ISSN 0976 – 1411

Available online at www.internationaleJournals.com

International eJournals

INTERNATIONAL eJOURNAL OF MATHEMATICS AND ENGINEERING

International eJournal of Mathematics and Engineering 7 (2010) 86-95

www.InternationaleJournals.com

MODELING AND PREDICTING URBAN MALE POPULATION OF BANGLADESH: EVIDENCE FROM CENSUS DATA

Authors: [1]Md. Rafiqul Islam and [2]A.B.M.Rabiul Alam Beg

[1]Md. Rafiqul Islam
Department of Population Science and Human Resource Development
University of Rajshahi, Rajshahi-6205, Bangladesh.
E-mail: rafique_pops@yahoo.com

[2]A.B.M.Rabiul Alam Beg
James Cook University
Australia.
E-mail: rabiul.beg@jcu.edu.au

Corresponding Author:
Md. Rafiqul Islam

Abstract.

This paper predicts the urban male population of Bangladesh using geometric growth rate method. The predictions are computed in three stages. In the first stage, the urban male population for the years 1981 and 1991 was predicted using the smoothed data for those years. The first stage predictions were obtained by using generalized negative exponential model estimated by nonlinear least squares method. While the urban male population for the year 2001 was predicted by a linear model. Using the cross validation predictive power (CVPP) and R^2 this article constructed shrinkage coefficient which determines the adequacy of the first stage predicted values. These predicted values are then used in the second stage to estimate the geometric growth rate for different age groups. Finally, considering year 2001 urban male population as the base period and using the estimated geometric growth rate of the second stage the predictions of the urban male populations are computed for 2002 through to 2031.

Key words: Urban population, linear forecast model, negative exponential forecast model, nonlinear least square estimation, geometric growth rate, cross validity predictive power (CVPP), R^2, shrinkage

Running Head Line: Modeling and Predicting Urban Male Population of Bangladesh

1. Introduction

Bangladesh is a poor country mainly depends on foreign aid. People of Bangladesh always struggle for survival. Rural people are generally hard hitted financially than the urban Bangladesh. There is a trend of people moving from rural area to urban Bangladesh for better life. Population in urban area is growing since the independence in 1971. It is to be noted that the male population of Bangladesh are the main driving force of income generation. The heavy influx of people in the urban areas is a burden to the Bangladesh government. Thus, the Bangladesh government needs firm policy to accommodate its fast growing urban population by providing shelter, education, health, clean water and etc.

The objective of this paper is to provide efficient prediction of urban male population at different age groups using a three stage procedures for the government's infrastructural policy decisions.
Since the pattern of age structure of population may change from various demographic variables and from country to country. Islam et al. (2003) found that the age structure of Bangladesh follows a modified negative exponential model. While the age structure for the population of both sexes of Bangladesh follows a generalized negative exponential model (Islam, 2005). Islam and Beg (2009) predicted the population of Dhaka District of Bangladesh using cubic polynomial during 2002-2031. A study by Bangladesh Bureau of Statistics (BBS, 1994) used logistic function to project life expectancy at birth up to year 2006. But the objective of this paper is to predict the urban male population. Therefore, following Islam et al. (2003), Islam (2005) and Islam and Beg (2009), this paper uses the most recent data and the geometric growth rate approach in three stages to predict the urban male population of Bangladesh for the years 2002 through to 2031. This paper also estimates the number of years needed to double the urban male population for different age groups using the predicted values of 1991 and 2001. So far none of the previous studies of Bangladesh considered this kind of analysis.
This paper is organized as follows. Section 2 presents the data and the data sources. The models and the methodological issues and various tests are described in section 3. Empirical results and discussion of the results are reported in section 4. Finally section 5 concludes the paper.

2. Data and Data Sources

To fulfill the objective of this paper the secondary quinquennial age data on urban male population of Bangladesh have been taken from various issues (BBS 1984, 1994, 2003) of Bangladesh census staring from 1981. Note that the census happens in Bangladesh in every ten years and the last census took place in 2001. Using the last three census data this paper forms the predictions for the years 2002 through to 2031. The age group considered in this study are (0-4, 5-9, 10-14, 15-19, ………, 65-69, 70 and above). The population values are in thousand. The data are shown in Table-1

Table 1. Observed urban male population (in thousands) by age group of Bangladesh for the census years 1981, 1991, and 2001.

