquiries provide data on characteristics of all enumerated women or men who have ever had children or are of child bearing age.

In (Henry *et al* 1978), the following three measures of fertility are frequently available from census reports.

a. Ratio of young children to women of child bearing age.
b. Classification of women by the number of their own young children living in the same household
c. Number of children ever born to women.

Migration is the third basic factor affecting change in the population of an area. The importance of migration with respect to how it affects growth and decline of populations which subsequently leads to change in demographic characteristics of the areas of origin and the area of destination has long been recognized. According to (Henry *et al,* 1978), migration is an important element in the growth of the population and the labor force of an area.

One may distinguish six classes of migration data corresponding to these several sources as follows.

a. Statistics collected on the occasion of movement of people across international borders, mostly as a bye product of the administrative operations of frontier control.
b. Passenger statistics obtained from the list of passengers on sea or air transport manifests
c. Statistics of passports and applications for passports, visa, work permits, and so on
d. Statistics obtained in connection with population registers
e. Statistics obtained in census or periodic national surveys on the basis of inquiries regarding previous residence, place of birth, or citizenship.
f. Statistics collected in special or periodic inquiries regarding migration, previous and present residence, or citizenship, such as a registration of aliens or a count of citizens overseas.

Population Projection by the Component Method

Calculation of the expected population size and age-sex composition is performed according to three scenarios, reflecting different assumptions about possible changes in fertility, mortality and migration levels. This study adopts the cohort component projection method, which consists of the following steps:

a. starts off with the population by age/sex as of a base year,
b. then applies the assumed survival rate by age/sex,
c. apply the international immigration rate by age/sex,
d. apply female fertility rate by age, in addition to the sex ratio at birth, to determine the future population. The basic procedure for the cohort component method is conceptualized in figure 1

To calculate the population in (t+1)th year from the known population in t^{th} year. The population aged one-year old and older in year t+1 can be found by applying the pertinent survival and international migration rates both by age to the population in year t. The female population multiplied by its age-specific fertility rate and sex ratio at birth yields the number of births by sex. More specifically, with regards to each population by age/sex from age 1 to 99 and “100 or over,” the population of age *x* is multiplied by the assumed survival rate for that age group, and is adjusted by the number of international immigrants from age *x* to age x+1, to find the population of age x+1 on the following year.

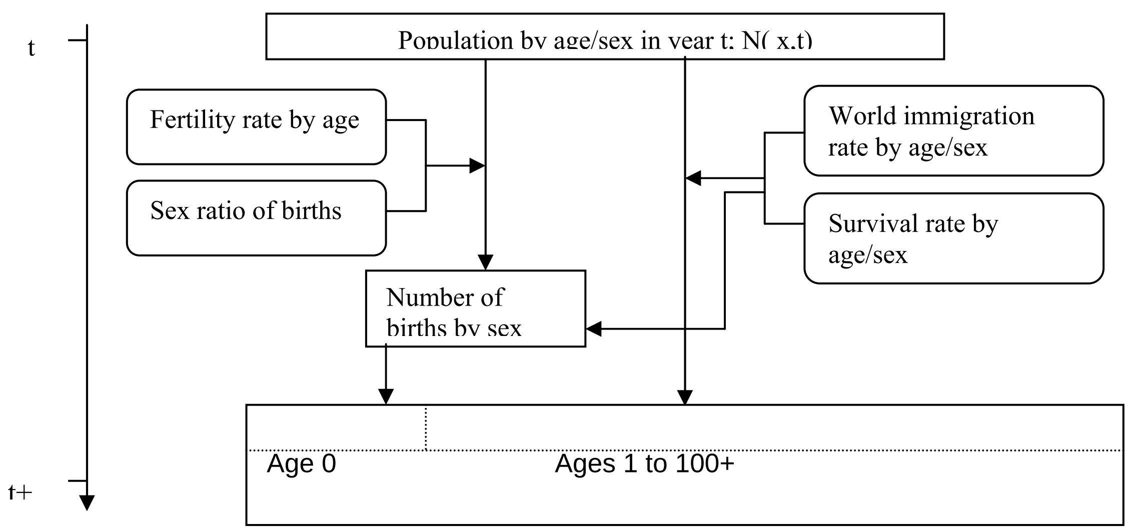

Figure 1. Procedures for Cohort Component Population projection

System Design

The conceptual diagram of ProjectionManager presented is viewed as a system comprising of the Knowledge Base, Inference engine and Decision Support engine

