Modeling and Forecasting of Some Fertility Parameters in Urban Area of Bangladesh

Md. Rafiqul Islam

Abstract. *The aim of the present study is to build some mathematical models and then to forecast some fertility parameters in urban area of Bangladesh. For this purpose, the secondary time series data on Crude Birth Rate (CBR), Total Fertility Rate (TFR), Gross Reproduction Rate (GRR) and Net Reproduction Rate (NRR) of various issues duly published by Bangladesh Bureau of Statistics (BBS) have been used in the present study. A few mathematical time trend models have been fitted to time series data of CBR, TFR, GRR and NRR. It is found that the CBR follows quadratic (i.e. parabolic) polynomial model while the TFR, GRR and NRR follow simple linear regression model. Model validation technique such as Cross-Validity Prediction Power (CVPP), ρ^2_{cv}, is applied to these models to verify how much these models are valid or not. It was found that all these models are more than 95%, 79%, 82%, and 72% stable respectively and their shrinkages are only 0.007397, 0.032133, 0.027916, and 0.0424229, respectively. These rates have been forecasted during 1999-2005 using these time trend models.*

Keywords: Crude Birth Rate (CBR), Total Fertility Rate (TFR), Gross Reproduction Rate (GRR), Net Reproduction Rate (NRR), Modeling, Cross Validity Prediction Power (CVPP) and F-test, Bangladesh.

1. INTRODUCTION

Bangladesh is located in southern Asia, borders the Bay of Bengal, between Burma and India, and occupies a total of 144,000 square kilometers. It is situated latitudinally between $21^0 5'$ and $26^0 4'$ North and

longitudinally between $88^0 5'$ and $92^0 5'$ East. The country has a tropical climate with cool, dry weather conditions from October to March. Summers are hot and humid. The monsoon season is characterized by heavy rainfall from June to October. It is frequently affected by natural calamities such as floods, cyclones, and droughts. Bangladesh is hilly in the southeast consisting mostly of flat alluvial plains. About 73 percent of the land is arable. The landscape has an extensive network of rivers that are very important for the socioeconomic condition of the nation. Among them are the Ganges-Padma, Brahmaputra-Jamuna, and the Megna.

Bangladesh is one of the most developing countries containing accelerated population growth in the world. The population of Bangladesh increased from 42 million in 1941 to 129.25 million in 2001 (BBS, 2001). It is to be mentioned here that Bangladesh is one of the most densely populated country containing 129.25 million citizens of which 65.84 million are males and 63.41 millions are females (BBS, 2001). It is the ninth most populous country in the world having population density 834 persons per square kilometer. The gross domestic product (GDP) per capita in Bangladesh is $331 in 2002 and the gross national product (GNP) per capita is $467 in 2005. The population is growing at a rate of 1.47 percent per year during the period 1991 to 2001 (Mitra and Associates, 2001).

It is to be noted that a number of works has been carried out for the estimation of demographic parameters of Bangladesh. Some have used raw data to estimate the parameters and some have used indirect techniques to do these. These estimates suffer too much for their qualities of their raw data. Different estimates for the parameters are obtained from different sources at various point of time. Thus, it is found that all parameters are functions of time. Hence, we can have a time trend model that can provide us the estimates with the changes of time. Therefore, an attempt has been made here to find out some time trend models to estimate the parameters CBR, TFR, GRR, and NRR in urban areas of Bangladesh.

The relevant review of literature of the present study is as follows:

Ali (1994) reviewed the relationship of total separation rates and separation rates due to death with their age variable and found a semi-log function of the type $\log_e y = \alpha + \beta x$. Islam and Ali (2004) found that age specific fertility rates (ASFRs) follow slightly modified bi-quadratic polynomial model where as forward and backward cumulative ASFRs follow quadratic and cubic polynomial model, respectively in the rural areas of Bangladesh. Age structure, Age Specific Death Rates (ASDRs), and the

number of persons surviving at an exact age x (l_x) in life table for male population of Bangladesh in 1991 follow modified negative exponential model, 4th and 3rd degree polynomial model, respectively (Islam, Islam, Ali and Mostofa, 2003). Islam (2004) showed that ASFRs and proportion of married women in the reproductive span follow 3rd degree polynomial model where as forward cumulative ASFRs follow quadratic polynomial model, respectively of Bangladesh. Islam (2004) observed that infant mortality rate (IMR) of Bangladesh followed semi-log linear time trend model. On the other hand, Crude Death Rate (CDR) and Life Expectancy at birth for male and female follow simple linear regression time trend (E_0) models.

