MATEMATIKA: MJIAM, Special Issue, December 2019, 65–77

Comparison of Lee Carter Model and Cairns, Blake and Dowd Model in Forecasting Malaysian Higher Age Mortality

Haneef Zulkifle, Fadhilah Yusof* and Siti Rohani Mohd Nor

Department of Mathematical Sciences, Universiti Teknologi Malaysia
81310 UTM Johor Bahru, Malaysia

*Corresponding author: fadhilahy@utm.my

Article history
Received: 10 November 2019
Received in revised form: 21 November 2019
Accepted: 19 December 2019
Published online: 31 December 2019

Abstract Demographers and actuaries are very much conscious of the trend of mortality in their own country or in the world in general. This is because mortality is the basis for longevity risk evaluation. Mortality is showing a declining trend and it is expected to further decline in the future. This will lead to continuous increase in life expectancy. Several stochastic models have been developed throughout the years to capture mortality and its variability. This includes Lee Carter (LC) model which has been extended by various researchers. This paper will be focusing on comparing LC model and another mortality model proposed by Cairns, Blake and Dowd (CBD). The LC uses the log of central rate of mortality and CBD uses logit of the mortality odds as dependent variable. Analysis of comparison is done using a few techniques including Akaike information criteria (AIC) and Bayesian information criterion (BIC). From the overall results, there is no model better than the other in every aspect tested. We illustrate this via visual inspection and in sample and outof sample analysis using Malaysian mortality data from 1980 to 2017.

Keywords Lee Carter Mortality Model, Cairns, Blake and Dowd Mortality Model, Malaysia Mortality Modelling

Mathematics Subject Classification 91B70, 62P05

1 Introduction

In recent times, there has been a growing interest in the area of mortality by demographers. The death statistics collected by the Department of Statistics is usually used to monitor mortality in a country. There are different kind of death statistics specifically for different purpose. Some of them includes death rate, infant mortality, age specific death rate and others. These data provide fundamental information about the population. Monitoring mortality is useful for the authorities. It can be used as a basis for planning retirement schemes, pricing of life insurance and annuity products or to manage public health. For example, infant mortality data can give indication of maternity health and wellbeing, mortality between races can show disparity among health service that each community receive and higher age mortality can provide information on how life insurance and annuity products can be priced.

Given the increase in mortality's relevance to the government plan and social security, the mortality forecasting area is ever developing.Medical technology advancement in recent decade has improved human's health generally. This led to prolonged lifespan and rise in life expectancy in most countries. Overall health status and condition measures can be summarized with the indicator of mortality rates and life expectancy. According to Department of Statistics, the life expectancy of a Malaysian newborn in 2018 has increased by 2.8 years from 72.2 years in 2000 to 75.0 years in 2018.

One of the most popular extrapolative method was introduced by Lee Carter, (LC) in 1992. The LC model excel in terms of simplicity and robustness in explaining linear trends in age-specific death rates [1]. To be specific, LC model uses a log-bilinear form for age-specific mortality which comprises of two factor that is time and age. Then, a matrix of singular value decomposition is utilized to extract a time-varying mortality index. This time varying index will later be modeled and forecasted via random walk with drift (a time series model). LC model does not include the underlying factor that could influence mortality changes like medical background, behavioral or social status of a population [2]. However, LC model has proven to be successful in forecasting mortality of many countries, but some countries do have problems like incomplete data or short time series data. This opens up opportunities for modifications, extension or new models to be introduced by other researchers.

One different model introduced was Cairns, Blake and Dowd (CBD) model. Nocito [3] highlighted that both of these models represent two distinct parametric mortality model. The LC uses the log of central rate of mortality and CBD uses logit of the mortality odds as dependent variable.CBD model was developed on the base of special case known as Perks model. The model fits well with the England and Wales data at higher age. The authors, Cairns, Blake and Dowd focused on modeling higher mortality because they realise that improvement in mortality has impacted the higher ages population financially. The unpredictable nature of mortality improvement meant that either pensioners outlive their saving plus life annuity or pension plan sponsor having to overpay because of under projected life expectancy of a person. The model has been accepted widely and this has enabled the formulation of pension funds, life insurance and private annuity for concerning companies [4]. It is also used for pricing longevity bonds [5].

