Modeling and Forecasting Mortality in Sri Lanka

Article *in* Sri Lankan Journal of Applied Statistics · December 2014
DOI: 10.4038/sljastats.v15i3.7794

Modeling and Forecasting Mortality in Sri Lanka

Wasana Aberathna[1], Lakshman Alles[2], W. N.Wickremasinghe[1*] and Isuru Hewapathirana[1]

[1]Department of Statistics, University of Colombo, Colombo 3, Sri Lanka
[2]School of Economics and Finance, Curtin University, Australia
[*]Corresponding Author: wnw@stat.cmb.ac.lk

Received: 28, January 2014 / Revised: 15, August 2014 / Accepted: 27, August 2014

ABSTRACT

The purpose of this study is to develop sex-specific mortality estimation models using historical mortality data for Sri Lanka, based on the statistical time series techniques attributed to Lee and Carter (1992). Historical mortality data was analyzed in the light of significant historical episodes. Several alternative univariate time series models were examined for modeling males and females, as well as a bivariate vector autoregressive (VAR) model. The VAR model when fitted to the first differenced series performed better than the univariate models and hence used for forecasting purposes. From the estimated VAR model, mortality forecasts were generated for the period up to 2030 and life tables were generated for the selected periods of 2006-2008.

Keywords: Ageing, Mortality forecasting, life tables, Sri Lanka

1. Introduction

1.1 Background

The life expectancy at birth in Sri Lanka as at 2007 is 70.3 years for males and 77.9 years for females (Central Bank of Sri Lanka, 2009) and this has been increasing over time. As a result of increasing life expectancy and decreasing mortality, the population in Sri Lanka and the world in general, is ageing. An ageing population is commonly defined as one with an increasing proportion of the population in the elderly groups. The speed of ageing is relatively higher in Asia (Borsch, 2009) and in fact, Sri Lanka has the fastest-ageing population in South Asia (Abeykoon, 1996; Sri Lanka Aging Survey, 2006). Elderly Sri Lankans (aged ≥ 60 years) constituted 9.1 percent of the population in the 2001 census, an increase of 2.5 percentage points from the previous census of 1981 (Department of Census and Statistics of Sri Lanka). The latest estimates for 2006 put this figure at 9.4 percent (Ostbye *et. al*., 2010).

An ageing population results in an increase in the number of pensioners and retirees. The traditional custom in Sri Lanka is for the elderly to be cared for by their children and to be supported financially by them, at least partially. Due to factors such as increased migration from rural areas to urban areas, and migration overseas for employment, the family support base for the elderly in Sri Lanka is increasingly diminishing (De Silva, 2005). Gamaniratne (2007) reports that the old age dependency ratio[1] which stood at 10% in 2000 is projected to increase to 18% in 2020 and 27% in 2040. These trends highlight the need for the elderly to be increasingly self-supportive in the future. Gamaniratne (2007) further reports that Sri Lanka's social security system covers only 25% of the working age population. The expected growth in the elderly population and the need to extend the coverage of social security to a larger proportion of the population underscore the need to strengthen social security programs in the future. In such a context the government of Sri Lanka would need to deal with the issues of providing retirement benefits to an increasing population of retirees, the fiscal affordability of doing so, and optimizing the design of pension systems for achieving these policy objectives.

There are three components that affect population demographics; fertility, mortality and migration. Sri Lanka, like many other developing countries in the Asia Pacific region, has entered the third stage of demographic transition, the phase of declining fertility and mortality (De Silva, 1994). Fertility control policies, increasing education of reproductive practices, increases in the marital age of females with education have contributed to the decline in fertility. Effective application of DDT in national efforts to eradicate malaria, improvements in the health care system, improvements in agricultural production and subsidized distribution of food items and the expansion of free educational services, all have directly or indirectly contributed to mortality decline in Sri Lanka. The 30-year war in Sri Lanka also took a heavy toll on the lives of younger age groups. With the end of the war in 2009, it is expected that the demographic patterns will change in the future. The population size in Sri Lanka may also be impacted as migratory patterns change after the end of the war.

[1] old age dependency ratio is the ratio of people aged 60 and over to those between aged 15 to 64

The three components affecting population demographics, fertility, mortality, and migration, need to be studied and analyzed for formulating policies and mechanisms for enhancing the financial security of retirees in Sri Lanka. Mortality projections are an essential pre-requisite for any exercise relating to government pension policy formulation and pension modeling.

Mortality projections are also important for public and private sector financial institutions such as insurance companies, pension plans and provident funds that offer long term financial products such as retirement pensions, annuities and life and health insurance policies. The financial viability of these institutions are crucially dependent on properly managing their balance sheet assets and liabilities, and the risks related to those assets and liabilities. The risk of mortality changes has a significant impact in assessing and modeling the liabilities relating to such long term financial products. Significant improvements in longevity observed in the recent past indicate that modeling and projecting longevity requires a stochastic modeling approach, rather than the traditionally used deterministic approaches.

Population projections for Sri Lanka have been compiled by De Siva (2007). These projections are based on a demographic cohort component approach, where age specific five year survivorship ratios and population data for mid-year five year age groups are utilized to project future population in five year age brackets.

1.2 Need for Modeling Mortality in Sri Lanka

As far as the authors are aware, no studies have so far utilized such statistical modeling techniques based on historical data for projecting mortality in Sri Lanka. The Lee-Carter (1992) statistical modeling technique differs from the approach of forecasting age-specific rates independently. It exploits the high degree of inter-temporal correlation structure across the ages, by making all death rates functions of the same time varying parameter.

The objective in this paper is to apply statistical modeling techniques similar to those applied by Lee and Carter (1992) to model mortality in Sri Lanka, and utilize the models to project future sex-specific and age-specific mortality in Sri Lanka, and to provide probabilistic confidence regions for the projections. Life tables are also constructed and interpreted using 2006, 2007 and 2008 data. To achieve this objective, first the historical information on the population size and mortality rates is gathered and then the Lee-Carter (1992) modeling technique is applied to Sri Lankan data, in order to develop an appropriate time series

mortality estimation model. The efficacy of the model is evaluated with appropriate diagnostics.
Thus, the results of this study will fill a serious information deficiency currently faced by financial institutions and government policy makers in Sri Lanka.

2. Methodology

2.1 Data Collection

In this study, the data collection process was the most difficult part. To develop life tables and mortality models, age-specific mortality rates must be available. Unfortunately the data available (reported) in Sri Lanka was incomplete, and the data gaps needed to be filled. This study is based on data collected from several sources for the period 1950 – 2010. This consists of age-specific and sex-specific midyear population sizes for 1950 – 2010 collected from statistical abstracts published by the Department of Census and Statistics, and age-specific, sex-specific number of deaths registered from 1950 to 2008 at the Registrar General's office.
In 1990, the number of deaths was not reported. Therefore it was estimated as the average of the number of deaths in 1989 and 1991. In some cases population sizes and number of deaths available in various age groups did not correspond to our age groups. Therefore, the data needed for our age groups had to be extrapolated from available records. To build life tables, 2006, 2007 and 2008 sex-specific, age specific death rates were used. These were calculated from death rates using age-specific midyear population sizes and age-specific number of deaths. As the death rates for the age groups 0 - 1 and 1 – 4 were not consistently available they were calculated as follows:
Number of deaths in 0 - 1 age group was calculated for each year using number of deaths registered in 3-month intervals (which was available) below 1 year. Midyear population size was available for the 0 – 4 age group, and 0 - 1 age group as a proportion of 0 – 4 age group was obtained from the census of population and housing – 2001. Therefore, the death rate for 0 – 1 age group was estimated as the ratio of number of deaths in that interval to the midyear population size (calculated as a proportion) for that interval. Number of deaths in 1 – 4 age group was calculated as the total number of deaths reported for ages 1, 2, 3, and 4. Then the midyear population size for 1 – 4 age group was obtained as a proportion of the midyear population size for 0 – 4 interval from the population and housing census – 2001. Then, the death rates for 1 – 4 were estimated using the corresponding ratio as before.

2.2 The Lee–Carter (LC) Approach

Lee-Carter model (Lee and Carter, 1992) has been considered a gold standard for mortality forecasting in recent times (Lee and Miller, 2001; Li *et. al.*, 2004; Li and Lee, 2005; Li and Chan, 2007; Li *et. al.*, 2011). The original work by Lee and Carter (1992) combines a rich yet parsimonious demographic model with statistical time series methods to forecast age-specific mortality in USA for the period 1990 to 2065. They have used historical age-specific death rates for the entire US population for the period 1933-1987 for the purpose of model fitting. This method is extrapolative and makes no effort to incorporate knowledge about medical, behavioral, or social influences on mortality change. Its virtues are that it combines a rich yet parsimonious demographic model with statistical time series methods. Annual age-specific death rates are used for this method.