The specific objectives of this study are in the following:

i) to fit appropriate mathematical time trend models to CBR, TFR, GRR and NRR in urban areas of Bangladesh, and
ii) to forecast these fertility parameters in urban areas of Bangladesh up to 2005.

2. METHODOLOGY

2.1 Sources of Data

The secondary time series data of CBR, TFR, GRR, and NRR of Bangladesh were taken from the Statistical Year Book of Bangladesh (BBS, 1980, 1986, 1999 and 2001). Hence, these rates have been used to fit some mathematical time trend models and then to forecast these parameters of Bangladesh up to 2005 using these time trend models. These rates are presented in Table 1.

3. METHODOLOGY

3.1 Model Fitting

i) The CBR for different years have been plotted in graph paper. It is seen that there are some sort of distortions that is unexpected. Before going to use this data, an adjustment has been made using the Package Minitab Release 12.1 by the latest smoothing method named "4253H, twice" (Velleman, 1980). Thus, smoothed data has been used for time trend model for CBR. From the scattered plot of CBR by years (Figure 1) it seems that

CBR follows polynomial model. In this case, an *nth* degree polynomial model is treated and the model of the nth degree polynomial is

$$y = a_0 + \sum_{i=1}^{n} a_i x^i + u$$ (Montgomery and Peck, 1982).

where, x is years; y is CBR; a_0 is the constant; a_i is the coefficient of x^i (i =1, 2, 3,..., n) and u is the error term of the model. Here we have to choose a suitable n for which the error sum of square is minimum.

ii) The scattered plot of TFR by years is shown in Figure 2. It is observed that TFR is fitted by using a simple linear regression model. The form of the model is

$$y = a_0 + a_1 x + u$$

where, x represents the years; y represents TFR; a_0, a_1 are parameters and u is the disturbance term of the model.

iii) The dotted plot of years and GRR is shown in Figure 3. It appears that GRR follows simple linear relation. Therefore, a simple linear regression model is fitted and the form of the model is found to be:

$$y = a_0 + a_1 x + u$$

where, x represents the years; y represents GRR; a_0, a_1 are parameters and u is the disturbance term of the model.

iv) Again, from the scattered plot of NRR by years (Figure 4), it is observed that NRR can also be fitted by simple linear regression model. Therefore, a simple linear regression model can be fitted and the form of the model is found to be:

$$y = a_0 + a_1 x + u$$

where, x represents the years; y represents NRR; a_0, a_1 are parameters and u is the disturbance term of the model.

It should be mentioned here that all these time trend models were estimated using the econometric software, E-Views.

3.2 F-test

To find out the measures of overall significance level of the fitted time trend model as well as the significance of R^2, the F-test is employed here. The F-test is given as follows:

$$F = \frac{R^2/(p-1)}{(1-R^2)/(n-p)}$$ with (p-1, n-p) degrees of freedom (d.f.);

where p = the number of parameters to be estimated, n is the number of cases and R^2 is the coefficient of determination in the model (Gujarati, 1998).

3.3 Model Validation

To check, how much these models are stable over the population, the Cross Validity Prediction Power (CVPP), ρ^2_{cv}, is applied. Here,

$$\rho^2_{cv} = 1 - \frac{(n-1)(n-2)(n+1)}{n(n-k-1)(n-k-2)}(1-R^2);$$

where, n is the number of cases, k is the number of regressors in the model and the cross-validated R is the correlation between observed and predicted values of the dependent variable. Using the above statistics, it is concluded that if the prediction equation is applied to many other samples from the same population, then (ρ^2_{cv} X100)% of the variance on the predicted variable would be explained by the model. The shrinkage of the model is the absolute value of the difference of ρ^2_{cv} and R^2. Moreover, the stability of R^2 of the model is equal to (1- shrinkage) (Stevens, 1996).

