

Modelling Aging Population in Sri Lanka

Amalie Jayatilake^{1, a}, Sarath Peiris^{2, b}

¹ Department of Business Management,
Faculty of Management Studies,
Sabaragamuwa University of Sri Lanka, Sri Lanka

²Department of Mathematics,
Faculty of Engineering,
University of Moratuwa,
Sri Lanka

^aamaliechathurika@gmail.com, ^bsarathp@uom.lk

Abstract

Population ageing is a universal phenomenon and it is expected to be among the most prominent global demographic trends of the 21st century. In Sri Lanka, there was a rising trend of aging population throughout the past years and has recorded the highest number of old people in South Asia. However, no sound statistical or mathematical models exist to project aging population in Sri Lanka. Using the population of people aged 60 years and above in Sri Lanka during 1950-2016, three types of statistical models: (i) ARIMA (0, 2, 1), (ii) exponential trend model, and (iii) double exponential smoothing model were developed. The models were compared using various statistical indicators and some statistical diagnostics tests. The comparison was done for both training set as well as a validation set. Among these models, the double exponential smoothing model was found to be the best. According to the forecast derived from this model, it was found that the increasing trend of aging population in the country will continue in the future and there will be approximately 2,936,000 old people in Sri Lanka in 2020. The information forecast in this study is beneficial for planners and decision makers in the government sector and other relevant organizations to cater to the needs of the increasing aged population.

Keywords: Aging Population, ARIMA, Demographic, Exponential Smoothing

Introduction

Aging population (60+ years) is a shift in the distribution of a country's population towards older people. This is usually reflected in an increase in the population's mean and median ages, a decline in the proportion of the population composed of children and a rise in the proportion of the population that is older (Gavrilov and Heuveline, 2003). Increase of aging population has become a universal phenomenon, which has received attention throughout the world. United Nations has mentioned that Sri Lanka has the fastest aging population among South Asia and consequently various authors such as Abeyratne et al, (2014) have highlighted the importance of forecasting aging population in spite of lack of consistent annual data.

De Silva (1994) notes that with rapid decrease of fertility which occurred early and the increasing trend of international migration within the working age group has influenced to accelerate the onset of the proportional increase of the old. Aberathna et al. (2014) point out that the factors such as fertility control policies, vast education of reproductive practices, and increases in the marital age limit of females has contributed to the decline of fertility reducing the younger population.

Since forecasting of aging population is also a part of the population projection, there is no doubts that these projections will be helpful for government to plan necessary actions for better care of elderly population. Li et al (2009) also claimed that it is necessary to predict the level of population aging, to inform policy debates about the likely effect of different intervention strategies.

The two common methods used to project national populations are: (i) mathematical model and (ii) cohort component method. Even though population aging related topics have been discussed by many authors in different countries (Manike, 2014a; Manike, 2014b; Siddhisena, 2005; Prasannath, 2014) in Sri Lanka only a single study (De Silva, 2007) has been carried out for the projection of future aging population using cohort component method. It was found that this method has various drawbacks from both mathematical and statistical points of view. In view of the above, this study was initiated to develop a statistical model for aging population for short-term prediction.

Data and Methodology

The study uses the secondary data from annual aging population statistics from 1950 to 2016. Data from 1950 to 1991 were acquired from statistical abstracts published by the Department of Census and Statistics (DCS) Sri Lanka (2015) and the corresponding data for the period of 1992 to 2016 were acquired from the Registrar General's Department (RGD) in Sri Lanka (2016). Data consisted of both the censured and estimated aging population data as the DCS has carried out only six housing and population census in 1953, 1963, 1971, 1981, 2001 and 2012 during the entire period from 1950-2016.

The statistical models techniques used are ARIMA models, growth model and double exponential smoothing models. All models were trained using data from 1950 to 2012 (63 years) and validated for the period of 2013-2016. EVIEWS and Minitab statistics software were used for data analysis.

Results and Discussion

ARIMA Model

The annual aging population can be considered as a time series at equal intervals and thus ARIMA models were developed. Augmented Dickey-Fuller (ADF) test confirmed (Table 1) that second differenced series is not significantly deviated from stationary.