Year

Knowledge Base

The Knowledge Base is composed of the quantitative (structured) and qualitative (unstructured) knowledge of population dynamic acquired by the human. The knowledge base is a network of semantically related static and dynamic objects, each of which is modeled, in a relational form. The structured knowledge is concerned with facts, rule and events of human population dynamics, which are commonly agreed upon by experts in demography. The unstructured knowledge is that knowledge which is acquired by demographic expert from experience and population census survey. It is the heuristic knowledge or that which is acquired by good practice, guess and judgment.

The ProjectionManager uses one or more data stores (databases) to provide relevant information to the knowledge base of the system. Data used for the implementation of the ProjectionManager includes the human life table, population distribution by age and sex. ProjectionManager acquires data from various internal and external sources. Internal data refer to databases from secondary data of demographic data survey, such as the results from National Demographic Health Survey (NDHS), Human Life Table, fact about population, and so on. External data refers to sources from primary sources for example the population census.

Inference Engine

The inference engine is concerned with the adoption of an appropriate line of reasoning, leading to the solution of a given demographic problem or the formulation of a body of consultative advice on a given demographic phenomenon. Carrying out population projection program involves a determination of the number of different levels of geographic details.

The inference procedure generates the following statistics and population distribution by Age Group and Sex.

a. Population distribution by Age Group and Sex.
b. Human life Table
c. Survival rate by age on base population
d. Fertility rate by age on base population
e. Migration rate by age on base population

The demographic statistics for population distributions by sex can largely be determined by the age group of sampled population, such that there is the existence of a function F_k , k = 1, 2, 3, …, n which is a vector defined on age group of population denoted by g_i, i = 1, 2, 3, …, m where n is the number of statistical/demographic information

m is the number of age group of the sample population

Let F_k represent the function defined on g_i to generate the population distribution by age group and sex. Hence, $F_k(g_i)$ produces a matrix defined by:

$$A_{i,j} ; \quad i = 1, 2, \ldots, m ; \quad j = 1, 2$$

where m is the number of age group and $A_{i,j}$ is defined by:

$$A_{i,j} = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \\ " & " \\ " & " \\ " & " \\ a_{m,1} & a_{m,2} \end{pmatrix}$$

$a_{i,j}$ represents the frequency of ith age group for jth type of sex and it is generated by

$$a_{i,j} = \sum_{k=1}^{n} h_k(g_i); \quad \text{from population census data}$$

where

$i = 1, 2, 3, 4, ..., 16; \quad j = 1, 2$

n represent the population census sample size

Let there exist a function age specific death statistics denoted by D_k , k = 1, 2, 3, …, m which is a vector defined on age group denoted by g_i i = 1, 2, 3, …, m , where m is the member of age group

Let T_k represents the function defined on g_i to generate the Human life Table from the given D_k.