4. RESULTS AND DISCUSSION

The polynomial model is considered to fit the model for CBR on time. It is seen that a second-degree polynomial is fitted and the fitted model is $y=27.21793-0.031904x^2$. The coefficient of determination of the model is $R^2=0.959844$ and its $\rho^2_{cv}=0.952447$.s

Fitting a regression of GRR on time it is found that the regression line is

$$y = 1.762281 - 0.033649x$$

giving $R^2 = 0.848457$ and $\rho_{cv}^2 = 0.820541$.

Lastly, a regression line of NRR on time gives the relation

$$y = 1.441930 - 0.021509x$$

with $R^2 = 0.769671$ and $\rho_{cv}^2 = 0.727242$.

It should be mentioned here that the usual models, i.e. Gompertz model or Makeham model, exponential model, log-linear model, and logistic model were also applied but seems to be worse fitted with respect to their shrinkages. Therefore, the results of these models were not shown here.

Table 1
CBR, TFR, GRR, AND NRR IN URBAN AREA OF BANGLADESH DURING 1980-1998

Year	CBR	Smoothed CBR	TFR	GRR	NRR
1980	29.2	26.89	3.87	1.85	1.60
1981	24.8	26.54	3.20	1.57	1.32
1982	22.9	26.26	3.01	1.48	1.26
1983	27.1	25.99	3.45	1.66	1.41
1984	25.0	25.76	3.10	1.50	1.30
1985	28.0	25.67	3.52	1.72	1.41
1986	25.9	25.55	3.26	1.57	1.28
1987	24.8	25.26	3.05	1.48	1.17
1988	24.9	24.95	2.9	1.53	1.22
1989	24.4	24.69	2.89	1.41	1.20
1990	24.6	24.39	2.95	1.42	1.20
1991	23.9	23.91	2.87	1.38	1.19
1992	23.7	22.94	2.88	1.37	1.18
1993	21.0	21.53	2.62	1.32	1.15
1994	20.2	20.29	2.58	1.25	1.12
1995	19.4	19.35	2.50	1.21	1.09
1996	19	18.13	2.48	1.19	1.08
1997	16.2	16.26	2.28	1.1	1.07
1998	14	13.98	2.24	1.08	1.06

Source: Statistical Year Book-1980, 1986, 1998, 1999.

Figure 1
ACTUAL, FITTED AND RESIDUAL OF CBR IN URBAN AREA OF BANGLADESH

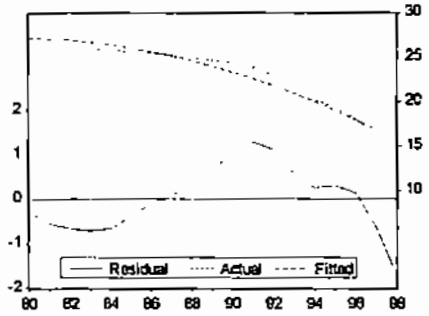


Figure 2
ACTUAL, FITTED AND RESIDUAL OF TFR IN URBAN AREA OF BANGLADESH OF BANGLADESH

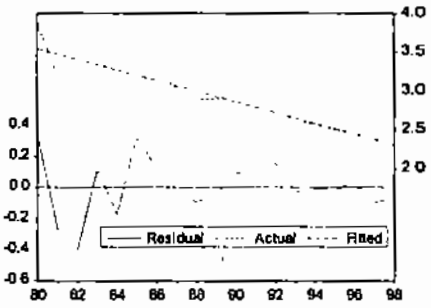


Figure 3
ACTUAL, FITTED AND RESIDUAL OF GRR IN URBAN AREA OF BANGLADESH

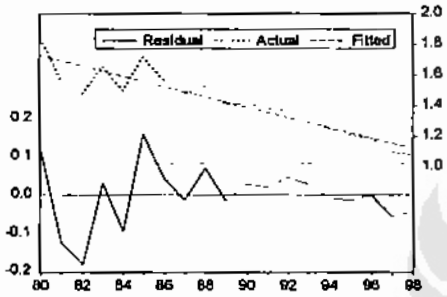
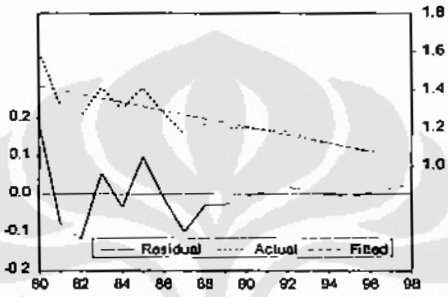