Table 1: Results of ADF test of ageing series

Series	Augmented Dickey-Fuller test statistic	
	t- statistics	Probability
Original series	-0.797412	0.9601
First differenced series	-5.76380	0.6701
Second differenced series	-9.99196	0.0000

In order to decide possible ARIMA models to the stationary series, the observed patterns of ACF and PACF of stationary series were compared with the theoretical ACF and PACF of AR(1), MA(1) and ARMA(1,1). Three models: ARIMA (1, 2, 0), ARIMA(0,2,1) and ARIMA(1,2,1) were considered as possible parsimonious models and tested for the significance of the parameters of the models (Table 2).

Table 2: Validation of the parameters of three parsimonious models

Model	Variable	Coefficient	t-Statistic	Prob.
ARIMA (1,2,0)	C	6.531	0.656	0.5141
	AR(1)	-0.493	-3.303	0.0016
ARIMA (0,2,1)	C	1.575	1.389	0.1698
	MA(1)	-0.945	-30.690	0.000
ARIMA (1,2,1)	C	1.550	1.396	0.167
	AR(1)	-0.068	-0.365	0.716
	MA(1)	-0.944	-29.538	0.000

According to the results in Table 2, the model ARIMA (1, 2, 1) was rejected as the AR parameter is not significant. It was then confirmed that errors of both models: ARIMA (1,2,0) and ARIMA(0,2,1) are random using Box-Pierce statistics at different lags. Furthermore, it was found that there is no serial correlation in ARIMA (0,2,1) since the Breusch-Godfrey Serial Correlation Lagrange Multiplier (LM) test statistic was not significant at 5% level. Also it has the lowest Akaike Information Criteria (AIC). Thus, ARIMA (0, 2, 1) model was identified as the best fitted ARIMA model for aging data. The percentage errors for the validation data was also found to be below 10%.

Growth Model

Growth model in the form of: $y_t = a * b^t$ was fitted and it was found that both the model and parameters are significant. The fitted model is $Y = 350.833 * (1.029)^t$ ($R^2 = 91.7\%$, $AdjR^2 = 91.5\%$). The residuals of the fitted growth model were found to be white noise. The percentage errors through the developed growth model for the

validation dataset were also found to be below 10% for all four years. Thus, this model can also be used to predict future ageing population.

Double Exponential Smoothing Model

According to Siregar (2016), double exponential smoothing method smoothed trend component separately using the two parameters namely α and β . He further mentioned that double exponential smoothing uses a dynamic trend component that works well for the series having shift in the trend. Smoothing constants are the key to success of exponential smoothing. Therefore, prior to applying double exponential smoothing it is necessary to decide those two smoothing constants. Thus, in order to select the most appropriate constants, a simulation study through trial and error method was carried out by changing α starting from 0.063613 to 0.963613 with an increment of 0.1 and that of β starting from 0.944218 to 0.044218 with an increment of 0.1. The best combination of α and β selected based on the minimum mean absolute percentage error (MAPE).

Having decided the two smoothing constants, the initial starting point values and the initial forecast were considered as very important. However, in this case, the default option in Minitab software was used. According to calculations, it was found that the initial value of level is 409.03 and initial value of trend is 19.24. Therefore, the initial forecast for the first observation of the series was taken as $409.03 + 19.24 = 417.27$. The double exponential smoothing plot obtained from Minitab is shown in Figure 1.