Let *m(x, t)* denote the central death rate for age x in year t.

The original Lee-Carter approach goes on to forecast *m(x, t)* for given *t*, using the model

$$\log_e[m(x,t)] = a_x + b_x k_t + \varepsilon(x,t) \qquad (2.1)$$

where $\varepsilon(x,t)$ is a random error component. According to Lee and Carter (1992), the model parameters a_x, b_x, and k_t are estimated using a rank-1 approximation of the singular value decomposition (SVD) of the matrix $\{\log_e[m(x,t)] - a_x\}$ as discussed by Good (1969); and Johnson and Wichern (2002), and the final 'adjusted' estimate of k_t (as explained by Lee and Carter, 1992) is then modeled using a standard time series approach. In particular, the LC approach consists of six key steps: Namely,

1. Estimating model parameters using rank–1 approximation of the SVD of the matrix $\{\log_e[m(x,t)] - a_x\}$. Here, the success of long term forecasting totally depends on the goodness-of-fit of the rank-1 approximation.
2. Re-scaling b_x and k_t so that b_x sum to unity and k_t sum to 0.
3. Adjusting k_t, the time-varying index, so that actual total number of deaths is equal to the fitted total number of deaths, for each t .
4. Modeling adjusted k_t using standard time series approach.
5. Forecasting k_t , and hence *m(x, t)* , for future time periods, and finally
6. Constructing life tables and forecasting life expectancy at birth (or at specific ages).

 ISBN-1391-4987

2.3 Singular Value Decomposition (SVD) in the LC Approach

Let A be an $m \times k$ real matrix. Then there exists orthogonal matrices $U_{m\times m}$ and $V_{k\times k}$ such that $A = U\,\Lambda_{m\times k}\,V'$ where $U'U = I_m$ and $V'V = I_k$. The '*m*' columns of *U* are the 'left singular vectors'; the 'k' rows of V' are the 'right singular vectors'.

The normalized eigenvectors of AA' make up the columns of *U* and the normalized eigenvectors of $A'A$ make up the columns of *V*. The 'singular values' (λ_i's) in Λ arranged in descending order, are the square roots of eigenvalues (λ_i^2) of AA' or $A'A$, since they have the same eigenvalues (Good, 1969; Johnson and Wichern, 2002).

Rank r approximation of SVD

This is given by $A \approx U_{m\times r}\,\Lambda_{r\times r}\,V'_{r\times k}$ where $U = (\underline{u}_1, \underline{u}_2, ..., \underline{u}_r)$; $V = (\underline{v}_1, \underline{v}_2, ..., \underline{v}_r)$ and Λ is a diagonal matrix with the *r* singular values $\lambda_1, \lambda_2, ..., \lambda_r$ along the diagonal in descending order. This can be written as $A = \lambda_1\,\underline{u}_1\,\underline{v}_1' + \lambda_2\,\underline{u}_2\,\underline{v}_2' + ... + \lambda_r\,\underline{u}_r\,\underline{v}_r'$.

Thus, rank-1 approximation of the SVD is $A \approx U_1\,\Lambda_1\,V_1'$ where $U_1 = \underline{u}_1$; $V_1 = \underline{v}_1$ and $\Lambda_1 = \lambda_1$ so that $A \approx \lambda_1\,\underline{u}_1\,\underline{v}_1'$.

SVD on MATLAB®

In MATLAB®,

[*U*,*S*,*V*] = svd (*X*)

produces a diagonal matrix *S* of the same dimensions as *X* with non-negative diagonal elements in decreasing order, and *U* & *V* such that $U'U=I_m$ and $V'V=I_k$.

[*U*,*S*,*V*] = svd (*X*, 0) produces 'economy size' decomposition. If *X* is $m \times k$ with

$m > k$, then only the first *k* columns of *U* is printed and *S* is $k \times k$ now. If $m < k$, [*U*,*S*,*V*] = svd (*X*, 'econ') produces only the first m columns of *V* and *S* is $m \times m$ now.

Example

Consider the following hypothetical example of 5×3 matrix of *m(x, t)* values:

$$A_1 = \begin{bmatrix} .00930 & .00780 & .00665 \\ .00030 & .00027 & .00022 \\ .00018 & .00015 & .00013 \\ .00020 & .00017 & .00015 \\ .00065 & .00060 & .00057 \end{bmatrix}$$

Let A be the matrix representing $\log_e[m(x,t)]$,

$$\text{Then } A = \begin{bmatrix} -4.6777 & -4.8536 & -5.0131 \\ -8.1117 & -8.2171 & -8.4219 \\ -8.6225 & -8.8049 & -8.9480 \\ -8.5172 & -8.6797 & -8.8049 \\ -7.3385 & -7.4186 & -7.4699 \end{bmatrix}.$$

[*U*,*S*,*V*] = svd (*A*, 0) gives,

$$U = \begin{bmatrix} -.2805 & -.7755 & -.2260 \\ -.4772 & -.1062 & .8715 \\ -.5085 & -.0722 & -.3106 \\ -.5013 & .0849 & -.2738 \\ -.4285 & .6123 & -.1337 \end{bmatrix}; S = \begin{bmatrix} 29.9471 & 0 & 0 \\ 0 & .1499 & 0 \\ 0 & 0 & .0477 \end{bmatrix}; \text{ and}$$

$$V = \begin{bmatrix} .5671 & -.6971 & -.4387 \\ .5774 & -.0434 & .8153 \\ .5874 & .7157 & -.3779 \end{bmatrix}.$$

Now, rank-1 approximation to SVD of A is given by

$$A \approx (29.9471)\underline{u}_1\,\underline{v}_1' = (29.9471) \begin{bmatrix} -.2805 \\ -.4772 \\ -.5085 \\ -.5013 \\ -.4285 \end{bmatrix} [.5671 \quad .5774 \quad .5874].$$

The goodness-of-fit of this approximation is = 29.9471/ 30.1447 = .99, or 99% (where 30.1447 is the trace of *S*).

Now let $\underline{b}_X = \dfrac{\underline{u}_1}{-2.1960} = [.1277 \quad .2173\; .2316 \quad .2283 \quad .1951]'$ where -2.1960 is the sum of the elements of $\underline{u}_1$; Then, $\sum_X b_X = 1$.

Let $\underline{k}_t^0$ be equal to (-2.1960)(29.9471) $\underline{v}_1'$ =[-37.2947 -37.9720 -38.6297] which does not sum to 0. But if we define $\bar{k}$ to be the mean of $\underline{k}_t^0$, then $\bar{k}$ = -37.9655 and now $\underline{k}_t = \underline{k}_t^0 - \bar{k}$ = [.6708 -.0065 -.6642], so that $\sum_t \underline{k}_t$ = 0.

Thus, $A \approx \underline{b}_x(\underline{k}_t + \bar{k}) = \underline{b}_x \underline{k}_t + \bar{k}\,\underline{b}_x$ where $\bar{k}\,\underline{b}_x$ is equal to $\underline{a}_x$, so that the matrix
$\log_e(m(x,t)) - a_x \approx \underline{b}_x\, \underline{k}_t$ such that it satisfies the conditions of L-C model.

Lee and Miller (2001) make a careful assessment of the performance of LC-method for forecasting mortality far into the future. They have not followed standard time series diagnostic methods to choose an optimum ARIMA model. Instead, they have assumed the random walk with a drift approach. According to Lee and Miller (2001), the LC-method still performed better than subjective expert judgment in most cases, but change in population ageing patterns, thus mortality rates in older populations, may pose problems for LC approach. Li *et al* (2004) discuss ways in which the LC approach can be used as a method to reduce the role of subjective judgment particularly in countries with limited data at unequal intervals.
In most applications to date, it has been found that a random walk with a drift (Lee and Carter, 1992) fits well though it is not always the best model overall. Unless some other time series model is found to be substantially better, it is advisable to use the random walk with drift because of its simplicity and straight forward interpretation.

2.4 Life Tables

A life table is simply an elegant and convenient way of analyzing age-specific death rates. The technique is now used in several other areas, where it is often called “survival analysis”. Life tables answer many questions that cannot be answered with simple measures of rates and ratios. Questions of this nature have immense practical importance in many disciplines, such as education, health, insurance, and actuarial science. Life tables provide summary measures of the level of mortality, independent of the age composition which can therefore be used for comparing the mortality levels of different populations.
The basic inputs for construction of life tables are age-specific mortality rates from which all other columns are derived. The data required to obtain such mortality rates are, (i) the distribution of population by age and sex and, (ii) the distribution of deaths by age and sex. A complete life table contains data for

every single year of age. An abridged life table, on the other hand, typically contains data by 5-year age intervals.