In Malaysia, the age for pension for government and non-government worker has been changed three times. Starting pension age was 56 which has been changed to 58 in 2008 and again changed to 60 years in 2012. Longevity risk is a big factor faced by any retiree. There will always be a risk that a person's saving will diminish before his death. As a consequence, there is a need to model and forecast the population of higher ages in Malaysia. For this purpose, Lee Carter and Cairns, Blake and Dowd model will be used. Comparison will be made between the models and the best model that suits higher age Malaysia mortality data will be determined using goodness of fit test and forecast performance.

As far as the author knows, research on mortality in Malaysia is still limited especially one that focuses on comparing and forecasting higher age mortality. As such, the purpose of this study is to compare mortality models which are most suited to forecast higher ages or near pension ages mortality. The age in discussion will be 55 years until 80, as data permitted. The models that will be compared will be the mortality model introduced by Lee Carter in 1992 and Cairns, Blake and Dowd mortality model introduced in 2006.

2 Methodology

In this section, we will discuss the Lee Carter and Cairns, Blake and Dowd mortality model using Malaysia mortality data from 1980 to 2017. Data was obtained from Department of Statistics Malaysia. The data was separated into two parts. Mortality data from 1980 to 2010 was used in sample fit and data from 2011 to 2017 was used for out of sample fit. The models were separately fitted to male and female population of the higher ages, between age 55 to 80.

2.1 Lee Carter Mortality Model (LC)

The Lee Carter Mortality Model was introduced in 1992. It is based on a two factors model which are age and time. The two factors are described as having a linear relationship with the logarithmic of central death rate. With its simplicity and robustness, it manages to explain variability in mortality and produce good forecast. The Lee Carter model can be defined as follows:

$$\log\left(m_{x,t}\right) = a_x + b_x k_t + \varepsilon_{x,t} \tag{1}$$

Where,

$$\begin{aligned} m_{x,t} &= \text{central death rate at age } x \text{ and year } t \\ a_x &= \text{average age-specific mortality} \\ b_x &= \text{age component which is relative speed of change of each age} \\ k_t &= \text{mortality index at time } t \\ \varepsilon_{x,t} &= \text{error term at age } x \text{ and time } t \text{ with mean 0 and variance } \sigma^2 \end{aligned}$$

To find unique solution, Lee and Carter impose some constraints which are $\sum_t k_t = 0$ and $\sum_x b_x = 1$. To estimate the parameters, we adopted Maximum Likelihood Estimation (MLE) to the Poisson distribution for the number of deaths

$$D_{x,t} = Poisson\left(\lambda_{x,t}\right) \text{ with } \lambda_{x,t} = E_{x,t}\mu_{x,t} \text{ and } \mu_{x,t} = e^{a_x + b_x k_t},$$

where, $D_{x,t}$ is the number of death, $E_{x,t}$ is the number of exposure at age x and time t. The likelihood function is expressed as

$$L_{x,t}\left(\theta; d_{x,t}\right) = \frac{\lambda_{x,t}^{d_{x,t}} e^{-\lambda_{x,t}}}{d_{x,t}!} \tag{2}$$

So, the log likelihood function is

$$\begin{aligned} \ln L_{x,t}\left(\theta; d_{x,t}\right) &= \ln\left(\frac{\lambda_{x,t}^{d_{x,t}} e^{-\lambda_{x,t}}}{d_{x,t}!}\right) \\ &= \ln\left(\lambda_{x,t}^{d_{x,t}} e^{-\lambda_{x,t}}\right) - \ln\left(d_{x,t}!\right) \\ &= d_{x,t}\ln\left(\lambda_{x,t}\right) - \lambda_{x,t} - \ln\left(d_{x,t}!\right) \end{aligned} \tag{3}$$

The joint log likelihood can be expressed as follows and will be maximised as with respect to a_x, b_x and k_t.

$$l(\theta) = \sum_{x,t} \left(d_{x,t} \ln(\lambda_{x,t}) - \lambda_{x,t} - \ln(d_{x,t}!)\right)$$
$$= \sum_{x,t} \left(d_{x,t} \ln(E_{x,t}\mu_{x,t}) - E_{x,t}\mu_{x,t} - \ln(d_{x,t}!)\right) \quad (4)$$

Then, the time series index k_t was modelled using ARIMA process. This will allow projection of the mortality. Lee and Carter suggested a random walk with drift to model k_t in which

$$k_t = k_{t-1} + d + \varepsilon_t, \quad (5)$$

where d is the drift parameter, and ε_t is error term.