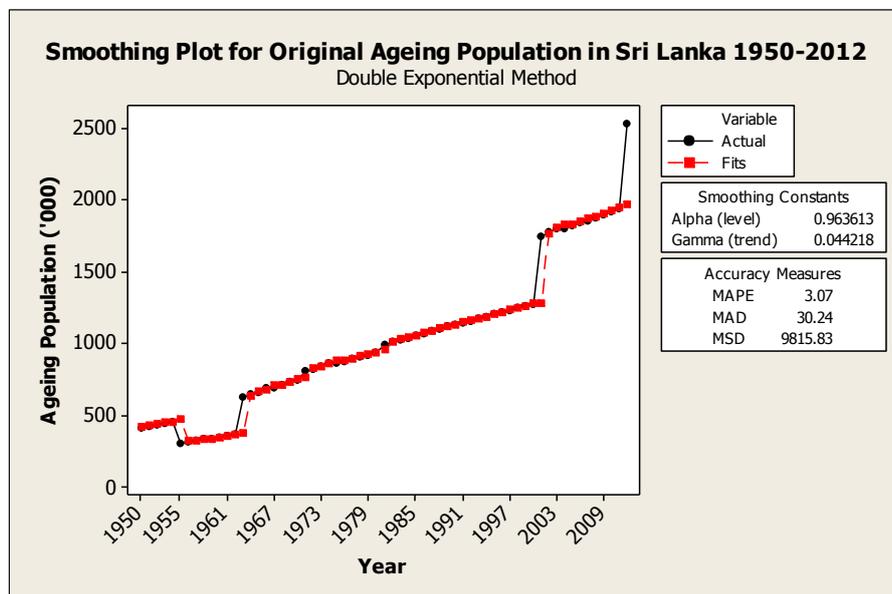


Figure 1: Double exponential smoothing curve for the ageing population in Sri Lanka

It was found that the percentage errors through the developed double exponential smoothing model for the validation dataset is very low (<4%) compared with other two models and the output is given in the Table 3.

Table 3: Percentage errors for the validation dataset

Year	2013	2014	2015	2016
Estimated ageing population (in '000)	2555	2609	2664	2718
Actual ageing population (in '000)	2548	2571	2593	2623
% Error	-0.3%	-1.5%	-2.7%	-3.6%

To confirm the suitability of the double exponential smoothing technique to forecast the future aging population in Sri Lanka, validation measurement such as MAPE was considered for both the training set and the validation set separately and it was found that those statistics are low compared with other two models.

Comparison of Three Models

The percentage errors for the validation dataset obtained using the three models were compared simultaneously (Table 4).

Table 4: Comparison of the percentage errors of three models for the validation set

Model	Ageing Population ('000)	2013	2014	2015	2016
ARIMA(0,2,1)	Estimated	2599	2678	2758	2840
	Actual	2548	2571	2593	2623
	% Error	-2.0%	-4.2%	-6.7%	-8.8%
Growth Model	Estimated	2249	2315	2383	2454
	Actual	2548	2571	2593	2623
	% Error	11.7%	9.9%	8.1%	6.5%
Double Exponential Smoothing	Estimated	2555	2609	2664	2718
	Actual	2548	2571	2593	2623
	% Error	-0.3%	-1.5%	-2.7%	-3.6%

As per results in Table 4, it is clear that the percentage errors of the ARIMA (0,2,1) and double exponential smoothing models are increasing with respect to time except the growth model. It can be seen that the percentage errors for all years are the highest for growth model and are lowest for double exponential smoothing method. Thus in respect to the percentage errors, it can be concluded that double exponential smoothing model is better than the other two. Similar results were obtained for the training sets of the three models. Siregar et al. (2016) mentioned that smaller the accuracy of measures the better the forecast. The three accuracy measures namely MAPE, MAD and MSD obtained for were compared simultaneously for the training set as well as for the validation set (Table 5).

Table 5: Comparison of the accuracy of the models through training set and validation set

Type of the dataset	Accuracy measurement	ARIMA(0,2,1) model	Exponential trend model	Double exponential smoothing model
Training dataset	MAPE	4.5%	13.6%	3.1%
	MAD	43.49	104.66	30.94
	MSD	9624	14677	10135
Validation dataset	MAPE	5.2%	9.1%	2%
	MAD	135	234	53
	MSD	22091	57049	3855

According to the Table 5, the two measurements MAPE and the MAD derived from the training set of the double exponential smoothing model are smaller, comparing with the same measurements derived from the two other models. Simultaneously the smaller MSD value of the training set belongs to the ARIMA (0, 2, 1) model. When considered about the validation set, all the three accuracy measurements (MAPE, MAD, and MSD) derived from double exponential smoothing model are comparatively small. Based on these reasons, it can be mentioned that the double exponential smoothing model is the best fitted model compared with the other two models. It should be highlighted that MAPE are less than 4% for the double exponential model. In fact, Siregar (2016) mentioned that if MAPE is less than 10% the fitted model is said to be excellent. Therefore, double exponential model is also recommended as the most suitable model.