Figure 4
ACTUAL, FITTED AND RESIDUAL OF NRR IN URBAN AREA OF BANGLADESH



The estimated CVPP (ρ^2_{cv}) of the fitted models with corresponding R^2 have been summarized in Table 2. From this table, it is seen that all the fitted models in equation (i) to equation (iv) possess the explanatory powers exceeding 0.959844, 0.825563, 0.848457, and 0.769671 respectively and their shrinkages are 0.007397, 0.032133, 0.027916, and 0.0424229 respectively. These imply that all these models are more than 95%, 79%, 82% and 72% stable respectively.

The results on model fitting of these time trend models have been presented in Table 3. From this table, it is observed that all the parameters of the fitted models are highly significant. The stability of R^2 of these models are more than 99%, 96%, 97% and 95%, respectively. These imply that the fitted models are well fitted to the data.

Table 2

ESTIMATED CROSS VALIDITY PREDICTION POWER, ρ^2_{cv}, OF THE PREDICTED EQUATIONS OF CBR, TFR, GRR AND NRR IN URBAN AREA OF BANGLADESH

Models	n	k	R^2	ρ^2_{cv}	Shrinkage
Equation 1	19	1	0.959844	0.952447	0.007397
Equation 2	19	1	0.825563	0.793430	0.032133
Equation 3	19	1	0.848457	0.820541	0.027916
Equation 4	19	1	0.769671	0.727242	0.0424229

Table 3
RESULTS ON MODEL FITTING

Models	Percentage of Variance Explained	Parameters	Significant Probability (p)
Model 1	95.9844		
		a_0	0.0000
		a_2	0.0000
Model 2	82.5563		
		a_0	0.0000
		a_1	0.0000
Model 3	84.8457		
		a_0	0.0000
		a_1	0.0000
Model 4	76.9671		
		a_0	0.0000
		a_1	0.0000

The calculated values of F-test of the models (1)- (4) are 406.3489, 80.45639, 95.17938, and 56.80747 with (1, 17) d.f. respectively but each of which corresponding tabulated is 8.18 at 1% level of significance. Therefore, from these statistics it is seen that these models and their corresponding R^2 are highly statistically significant.

To study the residual analysis of these models, the residuals have been plotted in the graphs that are shown in Figures 1 to 4. The forecasted values of these models have been presented in Table 4 and in Figures 5 to 8.

Table 4
THE FORECASTED VALUES OF CBR, TFR, GRR AND NRR IN URBAN AREA OF BANGLADESH DURING 1999-2005

Years	CBR	TFR	GRR	NRR
1999	14.45626	2.243158	1.089298	1.011754
2000	13.14819	2.174579	1.055649	0.990246
2001	11.77631	2.106000	1.022000	0.968737
2002	10.34062	2.037421	0.988351	0.947228
2003	8.841122	1.968842	0.954702	0.925719
2004	7.277817	1.900263	0.921053	0.904211
2005	5.650703	1.831684	0.887404	0.880702