Where,

$T_k(g_i) = [a_1, a_2\ a_3, a_4, a_5, a_6, a_7, a_8]$

a_1 = Central age specific death rate

a_2 = Probability of a person aged x dying within the interval x to x+n, where n could be 5

a_3 = Probability of a person with aged x surviving through the interval x to x+n

a_4 = Survivors of a cohort of live born babies to the exact age x

a_5 = Deaths experienced by the life table cohort within the interval x to x+n

a_6 = Number of person-years lived by the cohort during the interval between specified birthdays, that is, between x and x+n years

a_7 = Total stationary or life table population at age x and all higher ages

a_8 = Expectation of life remaining to persons who attain the exact age x

Hence, $T_k(g_i)$ produces a matrix defined by:

$A_{i,j}$ $i = 1, 2, \ldots, m$; $j = 1, 2, 3, 4, 5, 6, 7, 8$

where m is the number of age group and $A_{i,j}$ is defined by:

$$A_{i,j} = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} & a_{1,5} & a_{1,6} & a_{1,7} & a_{1,8} \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,4} & a_{2,5} & a_{2,6} & a_{2,7} & a_{2,8} \\ '' & '' & '' & '' & '' & '' & '' & '' \\ '' & '' & '' & '' & '' & '' & '' & '' \\ '' & '' & '' & '' & '' & '' & '' & '' \\ a_{m,1} & a_{m,2} & a_{m,3} & a_{m,4} & a_{m,5} & a_{m,6} & a_{m,7} & a_{m,8} \end{pmatrix}$$

$a_{i,j}$ represents the computational values of ith age group for jth life table attribute as defined above.

Typical life table is presented in Table 1. The table shows the behavior of death as in affect a cohort population.

In order to make a population projection, it is necessary to first establish the size and composition of the population at some point in time as the base for the projection. The component method can then be used to project the population from the base to future dates by considering the influence of mortality, fertility and migration. The data needed for the cohort component method used in projection are as follows:

Table 1. Evaluation of Population by Component Method

age group	mx	qx	px	lx	dx	nLx	Tx	ex
Under 1	0.184	0.163	0.837	10000.00	1632.40	8857.32	366821.20	36.68
1-4	0.053	0.186	0.814	8367.60	1556.98	29266.54	357963.90	42.78
5-9	0.010	0.046	0.954	6810.62	316.00	33263.09	328697.40	48.26
10-14	0.006	0.028	0.972	6494.62	179.34	32024.74	295434.30	45.49
15-19	0.006	0.031	0.969	6315.28	198.91	31079.13	263409.50	41.71
20-24	0.007	0.034	0.966	6116.37	207.44	30063.27	232330.40	37.99
25-30	0.007	0.032	0.968	5908.94	191.83	29065.10	202267.10	34.23
31-34	0.009	0.042	0.958	5717.11	240.66	27983.88	173202.00	30.30
35-40	0.010	0.047	0.953	5476.45	259.32	26733.93	145218.20	26.52
41-44	0.016	0.075	0.925	5217.13	389.24	25112.52	118484.20	22.71
45-50	0.024	0.112	0.888	4827.88	542.25	22783.77	93371.73	19.34
51-54	0.026	0.123	0.877	4285.63	525.02	20115.60	70587.95	16.47
55-60	0.031	0.144	0.856	3760.61	542.59	17446.58	50472.36	13.42
51-64	0.048	0.213	0.787	3218.02	685.73	14375.80	33025.77	10.26
65-70	0.063	0.273	0.727	2532.30	690.08	10936.28	18649.98	7.36
71-74	0.099	0.397	0.603	1842.22	732.16	7380.68	7713.69	4.19
75+	0.130	1.000	0.000	1110.05	1110.05	333.02	333.02	0.30

a. Population by age and sex.
b. Assumed survival rate by age and sex.
c. Assumed female age-specific fertility rate.
d. Assumed rate of international immigration by age and sex.
e. Sex ratio at birth.

Application of Survival rate by Age group on Base Population

Survival rate is applied on the base population so as to determine what the population will transform to due to the influence of mortality on the cohort within the time specification.