The notations used for different columns of a life table is given below:

x	: exact age
(x, x + n)	: age group with initial age x with the length of interval n
m(x, n)	: age – specific death rate for age interval (x, x + n)
q(x, n)	: probability of an individual aged x dying before the end of the interval (x, x + n)
l(x)	: number of survivors at age x in a life table with radix (starting population) of 100,000 persons
d(x, n)	: number of deaths in age interval (x, x + n)
L(x, n)	: number of person – years lived in age interval (x, x + n)
S(x, n)	: the proportion of the life table population in age group (x, x + n) who are alive n years later.
T(x)	: number of person – years lived at ages x and older.
e(x)	: expectation of life at age x, and
a(x, n)	: average number of years lived in the age interval (x, x + n) by those dying during that age interval.

Age-specific death rates are defined as:

$$m(x,n) = \frac{\text{number of deaths in the age group } (x,x+n)}{\text{Population of the age group at mid interval}}$$

Once the basic inputs are provided, the life tables can be calculated using most software packages. In this study, the freely downloadable software package MORTPAK4®[2] was used.

2. Descriptive Analysis

This section describes the data relating to population growth and the number of deaths registered in Sri Lanka from 1950 to 2008. Sex-specific death rates, considering both information of population and the number of death registered are also discussed.

3.1 Population Growth

Figure 3.1 shows the gender-wise midyear population 1950 up to 2010 (Source: Registrar General's Office and Department of Census and Statistics, Sri Lanka). Both have sharply upward trends, and the gap between the two graphs decrease

[2] See www.un.org/esa/population/mortpak/mortpakwebpage.pdf

with time. From 1950 up to 2001, male population is greater than female population. After 2001, it is lower than the female population.

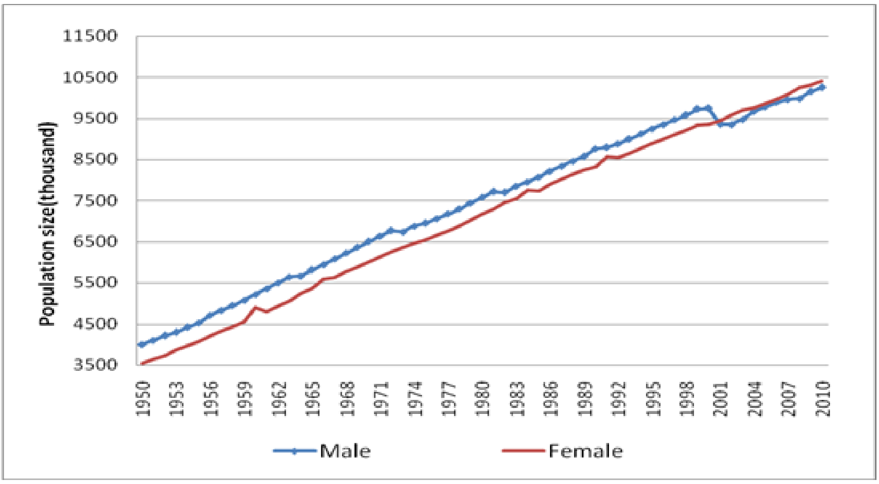


Figure 3.1: Time plot of midyear population size*
*Source: Department of Census and Statistics, Sri Lanka

Even though the Sri Lankan population size has been gradually increasing (7.54m in 1950 to 20.2m in 2010) the average annual growth rate is seen to be declining (3.0 in 1960 to 0.7 in 2010). An ageing population is commonly defined as one with an increasing proportion of the population in the elderly age groups. Table 3.1 (Source: Registrar General's Office and Department of Census and Statistics, Sri Lanka) clearly shows the changes in the percentage distribution of selected age groups.

Table 3.1: Percentage distribution of the population in selected age groups in Sri Lanka, for the period 1950 – 2010*

Year	Age group		
	0 -14 (%)	**15-54 (%)**	**55 and above (%)**
1950	37.22	54.90	7.87
1960	39.84	52.60	7.56
1970	41.91	49.61	8.49
1980	38.98	51.92	9.10
1990	35.00	55.60	9.41
2000	35.21	55.33	9.46
2010	26.30	60.50	13.20

*Source: Registrar General's Office, Department of Census and Statistics, Sri Lanka

In Table 3.1 the elderly group is defined as those who are 55 years or more. The reason for this is that in both the government and private sectors in Sri Lanka, the retirement age falls between 55 and 60 years. According to Table 3.1, the elderly population is gradually increasing but the child population has increased up to 1970 and decreased thereafter. Middle age group has decreased up to 1970 and increased thereafter.

3.2 Deaths

The registered number of deaths reveals the effects of some historical events. Figure 3.2 shows that the number of male deaths is always greater than the number of female deaths. There are four peak points in 1974, 1989, 1996 and 2005 for the male graph but only two peak points in the female graph, namely 1974 and 2005. The reasons for these may be the effects of the youth unrests that occurred in 1971 and in 1987-89, the separatist war that took heavy tolls from time to time and especially an escalation of events in 1996, and the tsunami disaster that occurred in 2004 in the South, respectively. Except for the tsunami that took a mix of both male and female lives (with a higher proportion of females), the other events took mostly male lives.

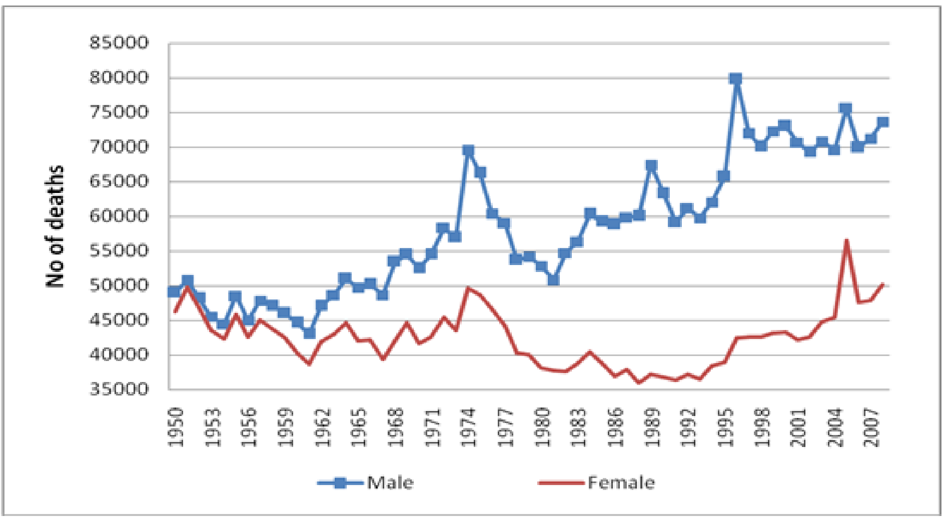


Figure 3.2: Time plot of gender wise total number of deaths*
*Source: Department of Census and Statistics, Sri Lanka

4. Results and Discussion

4.1 The LC model

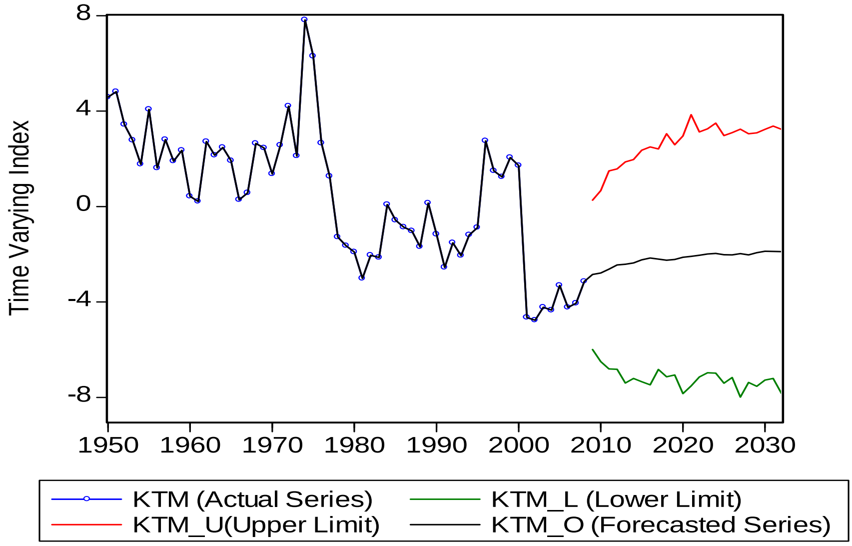
Rank-1 approximation to the exact SVD solution in the LC method was used for estimating k_t , a_x, and b_x of equation (2.1) for both male and female data, as explained in Sections 2.2 and 2.3. The rank-1 approximation gave a goodness-of-fit of 91% and 89% for the female and male data respectively. Table 4.1 shows the estimates of a_x and b_x for both sexes for the period 1950 – 2008.