Figure 5
FORECASTED VALUES OF CBR IN URBAN AREA OF BANGLADESH

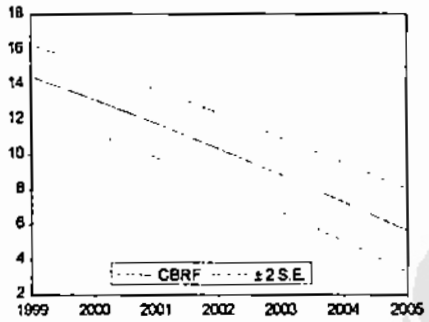


Figure 6
FORECASTED VALUES OF TFR IN URBAN AREA OF BANGLADESH

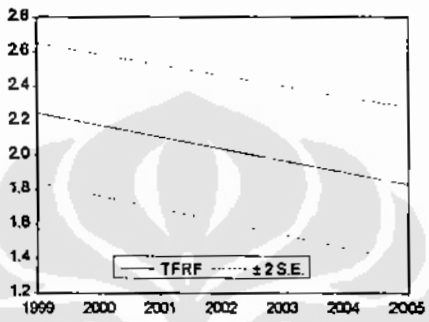


Figure 7
FORECASTED VALUES OF GRR IN URBAN AREA OF BANGLADESH

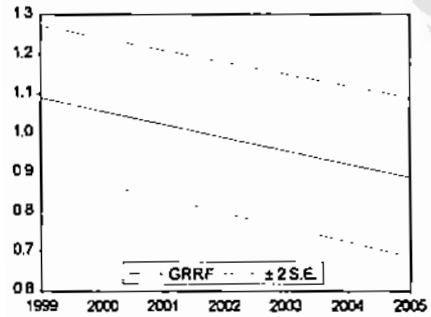
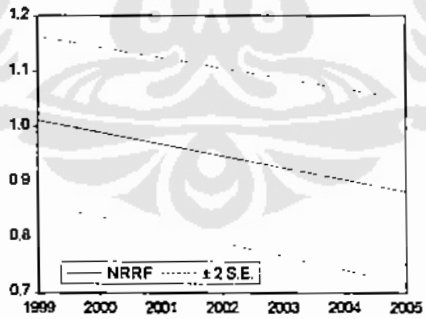


Figure 8
FORECASTED VALUES OF NRR IN URBAN AREA OF BANGLADESH



5. CONCLUSION

In the present study on the one hand, CBR in urban area of Bangladesh follows 2nd degree (i. e. quadratic) polynomial model but in which the term containing x is absent because of the insignificant of the parameter. On the other hand, TFR, GRR and NRR in urban area of Bangladesh follow simple linear regression model. These time trend models have been provided the forecasted values of these rates during 1999-2005.

It is hoped that this study may be useful to the researchers and academicians as well as the government planners to provide sound basis for comprehending any future plan of action for the socio-economic development and health care programs in the country.

References

Ali, M. Korban. 1994. "Modeling of Labour Force Dynamics in Bangladesh: An Evidence from 1981 Census", *The Rajshahi University Studies Part-B*, Vol. 22, Page-259-266.

Bangladesh Bureau of Statistic (BBS). 1980. *Statistical Year Book of Bangladesh 1980.*

------. 1986. *Statistical Year Book of Bangladesh 1986.*

------. 1999. *Statistical Year Book of Bangladesh 1998.*

------. 2001. *Statistical Year Book of Bangladesh 1999.*

------. 2001. *Population Census 2001: Preliminary Report.* Government of the People's Republic of Bangladesh, Dhaka.

Gujarati, Damodar N. 1998. *Basic Econometrics*, Third Edition, New York: McGraw Hill,

Islam, Md. Rafiqul, Md. Nurul Islam, Md. Ayub Ali and Md. Golam Mostofa. 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 and M. Korban Ali. 2004. "Mathematical Modeling of Age Specific Fertility Rates and Study the Reproductivity in the Rural Area of

Bangladesh During 1980-1998", *Pakistan Journal of Statistics*, Vol. 20(3), page-381-394.

Islam, Md. Rafiqul. 2004. "Indirect Estimation of Fertility Parameters of Bangladesh", *Journal of Indian Anthrop. Soc.* Vol.39, No. 2, Page-195-202.

Islam, Md. Rafiqul. 2005. "Empirical Forecasting of Mortality Parameters of Bangladesh", *South Asian Anthropologist, India*, Vol. 5, No.1, page-81-87.

Mitra, S. N. and Associates. 2001. *Bangladesh Demographic and Health Survey, 1999-2000.* National Institute of Population Research and Training (NIPORT), Dhaka, Bangladesh.

Montgomery, Douglas C. and Elizabeth A. Peck. 1982. *Introduction to Linear Regression Analysis.* New York: John Wiley and Sons.

Stevens, J. 1996. *Applied Multivariate Statistics for the Social Sciences*, Third Edition, New Jersey: Lawrence Erlbaum Associates, Inc., Publishers.

Velleman, P. F. 1980. "Definition and Comparison of Robust Nonlinear Data Smoothing Algorithms", *Journal of the American Statistical Association*, Volume 75. Number 371, Page 609-615.

Md. Rafiqul Islam, Associate Professor, Dept. of Population Science and Human Resource Development, Rajshahi University, Bangladesh.
E-mail: rafique_pops@yahoo.com