Age group	Census year		
	1981	**1991**	**2001**
0-4	955	1438	1588
5-9	935	1524	1662
10-14	947	1360	1882
15-19	744	1038	1684
20-24	741	1115	1594
25-29	709	1104	1508
30-34	528	859	1251
35-39	450	806	1127
40-44	360	597	909
45-49	259	418	655
50-54	233	329	514
55-59	130	195	284
60-64	153	212	301
65-69	65	102	155
70 and above	161	204	319
Total	7370	11301	15433

3. Models and the Methodological Issues

If the population due to ages is presented in the graph paper, then it is found that thre are some short of distortions exist in the data aggregate. For that reason, data needs to be adjusted. Therefore, the data have been smoothed out to eliminate any abnormalities from the data. The technique used is “4253H twice” of Velleman (1980). This is the default in the Minitab package - version 12.1. The smoothed series are then used to predict the urban male population of Bangladesh using negative exponential for the years 1981 and 1991 data and the linear model for the year 2001 data. The smoothed series for the years 1981, 1991, and 2001 for different age groups are given below.

Table 2. Smoothed urban male population

Age group	Census year		
	1981	**1991**	**2001**
0-4	955	1467	1595
5-9	934	1437	1645
10-14	881	1341	1674
15-19	805	1217	1667
20-24	735	1118	1607
25-29	659	1027	1480
30-34	554	910	1304
35-39	446	760	1106
40-44	353	595	897
45-49	275	438	684
50-54	206	312	483
55-59	158	228	339
60-64	140	189	281
65-69	138	181	274
70 and above	138	181	274
Total	7377	11401	15310

The models

Both the original and smoothed age structure of urban male population display negative exponential patterns in terms of different age groups for the years 1981 and 1991. Therefore, this study employs a generalized negative exponential model to the data for prediction purpose. While a linear prediction model fits the observed data well for the year 2001. The models considered for the first stage predictions are the following.

(generalized negative exponential) $y_i = c + e^{(-ax_i + b)} + u_i$ (1)

(Polynomial) $$y_i = \alpha_0 + \sum_{j=1}^{p} \alpha_j x_i^j + \eta_i \quad (2)$$

With $p = 1$ in (2) produces a linear model.

Where with reference to the present research, y_i represents the urban male population of the i-th age group, x_i is the mid value of the i-the age group, j =1,2…, p is the order of the polynomial, $c, a, b, \alpha_0, \alpha_1, ..., \alpha_p$ are parameters, and

u_i and η_i are normal random variables with mean zero and constant variance of model (1) and model (2) respectively.
The models are estimated by using nonlinear least squares and ordinary least squares methods available in STATISTICA. Note that the ordinary least squares is applied to estimating the polynomial model (2).

Geometric growth rate method

Geometric growth rate is estimated by using the following equation.

$$\hat{P}_{t_2}^{m-m+5} = \hat{P}_{t_1}^{m-m+5}\{1+r^{m-m+5}\}^{t_2-t_1} \qquad (3)$$

Where, $\hat{P}_{t_1}^{m-m+5}$ is the predicted initial population at time t_1 for the age group m to $m+5$; $\hat{P}_{t_2}^{m-m+5}$ is the predicted terminal population at time t_2 for the age group m to $m+5$; r^{m-m+5} is the intercensal annual growth rate of the age group m to $m+5$; and (t_2-t_1) is the time interval between intercensal period.
The r^{m-m+5} is computed for different age group from (3) as follows.

$$r^{m-m+5} = \text{Antilog}\left[\left(\frac{1}{t_2-t_1}\right)\log_e\left(\frac{\hat{P}_{t_2}^{m-m+5}}{\hat{P}_{t_1}^{m-m+5}}\right)\right]-1 \qquad (4)$$

Years 1991 and 2001 are considered as the initial and the terminal urban predicted male populations respectively in estimating the age specific growth rate by equation (4).
For prediction purpose, year 2001 census is treated as the base male population. The intercensal annual growth rate during 1991-2001 is used in this study assuming fertility and mortality remain unchanged during the forecast period. Estimation of the intercensal annul geometric growth rate for different age groups is computed based on the first-stage predicted data for 1991 and 2001.
Model evaluation criteria used are the usual regression t-test, and R^2. For model validation this paper uses cross validation predictive power (CVPP) denoted, ρ_{cv}^2 computed by $\rho_{cv}^2 = 1-\dfrac{(n-1)(n-2)(n+1)}{n(n-k-1)(n-k-2)}(1-R^2)$, where n is the number of classes, k is the number of regressors in the model, and R^2 is the coefficient of determination in the first stage of estimation. The shrinkage of the model is equal to the absolute value of $\lambda = (\rho_{cv}^2 - R^2)$, Steven (1996). The shrinkage criterion asserts that the better predictions are obtained if λ approaches zero.