Table 4.1: Estimates of a_x and b_x for 1950 – 2008 for males and females

Age Group	Males		Females	
	a_x	b_x	a_x	b_x
0 – 4	-4.68823	0.05897	-4.82197	0.05702
5 – 9	-6.78061	0.08529	-6.80158	0.08043
10 – 14	-7.06443	0.08886	-7.19796	0.08512
15 – 19	-6.51973	0.08201	-6.66236	0.07879
20 – 24	-6.05618	0.07617	-6.41775	0.07590
25 – 29	-5.94682	0.07480	-6.32310	0.07478
30 – 34	-5.87618	0.07391	-6.21934	0.07355
35 – 39	-5.54524	0.06975	-5.99724	0.07092
40 – 44	-5.28597	0.06649	-5.82504	0.06889
45 – 49	-4.89645	0.06159	-5.47108	0.06470
50 – 54	-4.57742	0.05757	-5.11910	0.06054
55 – 59	-4.15225	0.05223	-4.65699	0.05507
60 – 64	-3.82528	0.04811	-4.24522	0.05020
65 – 69	-3.33825	0.04199	-3.63993	0.04305
70 – 74	-2.91610	0.03668	-3.10025	0.03666
75 and over	-2.03447	0.02559	-2.06112	0.02437

The plot of adjusted k_t for the period 1950 – 2008 along with 30 year forecasts with 95% confidence bands for both sexes (with '$K_t^{(m)}$' for males and '$K_t^{(f)}$' for females) is shown in Figure 4.1(a) and 4.1(b) respectively. The forecasts were made using the models that performed better than the others considered (i.e. VAR model). One striking observation that was made from these forecasts was that even though the adjusted k_t series generally showed a downward trend in both graphs in agreement with that of LC for their study period, the forecasts for the period 2010 – 2030 showed a slightly upward trend. This upward trend is probably a result of the increase in the mortality index from 2000 to 2001 in both graphs, Figure 4.1(a) and 4.1(b).

(a) Adjusted k_t for males with forecasts



(b) Adjusted k_t for females with forecasts

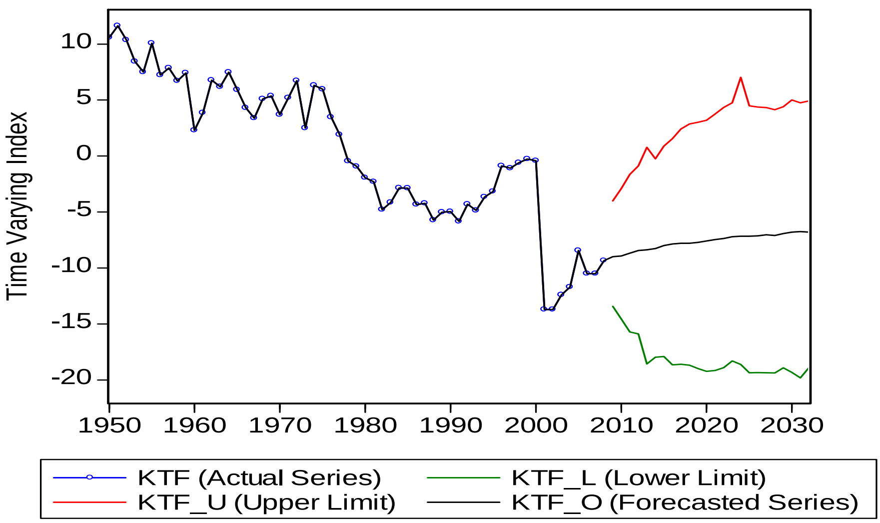


Figure 4.1: Adjusted time-varying index (k_t) for both males and females with forecasts.

Therefore, year 2001 looked like a turning point in our data set; however, this may be a short term turn and as we move on to the future when new data are available, the updated picture may be consistent with the general downward trend of LC. However, the lower bound of the 95% confidence interval based on the VAR model in both graphs 4.1(a) and 4.1(b), showed a downward trend over the period 2010 – 2030 indicating a possibility for the actual forecasts to follow the usually expected downward trend. These will be further discussed in Section 4.2.

4.2 Modeling Adjusted k_t Series

In our case, the adjusted k_t series (i.e. the time varying index) for both males (denoted by $K_t^{(m)}$) and females (denoted by $K_t^{(f)}$), were modeled using a standard time series approach. First of all, the augmented Dickey-Fuller unit root test (Dickey and Fuller, 1979) was applied to the two series, and it was found that the two series were integrated of order 1. After finding the order of differencing, ARIMA models were fitted to both series using the package EVIEWS® and based on Akaike Information Criterion (AIC), it was found that ARIMA (1,1,1) had the least AIC values for both male data (AIC = 3.89) and female data (AIC = 4.54). However, model adequacy tests failed to accept ARIMA (1,1,1) as a suitable model in both cases. Therefore, the two series were first modeled using a random walk with drift as well as random walk with drift in the presence of intervention, though the random walk models were not the best fitted. This was done in view of higher residuals at 1974 and 2001 for $K_t^{(m)}$ series and at 2001 for $K_t^{(f)}$ series, by following an approach similar to that of Lee and Carter (1992). Then a vector autoregressive (VAR) model that combines data from both series was fitted to first differenced series anticipating better results, and thus forecasts. These fitted models are shown below:

Random Walk Models for $K_t^{(m)}$ Series:

(with drift):

$$\mathbf{K_t^{(m)}} - \mathbf{K_{t-1}^{(m)}} = -0.1331 \quad (4.1)$$

(with drift and intervention at 1974 and 2001):

$$\mathbf{K_t^{(m)}} - \mathbf{K_{t-1}^{(m)}} = -0.1261 - 0.2029*WT \quad (4.2)$$

where WT = 1, when year = 1974 or 2001; and 0 otherwise.

Random Walk Models for $K_t^{(f)}$ Series:

(with drift):

$$K_t^{(f)} - K_{t-1}^{(f)} = -0.3436 \quad (4.3)$$

(with drift and intervention at 2001) :

$$K_t^{(f)} - K_{t-1}^{(f)} = -0.1167 - 13.1633*PT \quad (4.4)$$

Where PT = 1, when year = 2001; and 0, otherwise.

VAR models, traditionally designed for stationary series, have proven to be especially useful for dynamic forecasting of economic and financial time series and usually provide superior forecasts to those from univariate time series models under certain conditions. Recent studies show that VAR models have been used for mortality modeling as well. When there is no cointegration, standard VAR models are usually used whereas Error Corrected VAR models (VEC) are used in the presence of cointegration (Lütkepohl, 2005). Due to Lütkepohl (2005), VAR models can also be used when the cointegration structure is unknown. Since the order of integration was 1 in both series and there was no cointegration, standard VAR models were fitted to the vector of first differenced series using EVIEWS® package. The form of the VAR model was such that each variable has an equation explaining its evolution based on its own lags and the lags of the other variable (lag 1 in our case) and the fitted models (in terms of the current series) are shown below:

VAR Model for $K_t^{(m)}$ Series:

$$K_t^{(m)} = 0.1454* K_{t-1}^{(f)} + 0.4990* K_{t-1}^{(m)} + 0.0552 \quad (4.5)$$

Standard errors of the coefficients in model (4.5) are .0715, .1672, and .2230 respectively and R^2 was 67%.

VAR Model for $k_t f$ Series:

$$K_t^{(f)} = 1.0592* K_{t-1}^{(f)} - 0.3432* K_{t-1}^{(m)} - 0.1981 \quad (4.6)$$

Standard errors of the coefficients in model (4.6) are .1045, .2445, and .3260 respectively. The goodness-of-fit given by the R^2 was 87%. As we have observed from the beginning, the female series was much smoother than the male one and the SVD gave a better fit (91%) for the female data than for the male data (89%). This is further confirmed here for the fitting of VAR models. As a result more accurate forecasts are anticipated for female data than for male data in our study.

The models (4.5) and (4.6) did not show any notable violations of the diagnostics and it looked like the problem that arose with higher residuals at given years in the random walk models are now taken care of by the VAR models. Therefore, models (4.5) and (4.6) were used for forecasting in our study.

4.3 Calculation of Death Rates

For both male and female models that were fitted, ln[m(x,t)], i.e. the logarithm of the death rate, was calculated using the following equation:

$$\ln[\hat{m}(x,t)] = a_x + b_x \hat{k}_t \quad (4.7)$$

where estimates of a_x and b_x are obtained from Table 4.1, and k_t from equations (4.1) – (4.6). Figure 4.2 shows a comparison of actual death rates (in logarithm scale) for the years 1950, 1975, and 2000 with the forecasts for 2025 using model (4.5) for the male series.