Let λ represents the function defined on g_i to generate future population distribution by age group based on mortality factor in a cohort. The essential variables needed for the computation includes the following:

a. Initial population of the cohort
b. Survival ratio during interval t and t+n

Hence, $\lambda(g_i)$ produces a matrix defined by

$$\lambda_{i,j}; \quad i = 1, 2, \ldots, m, \quad j = 1, 2, 3$$

where m is the number of age group

$\lambda_{i,j}$ is defined by:

$$\lambda_{i,j} = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ " & " & " \\ " & " & " \\ " & " & " \\ a_{m,1} & a_{m,2} & a_{m,3} \end{pmatrix}$$

$a_{i,1}$ – Initial population of ith age group

$a_{i,2}$ – Computed value survival ratio of ith age group estimated between 't' and 't+n' from human life table

$a_{i,3}$ – Computed value final population of ith age group at the end of the interval 't' and 't+n', which can be obtained by:

$$a_{i,3} = a_{i,1} * a_{i,2}$$

Table 2 presents a report of the effect of mortality on a base population

Table 2. Population Projection by Component Method

Age group	Initial Population	Survival Ratio	Ending Population	Child bear Population	Age specific	Birth
01-04	29	0.8139	198.0312			
05-09	44	0.9536	23.6031			
10-14	63	0.9724	41.9584			
15-19	95	0.9685	61.2612	62.1306	0.146	45.35534
20-24	124	0.9661	92.0075	93.50375	0.026	12.15549
25-30	118	0.9675	119.7964	121.8982	0.259	157.8582
31-34	101	0.9579	114.165	116.0825	0.226	131.1732
35-40	71	0.9526	96.7479	98.87395	0.158	78.11043
41-44	56	0.9253	67.6346	69.3173	0.09	31.19279
45-50	44	0.8877	51.8168	53.9084	0.074	19.94611
51-54	41	0.8775	39.0588			
51-64	34	0.8557	29.9022			
55-60	38	0.7869	35.9775			
65-70	32	0.7275	29.0938			
71-74	24	0.6026	23.28			
75+	30	0	14.4624			

Application of Fertility rate by Age group on Base Population

The fertility rate is applied on the base population so as to determine the birth that could occur in the cohort within the time specification.

Let β represent the function defined on g_i to generate future population distribution by age group based on fertility factor in a cohort. The essential variables needed for the computation includes the following:

c. Initial population of the cohort
d. Average annual age specific fertility rate during interval 't' and 't+n'

Hence, $\beta(g_i)$ produce a matrix defined by

$$\beta_{i,j} \quad i = 1, 2, \ldots, m, \quad j = 1, 2, 3, 4$$

where m is the number of age group

$\beta_{i,j}$ is defined by:

$$\beta_{i,j} = \begin{pmatrix} b_{1,1} & b_{1,2} & b_{1,3} & b_{1,4} \\ b_{2,1} & b_{2,2} & b_{2,3} & b_{2,4} \\ " & " & " & " \\ " & " & " & " \\ " & " & " & " \\ b_{m,1} & b_{m,2} & b_{m,3} & b_{m,4} \end{pmatrix}$$

$b_{i,1}$ – Initial population of ith age group

$b_{i,2}$ – Computed value of average population exposed to childbearing of ith age group between 't' and 't+n', which can be obtained by:

$$b_{i,2} = \beta(b_{i1} + b_{i3})/2$$

$b_{i,3}$ – Computed value of average annual age specific fertility rate of ith age group during the interval 't' and 't+n'

$b_{i,4}$ – Computed value of births, which can be obtained by:

$$b_{i,4} = 5(b_{i,2} * b_{i,3})$$

Application of Migration rate by Age group on Base Population

A number of new problem enter population projection with the recognition of migration. Unlike fertility and mortality, which take time for the impact of decision on them to be felt, migration is directly influenced by government action and other socio-economic development and by wide variety of factors, which can easily change quite suddenly from one year to the next. For this reason, there is no really sound basis for projecting migration data into the future. In a sense, all information on migration data is historical and nothing more. It gives an idea of what has happened but gives no clues as to what will happen in the future. In this research work, the influence of migration has been taken to be zero.

Combining the influence of Mortality, Fertility and Migration on Base population

Once the projection on mortality, fertility and migration has been determined, then the total population projection for a future date can be obtained. This scenario can then be graduated to obtain the population for a later date.