4. Empirical Results and Discussion of the Results

This study used both the smoothed and the original series to fit generalized negative exponential model for the years 1981 and 1991 for all age groups. While a linear model is fitted to the smoothed and the original data for the year 2001 only. This study found that the smoothed data has the better predictions. Consequently this current study has used the smoothed series in the first stage of prediction. The estimated models are given below.

1981:

$$y_i = -916.170 + \exp(-0.0099x_i + 7.599)$$
$$t-ratio: (-9.45486) \quad (-11.05264) \quad (168.5038)$$

$$R^2 = 0.97454 \qquad \lambda = 0.0062236$$

1991:

$$y_i = -2882.9 + \exp(-0.0059x_i + 8.415)$$
$$t-ratio: (-23.10) \quad (-16.40506) \quad (319.0029)$$

$$R^2 = 0.97737 \qquad \lambda = 0.00564$$

2001:

$$y_i = 1961.4 - 25.086x_i$$
$$t-ratio: \quad (84.36) \quad (-1.95)$$

$$R^2 = 0.9271 \qquad \lambda = 0.01779$$

All of the estimated coefficients are statistically significant at the conventional level for all of the above estimated models. The R^2 of each model is high. Moreover, the above models provide good predictions based on the shrinkage criterion λ. Therefore, these models can be adopted for prediction.

The predicted values for the years 1991 and 2001 are then used to estimate the geometric growth rates for each age group. Which in turns these rates are used to predict the urban male population of Bangladesh for the years 2002 through to 2031 considering original 2001 census as the base.

This study has also estimated the number of years required to double the population at different age groups, which is provided in the following Table 3.

Table 3. Estimated geometric growth rate and the number of years required to double the urban male population at each age group.

Age group	The estimated geometric growth rate	Approximate number of Years required to double the urban male population
0-4	0.020	36
5-9	0.021	33
10-14	0.023	30
15-19	0.025	28
20-24	0.027	26
25-29	0.029	24
30-34	0.032	22
35-39	0.034	21
40-44	0.037	19
45-49	0.040	18
50-54	0.043	16
55-59	0.048	15
60-64	0.054	13
65-69	0.066	11
70-above	0.105	7

This table indicates that the aged male population will grow at a faster rate than the younger. The main reason for this may be due to the fact that there is a growing tendency of late marriage among males and also there is a tendency of the females to be involved in the payroll.

Using this estimated growth rate and 2001 census as the base this paper predicts the urban male population which is given below in the Table 4.

Table 4. Predicted urban male population (in thousands) for the years 2002 through to 2031

Age Group	Year				
	2002	2003	2004	2005	2006
0-4	619.025	1650.656	1682.905	1715.784	1749.306
5-9	1697.598	1733.958	1771.097	1809.031	1847.778
10-14	1925.920	1970.865	2016.858	2063.925	2112.091
15-19	1726.609	1770.296	1815.089	1861.015	1908.103
20-24	1637.564	1682.318	1728.295	1775.529	1824.053
25-29	1552.394	1598.095	1645.141	1693.573	1743.430
30-34	1290.607	1331.467	1373.621	1417.110	1461.976
35-39	1165.360	1205.025	1246.040	1288.452	1332.307
40-44	942.3066	976.8335	1012.626	1049.729	1088.192
45-49	680.9320	707.8907	735.9166	765.0522	795.3413
50-54	536.1559	559.2669	583.3740	608.5203	634.7505
55-59	297.5267	311.6976	326.5435	342.0964	358.3902
60-64	317.3269	334.5393	352.6854	371.8158	391.9839
65-69	165.2926	176.2686	187.9735	200.4556	213.7666
70 and above	352.5231	389.5691	430.5082	475.7496	525.7452