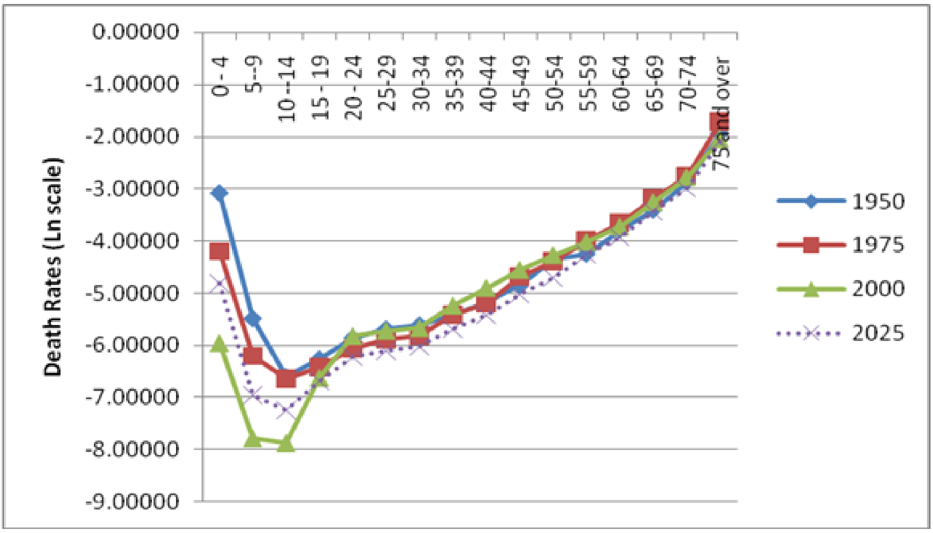


Figure 4.2: Comparison of actual and forecast death rates by age for males, at selected years (actual values in solid and forecasts in dots).

The goodness-of-fit of model (4.5) was tested using R^2 which was 67%. It was observed that the forecasts for 2025 given by (4.5) matched the overall shape of actual death rates in log scale for the given years 1950, 1975, and 2000 over the age groups well. In fact, the forecasts for 2025 were between those for 1975 and 2000 in the age groups 0 – 4, 5 – 9, and 10 – 14 and were very close (lying slightly lower) to those of 1950, 1975, and 2000 in all other age groups.

 ISBN-1391-4987

Figure 4.3 shows a comparison of actual death rates (in logarithm scale) for the years 1950, 1975, and 2000 with the forecasts for 2025 using model (4.6) for the female series.

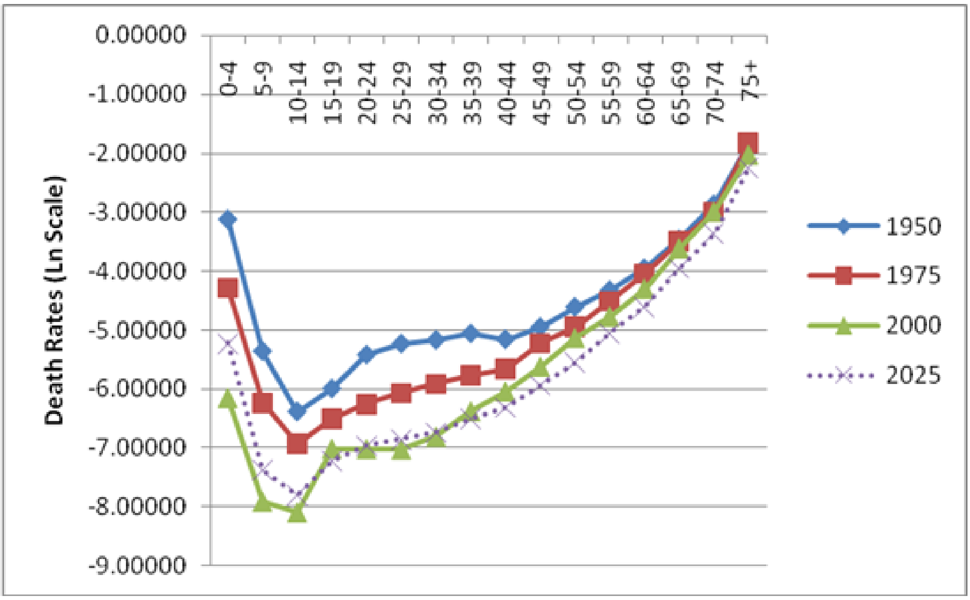


Figure 4.3: Comparison of actual and forecast death rates by age for females, at selected years (actual values in solid and forecasts in dots)

The goodness-of-fit of model (4.6) was tested using R^2 which was 87%. It was observed that the forecasts for 2025 given by (4.6) matched the overall shape of actual death rates in log scale for the given years 1950, 1975, and 2000 over the age groups well. In fact, the forecasts for 2025 were between those for 1975 and 2000 in the age groups 0 – 4, 5 – 9, and 10 – 14 and were almost equal to those of 2000 in the age groups 15 – 19, 20 – 24, 25 – 29, and 30 – 34; and were lying slightly below those of 2000 in all other age groups.

4.4 Calculation of Life Tables

Abridged life tables were generated using actual data for a selected period, namely 2006 – 2008 using the software package MORTPAK4®. Tables 4.2 and 4.3 show these generated life tables for the male and the female data respectively. In these tables, age-specific death rates were obtained for individual ages 0, 1, 5, 10, …, 75, 80, and 85, in the given period.

Table 4.2: Abridged life tables for males for the period 2006 – 2008 using actual death rates

Life table for males 2006-2008									
Age	m(x,n)	q(x,n)	l(x)	d(x,n)	L(x,n)	S(x,n)	T(x)	e(x)	a(x,n)
0	0.01098	0.01087	100000	1087	98994	0.98818	7035232	70.352	0.074
1	0.00059	0.00237	98913	234	395097	0.99777	6936238	70.124	1.620
5	0.00033	0.00165	98679	163	492989	0.99835	6541142	66.287	2.500
10	0.00033	0.00164	98516	161	492177	0.99740	6048153	61.393	2.500
15	0.00082	0.00410	98355	403	490899	0.99402	5555975	56.489	2.829
20	0.00160	0.00796	97952	780	487962	0.99046	5065077	51.710	2.698
25	0.00215	0.01072	97171	1041	483308	0.98940	4577115	47.103	2.552
30	0.00210	0.01043	96130	1003	478187	0.98832	4093807	42.586	2.543
35	0.00270	0.01342	95127	1277	472601	0.98407	3615620	38.008	2.621
40	0.00386	0.01912	93851	1794	465070	0.97539	3143019	33.490	2.668
45	0.00628	0.03095	92057	2849	453624	0.96318	2677949	29.090	2.662
50	0.00895	0.04384	89208	3911	436920	0.94204	2224325	24.934	2.669
55	0.01543	0.07447	85296	6352	411597	0.91058	1787405	20.955	2.657
60	0.02216	0.10523	78944	8307	374790	0.87652	1375808	17.428	2.601
65	0.03126	0.14536	70637	10268	328511	0.82251	1001017	14.171	2.597
70	0.04824	0.21593	60369	13036	270203	0.74485	672507	11.140	2.573
75	0.07174	0.30503	47334	14438	201261	0.61985	402304	8.499	2.548
80	0.12430	0.47138	32895	15506	124752	0.37947	201043	6.112	2.438
85	0.22793	...	17389	17389	76291	...	76291	4.387	4.387

In Table 4.2, the first entry of S(x,n) is for survivorship of 5 cohorts of birth to age group 0 – 4 = L(0,5) / 500000. Second entry of S(x,n) is for,
S(0,5) = L(5,5) / L(0,5), and the last entry of S(x,n) is S(80+,5) = T(85) / T(80).

Table 4.3: Abridged life tables for females for the period 2006 – 2008 using actual death rates

Life table for females 2006-2008									
Age	m(x,n)	q(x,n)	l(x)	d(x,n)	L(x,n)	S(x,n)	T(x)	e(x)	a(x,n)
0	0.00891	0.00884	100000	884	99184	0.99032	7685034	76.850	0.077
1	0.00049	0.00197	99116	195	395978	0.99819	7585850	76.535	1.510
5	0.00028	0.00137	98921	136	494264	0.99865	7189872	72.683	2.500
10	0.00026	0.00132	98785	131	493597	0.99821	6695609	67.780	2.500
15	0.00048	0.00241	98654	238	492713	0.99739	6202012	62.866	2.651
20	0.00055	0.00274	98417	270	491429	0.99688	5709299	58.012	2.576
25	0.00070	0.00349	98147	343	489896	0.99647	5217870	53.164	2.554
30	0.00072	0.00358	97804	350	488166	0.99598	4727974	48.341	2.556
35	0.00092	0.00459	97454	447	486205	0.99472	4239808	43.506	2.614
40	0.00125	0.00622	97007	603	483636	0.99169	3753603	38.694	2.678
45	0.00219	0.01088	96404	1049	479617	0.98629	3269967	33.919	2.707
50	0.00345	0.01710	95355	1630	473043	0.97689	2790350	29.263	2.710
55	0.00619	0.03051	93725	2859	462112	0.95948	2317307	24.724	2.722
60	0.01065	0.05197	90866	4722	443387	0.93575	1855195	20.417	2.683
65	0.01656	0.07978	86144	6872	414898	0.88805	1411807	16.389	2.698
70	0.03252	0.15117	79271	11983	368452	0.81312	996909	12.576	2.671
75	0.05222	0.23249	67288	15644	299595	0.67766	628458	9.340	2.645
80	0.10984	0.43179	51644	22300	203024	0.38265	328863	6.368	2.525
85	0.23319	...	29345	29345	125839	...	125839	4.288	4.288

In Table 4.3, the first entry of S(x,n) is for survivorship of 5 cohorts of birth to age group 0 – 4 = L(0,5) / 500000. The second entry of S(x,n) is for S(0,5) = L(5,5) / L(0,5), and
last entry of S(x,n) is S(80+,5) = T(85) / T(80).