Therefore, if P_{t+1} represents the Population of age group g_i, then the summation of all the components of population, will give the projected population P_i at time 't+n' as:

$$P_{i+1} = \lambda(a_{i,3}) + \beta(a_{i,4}) + y$$

Where

$\lambda(a_{i,3})$ = Final population of ith age group on the influence of death between time 't' and 't+n'.

$\beta(b_{i,4})$ = Births ith age group that occur within the population between time 't' and 't+n'

y = 0 since influx is assumed equal to exflux

The influence of migration between time 't' and 't+1' is assumed to be zero for the reasons given on migration rate.

Total population in the furture time t+n = $\sum_{i=1}^{m} p_i$

where i = 1, 2, 3, …, m, n is interval size

Table 3 presents a typical population projection while the graph represents a pictorial view.

Results of Analysis

Table 2. Population Projection for 4 consecutive Period using a Sampled population Structure

Age group	Initial Population 2004	2009	2014	2019	2024
1-4	29	198	184	152	158
5-9	44	24	161	150	124
10-14	63	42	23	154	143
15-19	95	61	41	22	149
20-24	124	92	59	40	21
25-30	118	120	89	57	38
31-34	101	114	116	86	55
35-40	71	97	109	111	82
41-44	56	68	92	104	106
45-50	44	52	63	85	96
51-54	41	39	46	56	76
55-60	38	36	34	40	49
51-64	34	30	28	27	32
65-70	32	29	26	24	23
71-74	24	23	21	19	18
75+	30	14	14	13	11
Total	944	1039	1107	1140	1182
% cent	17.4	19.2	20.4	21.1	21.8