Short-Term Forecasting

Using the developed double exponential smoothing model forecasting was carried out for the aging population in Sri Lanka for the years 2017, 2018, 2019 and 2020 and the results are shown in Table 6.

Table 6: Forecasted Aging Population in Sri Lanka from 2017-2020

Year	Forecasted Ageing Population [in '000]
2017	2,772
2018	2,827
2019	2,881
2020	2,936

Conclusions and Recommendations

The Double Exponential Smoothing model gave better forecast values for aging population in Sri Lanka during 1950 to 2016. The forecasted aging populations for 2017 to 2020 are 2.772, 2.827, 2.881 and 2.936 million respectively. Until a further model is developed, DES model can be used to forecast aging population in Sri Lanka. The forecasted aging populations up to 2020 would be very useful for policy makers to implement various projects to care of elderly persons in Sri Lanka.

The government should cater to the needs of the elderly population by rethinking and developing necessary welfare facilities. A better method needs to be developed to estimate aging population for non-census years due to the drawbacks found with relate to those estimations. Furthermore, it is suggested to find the possibility of neural network models and multivariate time series techniques (Vector Autoregression or Bayesian Autoregression models) by incorporating the factors such as mortality, fertility, births, deaths and migrations.

References

- Abeyrathna, W., Alles, L., Wickremasinghe, W. N., & Hewapathirana, I. (2014). 'Modeling and forecasting mortality in Sri Lanka. Sri Lanka'. *Journal of Applied Statistics*, Vol. 15, No 3, pp 141-170.
- Department of Census and Statistics Sri Lanka (2015). *Census of Population and Housing 2012*, ISBN:978-955-577-940-1.
- De Silva, W. I. (1994). 'How serious is ageing in Sri Lanka and what can be done about it?'. *Asia-Pacific Population Journal*, Vol. 9, No. 1, pp 19-36.
- De Silva, W. I. (2007). *A population projection of Sri Lanka for the new millennium 2001-2101 trends and implications*. Institute for health policy, Colombo.
- Gavrilov, L. A., & Heuveline, P. (2003). *The aging of population*. Manuscript of article in the encyclopedia of population, Retrieved from <http://www.galegroup.com/servlet/ItemDetailServlet?region=9&imprint=000&titleCode=M333&type=4&id=174029>
- Li, O., Reuser, M., Kraus, C., & Alho, J. (2009). 'Aging of a giant: a stochastic population forecast for China 2006-2060'. *Journal of Population Research*, Vol. 26, pp 21-50.
- Menike, H. R. A. (2014a). 'Age structure transition and demographic bonus in Sri Lanka. *International Journal of Social Sciences and Entrepreneurship*', Vol 1, No. 12, pp 421-431.
- Menike, H. R. A. (2014b). 'Important features of the elderly population in Sri Lanka'. *Research Process*, Vol 2, No. 2, pp 29-38.
- Prasannath, V. (2014). *Trends in aging population and its consequences in Sri Lanka*. Proceedings of second Annual Research Conference, Faculty of Commerce and Management, South Eastern University of Sri Lanka, 100-104.
- Registrar's General Department of Sri Lanka (2016). *Age and sex specific midyear population sizes of Sri Lanka - 1992-2016*.
- Siddhisena, K. A. P. (2005). *Social- economic implications of aging in Sri Lanka: an overview*, Oxford Institute of Aging Working Papers, 105. Retrieved from <http://www.ageing.ox.ac.uk/download/98>
- Siregar, B., Butar, I. A. B., Rahmat, R. F., Andayani, U., & Fahmi, F. (2016). 'Comparison of exponential smoothing methods in forecasting palm oil real production'. *Journal of Physics*, 801, pp 20-28.