Outcomes of the life tables

From Tables 4.2 and 4.3 that were based on 2006 – 2008 data, it appears that the life expectancy at birth for males in Sri Lanka is 70.3 years while for females it is 76.8 years. The column q(x,n) can be used for calculating the probability of an individual aged x surviving another n years (where n is the considered age group size). This probability gives important information as input for state sector decision makers as well as private sector financial institutions. For example, the probability that a 50-year old male surviving another 5 years is (1 – q(50, 5)) = 1- 0.04384 = 0.95616 from Table 4.2. A similar probability for a 50-year old female is = 1 – 0.01710 = 0.9829 from Table 4.3.

One may also be able to calculate the probability of a person aged x surviving a longer period than n years, especially a period which is a multiple of n, using information from the life tables.
For example, the probability of a male aged 55 years surviving another 20 years (${}_{20}p_{55}$, say) may be of interest to many financial institutions. Using the fact that events are independent, it can be calculated from Table 4.2, as follows:
${}_{20}p_{55} = (1 - {}_{5}q_{55})(1 - {}_{5}q_{60})(1- {}_{5}q_{65}) = (0.92553)(0.89477)(0.85464) = 0.71.$
A similar probability for a female aged 55 years is
= (0.96945)(0.94803)(0.92022) = 0.85.

4.5 Mortality Forecasts

Using the VAR models that were considered as the best fitted in our study, relatively accurate mortality forecasts (*i.e*. forecasts of time-varying index k_t) can be made for males as well as females using the two fitted models (4.5) and (4.6) respectively, on the assumption that conditions that prevailed during the study period will not drastically change over the forecast period. This attempt of forecasting for 30 years ahead depends on the ability of the rank-1 approximation to SVD to capture the trend in the (adjusted) mortality index k_t for both sexes. On the other hand, even though they were not the best fitting models, random walk models with drift (with and without intervention) were also used for forecasting mortality index. A comparison of forecasts using VAR models and random walk models for both series, male and female, are given in Tables 4.4 and 4.5. It seems that the widths of 95% confidence intervals for the forecasts from the random walk models (with or without intervention) appear to be unacceptably high when compared to those from the VAR models in both Tables 4.4 and 4.5. The confidence intervals for the forecasts during the period 2010 – 2030 obtained from the VAR models appear to be relatively more accurate even though the forecasts showed a slightly upward trend during the forecast period as discussed in section 4.1.

Table 4.4: Comparison of forecasts along with Lower Limit (L.L.) and Upper Limit (U.L.) of 95% Confidence Intervals for mortality index ($K_t^{(m)}$) for male data (2009-2032) using three different methods

	Method1[3]			Method2[4]			Method3[5]		
Period	**Forecast**	**L.L**	**U.L**	**Forecast**	**L.L**	**U.L**	**Forecast**	**L.L**	**U.L**
2009	-3.27312	-30.3411	23.79488	-2.89594	-6.13464	0.125302	-3.26613	-30.5686	24.03635
2010	-3.40624	-30.7027	23.89018	-2.76359	-6.68348	1.405362	-3.39225	-30.9251	24.14063
2011	-3.53936	-31.0623	23.98359	-2.57093	-6.61885	1.787354	-3.51838	-31.2798	24.243
2012	-3.67248	-31.4201	24.07515	-2.48342	-7.14966	1.936723	-3.6445	-31.6325	24.3435
2013	-3.8056	-31.7761	24.16491	-2.44648	-7.07417	2.197774	-3.77063	-31.9834	24.44218
2014	-3.93872	-32.1304	24.2529	-2.41709	-7.12387	1.645799	-3.89675	-32.3326	24.53909
2015	-4.07185	-32.4829	24.33917	-2.43113	-7.09737	2.072679	-4.02288	-32.68	24.63425
2016	-4.20497	-32.8337	24.42377	-2.39395	-6.9072	1.915434	-4.149	-33.0257	24.72773
2017	-4.33809	-33.1829	24.50671	-2.24717	-7.47018	2.40013	-4.27513	-33.3698	24.81954
2018	-4.47121	-33.5305	24.58806	-2.25574	-7.70979	2.565054	-4.40125	-33.7122	24.90974
2019	-4.60433	-33.8765	24.66783	-2.24312	-7.39284	2.595807	-4.52738	-34.0531	24.99835
2020	-4.73745	-34.221	24.74606	-2.12235	-7.04576	2.869309	-4.6535	-34.3924	25.08541
2021	-4.87057	-34.5639	24.82279	-2.14778	-7.37227	2.993446	-4.77963	-34.7302	25.17095
2022	-5.00369	-34.9054	24.89805	-2.12024	-7.6964	2.475982	-4.90575	-35.0665	25.25501
2023	-5.13681	-35.2455	24.97186	-2.03951	-7.18069	3.082094	-5.03188	-35.4014	25.33761
2024	-5.26993	-35.5841	25.04426	-2.01613	-7.31247	3.050106	-5.158	-35.7348	25.41879
2025	-5.40305	-35.9214	25.11528	-1.96135	-7.64126	3.549703	-5.28413	-36.0668	25.49857
2026	-5.53617	-36.2573	25.18494	-1.94547	-7.09025	3.163079	-5.41025	-36.3975	25.57698
2027	-5.66929	-36.5919	25.25327	-1.92611	-7.34813	3.414653	-5.53638	-36.7268	25.65406
2028	-5.80241	-36.9251	25.3203	-1.97274	-7.28994	3.143004	-5.6625	-37.0548	25.72981
2029	-5.93553	-37.2571	25.38605	-1.9846	-7.13241	3.387059	-5.78863	-37.3815	25.80428
2030	-6.06866	-37.5879	25.45054	-1.98787	-7.0372	3.359687	-5.91475	-37.707	25.87748
2031	-6.20178	-37.9174	25.5138	-1.92232	-7.10972	3.234121	-6.04088	-38.0312	25.94944
2032	-6.3349	-38.2457	25.57586	-1.83974	-7.2004	3.405044	-6.167	-38.3542	26.02018

[3] Random Walk Model with Drift : $K_t^{(m)} - K_{t-1}^{(m)} = -0.133121$

[4] Vector Autoregressive (VAR) Model: $K_t^{(m)} = 0.145375603 * K_{t-1}^{(f)} + 0.4989909555 * K_{t-1}^{(m)} + 0.05517314328$

[5] Random Walk Model with Drift + Dummy Variable : $K_t^{(m)} - K_{t-1}^{(m)} = -0.126125 - 0.202875*WT$

where WT = 1 for year =1974 and year =2001 and 0 otherwise

Table 4.5: Comparison of forecasts along with Lower Limit (L.L.) and Upper Limit (U.L.) of 95% Confidence Intervals for mortality index ($K_t^{(f)}$) for female data (2009-2032) using three different methods