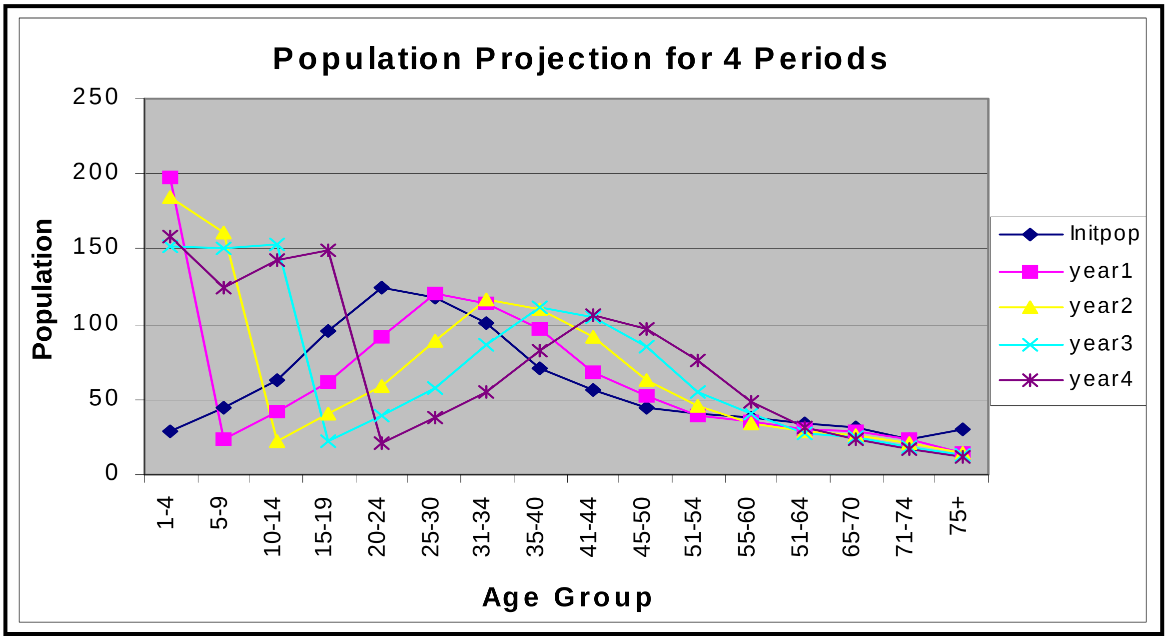

Figure 3. Line Graph of Population Projection by Component Method for 4 Periods

Decision Support Session

The Decision support system of the ProjectionManager is a procedure that provides communication between the user and other components of the ProjectionManager. It combines the knowledge base with the inference procedure to enhance strategic decision making. The DSS of the ProjectionManager applies the Add-in routine of the visual Basic to integrate both the Microsoft access and crystal report manager into the ProjectionManager.

The cohort component population projection method follows each cohort of people of the same age throughout its lifetime according to its exposure to mortality, fertility, and migration. Starting with a base population by sex and age, the population at each specific age is exposed to the chances of dying as determined by projected mortality levels and patterns by sex and age. Once deaths are estimated, they are subtracted from the population, and those surviving become older. Fertility rates are projected and applied to the female population in childbearing ages to estimate the number of births every year. Each cohort of children born is also followed through time by exposing it to mortality.

Finally, the component method takes into account any in-migrants who are incorporated into the population and out-migrants who leave the population. Migrants are added to or subtracted from the population at each specific age. The whole procedure is repeated for each year of the projection period, resulting in the projected population by age and sex, as well as birth and death rates, rates of natural increase, rates of population growth, and other summary measures of fertility, mortality, and migration for each year.

Conclusion

In order to bring out the implication of current trends in mortality, fertility and migration in population growth, it is useful to apply these factors to population of the country in successive years in the future, starting with the present population size and structure. Projections are like microscopes that magnify the differences in a given period and so help with analysis and the understanding of current rates.

In this study, a ProjectionManager for population projection is developed with focus on the people of a community. The study involves using the secondary data from a population census survey in (Olabode 2005). The system is designed and implemented using Visual Basic 6.0, Microsoft access and Microsoft excel.

Attempt has been made to design a ProjectionManager with embedded demographic model proposed for its practical implementation. In the Evaluation of Population by Component method, only one form of age specific fertility and mortality is considered, more work is still required on population projection by applying various the sets of mortality, fertility and migration rates.

References

Kpendekpo G. M. K. (1982). *Essentials of Demographic Analysis for Africa*. New Hampshire: Heinemann Educational Books Inc.

Olabode, O.. & Akinyokun, O. C. (2002). *Computer Aided System for Demography in Nigeria*, Master Thesis, Federal University of Technology, Akure.

United Nations. (2001). *The State of the Demographic Transition In Africa: Executive Summary*, Economic Commission for Africa, Joint ECA/OAU/ADB Secretariat in Collaboration with UNFPA. FSSDD/ICPD/FC.4/01/3, Yaounde, Cameroon.

Takahashi, S., Kaneko, R., Ishikawa, A., Ikenoue, M., & Mita, F. (1999). Population Projections for Japan: Methods, Assumptions and Results. *Review of Population and Social Policy, 8,* 75-115.

Heisel, D. F. (1969). Measuring Current Population Changes. *Population of Tropical Africa* (3rd ed., pp. 34-49). London: Lowe & Bryone (Printer) Ltd.

Henry, M. A. (1978). *Population of Tropical Africa* (3rd ed.). London: Lowe & Bryone (Printer) Ltd.

Kpendekpo G. M. K. (1987). *Essentials of Demographic Analysis for Africa*. New Hampshire: Heinemann Educational Books Inc.

---

[1] Dr. Olabode Olatubosun is with the Computer Science Department, Federal University of Technology, Akure, Nigeria. P.M.B 704, Akure, Ondo State, Nigeria. Email: olabode_olatubosun@yahoo.co.uk

[2] Dr. Olabode Bola Titilayo is with the Department of Mathematical Sciences, Federal University of Technology, Akure, Nigeria.