	Year				
Age Group	**2007**	**2008**	**2009**	**2010**	**2011**
0-4	1783.482	1818.326	1853.851	1890.070	1926.996
5-9	1887.354	1927.779	1969.069	2011.244	2054.322
10-14	2161.380	2211.820	2263.436	2316.258	2370.312
15-19	1956.383	2005.884	2056.637	2108.675	2162.030
20-24	1873.904	1925.117	1977.730	2031.781	2087.309
25-29	1794.755	1847.591	1901.982	1957.975	2015.616
30-34	1508.262	1556.013	1605.277	1656.100	1708.532
35-39	1377.655	1424.546	1473.033	1523.171	1575.015
40-44	1128.064	1169.398	1212.245	1256.663	1302.708
45-49	826.8295	859.5643	893.5952	928.9734	965.7522
50-54	662.1113	690.6516	720.4220	751.4757	783.8680
55-59	375.4600	393.3428	412.0774	431.7042	452.2659
60-64	413.2459	435.6612	459.2924	484.2054	510.4697
65-69	227.9615	243.0990	259.2416	276.4562	294.8139
70 and above	580.9948	642.0505	709.5224	784.0848	866.4828

	Year				
Age Group	**2012**	**2013**	**2014**	**2015**	**2016**
0-4	1964.644	2003.028	2042.161	2082.059	2122.737
5-9	2098.322	2143.265	2189.171	2236.060	2283.953
10-14	2425.627	2482.234	2540.161	2599.440	2660.103
15-19	2216.734	2272.823	2330.331	2389.293	2449.748
20-24	2144.354	2202.959	2263.165	2325.016	2388.558
25-29	2074.953	2136.038	2198.921	2263.655	2330.295
30-34	1762.624	1818.428	1876.000	1935.394	1996.668
35-39	1628.624	1684.057	1741.378	1800.649	1861.937
40-44	1350.441	1399.922	1451.216	1504.390	1559.512
45-49	1003.987	1043.736	1085.058	1128.017	1172.676
50-54	817.6566	852.9016	889.6658	928.0147	968.0167
55-59	473.8069	496.3739	520.0158	544.7837	570.7312
60-64	538.1586	567.3494	598.1236	630.5671	664.7704
65-69	314.3907	335.2674	357.5303	381.2717	406.5895
70 and above	957.5399	1058.166	1169.367	1292.253	1428.054

	Year				
Age Group	**2017**	**2018**	**2019**	**2020**	**2021**
0-4	2164.209	2206.491	2249.600	2293.550	2338.360
5-9	2332.872	2382.838	2433.875	2486.005	2539.252
10-14	2722.181	2785.708	2850.717	2917.244	2985.323
15-19	2511.733	2575.285	2640.446	2707.256	2775.756
20-24	2453.837	2520.899	2589.795	2660.573	2733.286
25-29	2398.897	2469.518	2542.218	2617.058	2694.102
30-34	2059.883	2125.098	2192.379	2261.790	2333.398
35-39	1925.312	1990.844	2058.606	2128.675	2201.129
40-44	1616.654	1675.890	1737.296	1800.952	1866.941
45-49	1219.103	1267.368	1317.544	1369.707	1423.935
50-54	1009.743	1053.268	1098.669	1146.027	1195.426
55-59	597.9146	626.3927	656.2273	687.4828	720.2270
60-64	700.8289	738.8433	778.9197	821.1699	865.7119
65-69	433.5885	462.3804	493.0842	525.8267	560.7436
70 and above	1578.125	1743.968	1927.238	2129.768	2353.582

Age Group	Year 2022	2023	2024	2025	2026
0-4	2384.045	2430.622	2478.109	2526.525	2575.886
5-9	2593.639	2649.191	2705.933	2763.890	2823.089
10-14	3054.991	3126.285	3199.242	3273.902	3350.304
15-19	2845.989	2917.999	2991.831	3067.532	3145.147
20-24	2807.985	2884.727	2963.565	3044.559	3127.766
25-29	2773.414	2855.061	2939.111	3025.635	3114.707
30-34	2407.273	2483.487	2562.114	2643.231	2726.915
35-39	2276.049	2353.519	2433.625	2516.459	2602.111
40-44	1935.347	2006.260	2079.771	2155.976	2234.972
45-49	1480.310	1538.916	1599.843	1663.182	1729.029
50-54	1246.955	1300.705	1356.772	1415.255	1476.259
55-59	754.5307	790.4683	828.1176	867.5601	908.8812
60-64	912.6699	962.1750	1014.365	1069.387	1127.392
65-69	597.9790	637.6870	680.0318	725.1884	773.3435
70 and above	2600.915	2874.241	3176.289	3510.080	3878.947