Period	Method1[6]			Method2[7]			Method3[8]		
	Forecast	L.L	U.L	Forecast	L.L	U.L	Forecast	L.L	U.L
2009	17.55644	-20.2753	55.38813	88.03428	60.85152	115.217	-9.04252	-13.4617	-4.76413
2010	17.67451	-20.4765	55.82547	89.57	62.16153	116.9785	-8.87906	-15.1067	-1.50658
2011	17.79259	-20.675	56.26016	91.10572	63.47338	118.7381	-8.57138	-15.8173	-1.35777
2012	17.91066	-20.8709	56.69226	92.64144	64.78703	120.4959	-8.40689	-16.2469	0.61996
2013	18.02874	-21.0644	57.12184	94.17717	66.10243	122.2519	-8.36354	-16.9413	0.218174
2014	18.14681	-21.2553	57.54895	95.71289	67.41955	124.0062	-8.21932	-17.8451	0.316045
2015	18.26489	-21.4439	57.97366	97.24861	68.73835	125.7589	-8.21429	-18.1698	0.971324
2016	18.38296	-21.6301	58.39603	98.78433	70.05878	127.5099	-8.15736	-17.9066	0.878609
2017	18.50104	-21.814	58.81609	100.3201	71.38082	129.2593	-7.98432	-18.1956	2.020721
2018	18.61911	-21.9957	59.23391	101.8558	72.70443	131.0071	-7.87947	-19.3038	2.307763
2019	18.73719	-22.1752	59.64953	103.3915	74.02956	132.7534	-7.75721	-19.2299	3.012978
2020	18.85526	-22.3525	60.06301	104.9272	75.3562	134.4982	-7.54641	-18.9038	2.952696
2021	18.97334	-22.5277	60.47438	106.4629	76.6843	136.2416	-7.61481	-20.0339	3.766461
2022	19.09141	-22.7009	60.8837	107.9987	78.01384	137.9835	-7.54753	-19.6202	3.709595
2023	19.20949	-22.872	61.29099	109.5344	79.34479	139.724	-7.35565	-19.0295	4.097897
2024	19.32756	-23.0412	61.69632	111.0701	80.67712	141.4631	-7.21228	-20.3787	3.946288
2025	19.44564	-23.2084	62.09971	112.6058	82.0108	143.2009	-7.09191	-18.9601	4.640648
2026	19.56371	-23.3738	62.5012	114.1416	83.3458	144.9373	-6.99333	-19.3068	4.956842
2027	19.68179	-23.5373	62.90084	115.6773	84.68211	146.6724	-6.91172	-19.4663	5.404002
2028	19.79986	-23.6989	63.29865	117.213	86.01969	148.4063	-6.97585	-19.5051	5.262581
2029	19.91794	-23.8588	63.69468	118.7487	87.35852	150.1389	-6.93265	-19.1	4.788668
2030	20.03601	-24.0169	64.08895	120.2844	88.69858	151.8703	-6.93248	-18.7512	6.046848
2031	20.15409	-24.1733	64.4815	121.8202	90.04319	153.5971	-6.74346	-19.4279	5.276962
2032	20.27216	-24.3234	64.86775	123.3559	91.38893	155.3228	-6.59377	-19.382	5.660622

[6] Random Walk Model with Drift : $K_t^{(f)} - K_{t-1}^{(f)} = -0.34362$

[7] Random Walk Model with Drift + Dummy Variable1: $K_t^{(f)} - K_{t-1}^{(f)} = -0.1166667 - 13.16333*PT$
where PT=1 when year = 2001 and PT=0 otherwise

[8] Vector Autoregressive (VAR) Model : $K_t^{(f)} = 1.059236473 * K_{t-1}^{(f)} - 0.343156097 * K_{t-1}^{(m)} - 0.1980843972$

Finally, Table 4.6 shows forecasts of age-specific death rates per 100,000 at 5-year intervals for the period 2010 – 2030 using our fitted VAR models for both male and female data. In Table 4.6, the forecasts for the period 2010 – 2030 shows an upward trend in each age group which is quite expected as a result of the upward trend in the mortality forecasts shown in Figures 4.1(a) and 4.2(b). However, the lower limits of 95% confidence intervals for the forecasts of age-specific death rates per 100,000 showed a downward trend in agreement with those of LC for the same period 2010 – 2030 suggesting the possibility of the actual death rates to show a downward trend. For example, for the forecasts of male death rates per 100,000, the lower bounds of the 95% confidence intervals for the years 2010, 2015, 2020, 2025, and 2030 were 65, 63, 63, 60, and 63 respectively in the 5-9 age group; 113, 110, 110, 105, and 110 respectively in the 25 – 29 age group; and 634, 618, 620, 598, and 620 in the 55–59 age group. Corresponding values for the female death rates per 100,000 were 33, 26, 25, 24, and 25 in the 5 – 9 age group; 63, 50 48, 48, and 48 in the 25 – 29 age group; and 382, 318, 304, 303, and 306 in the 55 – 59 age group.

Table 4.6: Forecasts* of age-specific death rates per 100,000 at five-year intervals, 2010 – 2030

age class	Males					Females				
	2010	**2015**	**2020**	**2025**	**2030**	**2010**	**2015**	**2020**	**2025**	**2030**
0-4	782	797	812	820	818	485	504	523	537	542
5-9	90	92	95	96	96	54	57	61	63	64
10-14	67	69	71	72	72	35	37	39	41	41
15-19	118	121	124	126	125	63	67	71	73	74
20-24	190	195	199	202	201	83	88	92	95	96
25-29	213	218	223	226	225	92	97	102	106	107
30-34	229	234	240	243	242	104	109	114	118	120
35-39	322	330	337	341	340	132	139	146	150	152
40-44	421	431	440	444	444	160	168	176	181	183
45-49	630	643	656	662	661	237	247	258	266	269
50-54	877	894	910	918	917	349	364	379	389	393
55-59	1361	1385	1408	1420	1418	582	604	627	643	648
60-64	1910	1940	1969	1985	1982	918	949	981	1004	1012
65-69	3161	3205	3247	3269	3266	1791	1844	1897	1935	1948
70-74	4893	4953	5009	5039	5034	3252	3333	3415	3473	3493
75+	12182	12286	12384	12435	12426	10254	10423	10592	10711	10753

* The above forecasts were made using the VAR models; (Rounded off to the nearest integer)

ISBN-1391-4987

Errors of Forecasts

In Table 4.6 above, one can compute approximate 95% confidence intervals for the forecasts of sex-specific death rates using LC approach as follows: Lower Bound = $(est.)e^{-b_x(m.e)}$; and Upper Bound = $(est.)e^{b_x(m.e)}$, where '*est.*' is the forecast given in Table 4.6; b_x is the corresponding entry read from Table 4.1; and '*m.e*' is the margin of error obtained from Table 4.4 or 4.5.

For example, approximate 95% confidence interval for the male death rate forecast in 2010 for the age group 0 – 4 is computed as follows:

'*est*' = 782; b_x = 0.05897; 'm.e' = 4.169. Thus, 95% confidence interval is: $(782e^{-(0.05897)(4.169)}, 782e^{(.05897)(4.169)})$ = (611, 999).

3. Conclusions

This study was mainly focused on modeling and forecasting mortality rates using Sri Lankan data and generating sex-specific life tables. Mortality was modeled using the Lee-Carter (1992) approach.

The number of registered deaths reveals historical episodes. There are peaks recorded in 1974, 1989, 1996 and 2005, due to two civil conflicts, the civil war and the tsunami incident, respectively. There are four peak points in 1974, 1989, 1996 and 2005 in the male graph, but only two peak points in the female graph in 1974 and 2005.

Out of several alternative time series models examined for modeling male and female mortality in Sri Lanka, a bivariate vector autoregressive (VAR) model performed better than the univariate models in this study, and produced relatively better forecasts.

From the 2006-2008 life table generated using estimated k_t values as input, some important information was evident. The life expectancy at birth for males was 70.3 years, and 76.8 for females. The probability of a male aged 55 years surviving another 20 years was 0.71, where as it was 0.85 for the females. Generally there was a downward trend in the mortality index for both sexes during the period 1950-2000 in agreement with similar studies elsewhere; but there was an upward trend from 2000-2008, and as a result the 30-year forecasts showed a slightly upward trend in both cases. However, the lower bounds of 95% confidence intervals for the forecasts showed a downward trend indicating a possibility for the actual forecasts to follow a downward trend.

There were certain limitations too, in this study: Data on age–specific mortality rates were incomplete as the population size and the number of deaths for ages above 75 and separate data for age groups 0 – 1 and 1 – 4 were not available in

Sri Lanka. The gaps in the data had to be estimated based on assumptions discussed earlier. Also, the time series models were fitted using only 59 data points available. Therefore, all data points were used for model fitting without keeping any as a test set.
This study will be further refined and extended in future work, with special emphasis on missing value estimation and outlier analysis.

References

1. Abeykoon A. T. P. L. (1996). "Demographic Implications of Health Care in Sri Lanka"*Asia-Pacific Population Journal*, 11(2), 47-58.
2. Borsch A. (2009). "Many Roads to Rome: Varieties of funded Pensions in Europe and Asia", *Pensions: An international Journal*, 14(3), 172.
3. Central Bank of Sri Lanka (2009). Annual Report.
4. De Silva I. W. (1994a). "Ahead of Target: Achievement of Replacement Level Fertility in Sri Lanka before the Year 2000", *Asia-Pacific Population Journal*, 9(4), 3-22.
5. De Silva I. W. (1994b). "How Serious is Ageing in Sri Lanka and What can be Done about it?",*Asia-Pacific Population Journal*, 9(1), 19-36.
6. De Silva I. W. (2005). "Family Transition in South Asia: Provision of Social Service and Social Protection", *Asia-Pacific Population Journal*, 13-46.
7. De Silva I. W. (2007). "Beyond Twenty Million: Projecting the Population of Sri Lanka 2001-2081", Research Studies: Demographic Transition and Pension Series No. 6, Institute of Policy Studies, Colombo, Sri Lanka.
8. Department of Census and Statistics, Sri Lanka (2010). *Statistical Abstract*, www.statistics.gov.lk.
9. Dickey D. and Fuller W. A. (1979). "Distribution of the Estimates of Autoregressive Time Series with a Unit Root", *Journal of the American Statistical Association* 74 (June) 427-31.
10. Gamaniratne N. (2007). "Population Ageing, Policy Responses, and Options to Extend Retirement Coverage: Case Study of Sri Lanka" Research Studies: Demographic Transition and Pension Series No. 7, Institute of Policy Studies, Colombo, Sri Lanka.

11. Good I. J (1969). “Some Applications of the Singular Decomposition of a Matrix”, *Technometrics*, 11(4), 823-831. DOI: 10.1080/00401706.1969.10490741

12. Johnson R. and Wichern D. (2002). Applied Multivariate Statistical Analysis Prentice Hall Press.

13. Lee R. D. and Carter L. R. (1992). “Modeling and Forecasting U.S. Mortality”, *Journal of the American Statistical Association*, 87, 659-671.

14. Lee R.D. and Miller T. (2001). “Evaluating the Performance of the Lee-Carter Method for Forecasting Mortality”. *Demography*, 28, 537-549. DOI: 10.1353/dem.2001.0036

15. Li N. and Lee R. (2005). “Coherent Mortality Forecasts for a Group of Populations: An Extension of the Lee-Carter Method”. *Demography,* 42, 575-594. DOI: 10.1353/dem.2005.0021.

16. Li N., Lee R., and Tuljapurkar S. (2004). “Using the Lee-Carter Method to Forecast Mortality for Populations with Limited Data”, *International Statistical Review,* 72(1), 19-34. DOI: 10.1111/j.1751-5823.2004.tb00221.x.

17. Li J.S.H. and Chan W.S. (2007). “The Lee-Carter Model for Forecasting Mortality Revisited” *North American Actuarial Journal,* 11, 68-89. DOI: 10.1080/10920277.2007.10597438.

18. Li J.S.H., Chan W.S., and Cheung S.H. (2011). “Structural Changes in the Lee-Carter Mortality Indexes: Detection and Implications”. *North American Actuarial Journal*, 15, 13-31. DOI: 10.1080/10920277.2011.10597607.

19. Lütkepohl, H. (2005). New introduction to multiple time series analysis - Cambridge University Press.

20. Ostbye T., Chan A., Malhotra R. and Kothalawala J. (2010). “Adult children caring for their elderly parents”, *Asian population Studies*, 6(1), 83-97. DOI: 10.1080/17441731003603504

21. United Nations. (2009). *Department of Economic and Social Affairs,* Population Division – World Population Ageing 2009.

Appendix

Appendix 1

Date: 04/03/13 Time: 15:01
Sample (adjusted): 1953 2008
Included observations: 56 after adjustments
Trend assumption: Linear deterministic trend
Series: KTF KTM
Lags interval (in first differences): 1 to 2

Unrestricted Cointegration Rank Test (Trace)

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None	0.174972	12.43147	15.49471	0.1374
At most 1	0.029218	1.660559	3.841466	0.1975

Trace test indicates no cointegration at the 0.05 level
* denotes rejection of the hypothesis at the 0.05 level
**MacKinnon-Haug-Michelis (1999) p-values

Unrestricted Cointegration Rank Test (Maximum Eigenvalue)

Hypothesized No. of CE(s)	Eigenvalue	Max-Eigen Statistic	0.05 Critical Value	Prob.**
None	0.174972	10.77091	14.26460	0.1661
At most 1	0.029218	1.660559	3.841466	0.1975

Max-eigenvalue test indicates no cointegration at the 0.05 level
* denotes rejection of the hypothesis at the 0.05 level
**MacKinnon-Haug-Michelis (1999) p-values

Appendix 2

Vector Autoregression Estimates

Date: 04/02/13 Time: 13:16

Sample (adjusted): 1951 2008

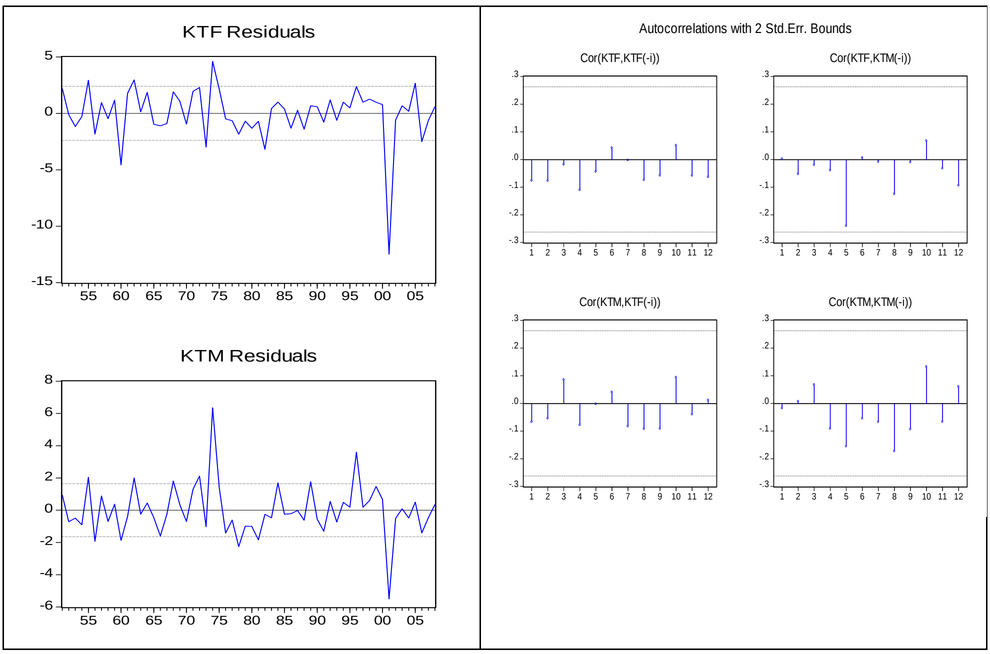
Included observations: 58 after adjustments

Standard errors in ( ) & t-statistics in [ ]

	KTF	KTM
KTF(-1)	1.059236	0.145376
	(0.10453)	(0.07150)
	[10.1334]	[2.03311]
KTM(-1)	-0.343156	0.498991
	(0.24454)	(0.16728)
	[-1.40326]	[2.98296]

C	-0.198084	0.055173
	(0.32601)	(0.22301)
	[-0.60760]	[0.24740]
R-squared	0.872755	0.672901
Adj. R-squared	0.868128	0.661007
Sum sq. resids	315.8837	147.8126
S.E. equation	2.396527	1.639360
F-statistic	188.6192	56.57251
Log likelihood	-131.4514	-109.4280
Akaike AIC	4.636256	3.876827
Schwarz SC	4.742831	3.983402
Mean dependent	0.067138	0.361897
S.D. dependent	6.599432	2.815652
Determinant resid covariance (dof adj.)		4.804374
Determinant resid covariance		4.320223
Log likelihood		-207.0328
Akaike information criterion		7.345958
Schwarz criterion		7.559107

Appendix 3



VAR Residual Portmanteau Tests for Autocorrelations
H0: no residual autocorrelations up to lag h
Date: 04/02/13 Time: 13:19
Sample: 1950 2032
Included observations: 58

Lags	Q-Stat	Prob.	Adj Q-Stat	Prob.	df
1	1.298190	NA*	1.320965	NA*	NA*
2	2.809366	0.5902	2.886112	0.5771	4
3	4.722088	0.7868	4.903164	0.7679	8
4	8.938492	0.7082	9.431894	0.6657	12
5	17.15056	0.3759	18.41869	0.3000	16
6	19.94141	0.4616	21.53155	0.3665	20
7	21.20482	0.6266	22.96838	0.5217	24
8	23.57777	0.7036	25.72099	0.5884	28
9	25.86684	0.7695	28.43051	0.6479	32
10	27.29328	0.8514	30.15413	0.7423	36
11	28.90659	0.9035	32.14502	0.8071	40
12	34.59551	0.8442	39.31800	0.6723	44

*The test is valid only for lags larger than the VAR lag order.
df is degrees of freedom for (approximate) chi-square distribution

Roots of Characteristic Polynomial
Endogenous variables: KTF KTM
Exogenous variables: C
Lag specification: 1 1
Date: 04/02/13 Time: 13:20

Root	Modulus
0.948177	0.948177
0.610051	0.610051

No root lies outside the unit circle.
VAR satisfies the stability condition.