Age Group	Year 2027	2028	2029	2030	3031
0-4	2626.211	2677.519	2729.831	2783.164	2837.539
5-9	2883.555	2945.317	3008.401	3072.837	3138.652
10-14	3428.490	3508.500	3590.377	3674.165	3759.908
15-19	3224.727	3306.320	3389.978	3475.752	3563.697
20-24	3213.246	3301.063	3391.280	3483.963	3579.179
25-29	3206.401	3300.795	3397.967	3498.000	3600.978
30-34	2813.249	2902.317	2994.204	3089.000	3186.798
35-39	2690.679	2782.262	2876.962	2974.885	3076.141
40-44	2316.864	2401.756	2489.758	2580.985	2675.555
45-49	1797.483	1868.647	1942.628	2019.538	2099.494
50-54	1539.893	1606.270	1675.508	1747.731	1823.067
55-59	952.170	997.521	1045.032	1094.806	1146.951
60-64	1188.545	1253.014	1320.980	1392.632	1468.172
65-69	824.696	879.459	937.859	1000.136	1066.549
70 and above	4286.579	4737.048	5234.855	5784.977	6392.909

The rate at which the urban male population is growing is a concern to the government of Bangladesh. Bangladesh Government, therefore, requires policies to provide maximum welfare to its first growing urban population. That is the government needs strategic planning to provide good health care, accommodation, roads and highways, educational institutions, and jobs to satisfy the minimum need of the people of Bangladesh.

5. Conclusions

This study observed that the age pattern of urban male population of Bangladesh follows generalized negative exponential model for the census years 1981 and 1991. But the linearity is maintained for the data of the census year 2001. The smoothed series are then used to estimate the geometric growth rate for each age group of urban males following Malthusian law of population growth. It is observed that the older urban males are growing faster to double the population than the younger. The projected growth of urban males is an early warning to the government of Bangladesh to take the matter seriously to accommodate its urban citizens with maximum welfare. Although this paper has used the generalized exponential and linear models for first stage population prediction there are, however a number of models e.g. logistic, Gompertz, Makeham models can also be applied for such predictions. But these models performed poorly in this study. Therefore, these models were omitted from the analysis. The research can be further explored to investigate the district wise growth of urban male population of Bangladesh. Specifically, the heavily populated districts e.g. Chittagong, Rajshahi, and Khulna are in our next research agenda.

References

BBS (1984). Bangladesh population census 1981, National Series, Government of the People's Republic of Bangladesh, Dhaka.

BBS (1994). Bangladesh population census 1991, Vol. 1, National Series, Government of the People's Republic of Bangladesh, Dhaka.

BBS. (2003). Bangladesh population census 2001, National report, Government of the People's Republic of Bangladesh, Dhaka.

Islam, Md. Rafiqul (2003). Modeling of demographic parameters of Bangladesh-an empirical forecasting, *unpublished Ph.D. Thesis*, Rajshahi University.

Islam, Md. Rafiqul, Islam, Md. Nurul, Ali, Md. Ayub & Mostofa, Md. Golam. (2003). Construction of male life table from female widowed information of Bangladesh, *International Journal of Statistical Sciences*, Vol. 2, Dept. of Statistics, University of Rajshahi, Bangladesh, Page 69-82.

Islam, Md. Rafiqul. (2005). Construction of female life table from male widowed information of Bangladesh, *Pakistan Journal of Statistics*, Vol. 21(3),Page 275-284.

Islam, Md. Rafiqul and A.B.M. Rabiul Alam Beg. Modeling and Predicting Population of Dhaka District of Bangladesh, International J. of Mathematics and Computation, Vol. 4, No. S09, pp. 69-80, September-2009.

Stevens, J. (1996). *Applied multivariate statistics for the social sciences*, Third Edition, Lawrence Erlbaum Associates, Inc., Publishers, New Jersey.

Velleman, P. F. (1980). Definition and comparison of robust nonlinear data smoothing algorithms, *Journal of the American Statistical Association*, Volume 75. Number 371, 609-615.