2.2 Cairns, Blake and Dowd Mortality Model (CBD)

The Cairns, Blake and Dowd model was introduced in 2006. It was built with the assumption that age, period and cohort effects are different in nature, in which there exist randomness between each year [6]. This model was designed specifically to cater for higher age mortality. A similarity of this model with Lee Carter model is that, they use extrapolation method to project the parameters into the future. However, in contrast, Cairns, Blake and Dowd model assumes smoothness in age, and at the same time allowing period effect [6].

The CBD model can be expressed as

$$\log\left(\frac{q_{x,t}}{1-q_{x,t}}\right) = k_t^1 + k_t^2(x - \overline{x})$$
$$\log\text{it}(q_{x,t}) = k_t^1 + k_t^2(x - \overline{x}) \quad (6)$$

where,

$q_{x,t}$ = Probability that a person age x, will die at time t

$\overline{x}$ = Average of the age range under consideration

k_t^1, k_t^2 = the period mortality indexes (mean and variance of mortality rates for a period).

To estimate the parameters, CBD model uses the same log likelihood as for LC, but since LC model uses forces of mortality or central death rate $\mu_{x,t}$ and CBD model uses mortality rate $q_{x,t}$,some modificationswere needed for a valid comparison between these models. Cairns *et al.* [6] suggested two assumptions which are given as

1. The central death rate is constant over every year of age (that is from age x to $x+1$) and over a calendar year.
2. The population is stationary.

The implication of the assumptions are

1. $m_{t,x} = \mu_{t,x}$,and
2. $q_{t,x} = 1 - e^{-\mu_{t,x}} = 1 - e^{-m_{t,x}}$

So, we could define $\mu_{t,x} = -\log(1 - q_{t,x})$ and use the same log likelihood as for LC for parameter estimation. To implement these mortality model, a package in R developed by Villegas *et. al* [7], StMoMo was used.

2.3 Goodness of Fit

Goodness of fit for a model is the measure of deviation between fitted and observed data. There are a couple of techniques that can be utilized to compare the goodness of fit between models. In this paper, we will focus on four techniques including the analysis of residuals, Akaike Information Criterion (AIC), Bayesian Information Criterion (AIC), root mean squared error (RMSE) and mean absolute percentage error (MAPE).

1. Residual analysis involves pattern observation in the residual heatmap. Regular patterns in the residuals indicate the inability of the model to describe all the features of the data appropriately. From the heatmap, we can see whether there is cohort or age effect and see whether a model does fit the data well.

2. Bayesian Information Criterion is used as a way to find the balance between lack of model fit and its complexity. It can be defined as

$$BIC = -2 \ln L + 2 \ln Nk \tag{7}$$

   where N is the number of samples for the analysis, L is the log likelihood and k is number of parameters.

3. By definition, Akaike Information Criterion (AIC) is the negative logarithm of likelihood of the model, adjusted according to number of model's parameter. Together with BIC, they can be used to determine which two rate function is better fit the data set [8]. The formula is given as

$$AIC = -2 \ln L + 2k \tag{8}$$

   where L is the log likelihood, and k is the number of parameters. The smaller the value of AIC, the better the model will fit the data.

4. Root Mean Square Error (RMSE) measures the standard deviation between observed and forecasted value and degree of coincidence among them [9]. It can be expressed as

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (p_i - o_i)^2}{N}} \tag{9}$$

   where,

$$\begin{aligned} p &= \text{projected or forecasted value} \\ o &= \text{observed or actual value} \\ N &= \text{number of samples} \end{aligned}$$

5. Mean Absolute Percentage Error (MAPE) measures how far is the difference between actual and forecasted value. The smaller this value, the closer the distance between actual and forecast value. The formula is given as

$$\frac{1}{N} \sum_{i=1} \left| \frac{o_i - p_i}{o_i} \right| \times 100 \tag{10}$$

3 Results and Discussion

3.1 In Sample Fit for Lee Carter (LC) and Cairns, Blake and Dowd (CBD) Comparison

3.1.1 Parameter Estimates

As mentioned before, the analysis was done using mortality data that was separated into two parts. Mortality data from 1980-2010 was used for in sample fit. Some classify in sample as training set. Another set of data from 2011-2017 was used for out of sample fit, also known as validation set.

Looking at the parameters LC plot, from Figure 1 and 2, a_x describes the shape of age of the death rates m_x, so it is always directly proportional to age. The parameter b_x explains how mortality at age x changes with respect to k_t. In both male and female, b_x takes positive values indicating declining of mortality at all ages under consideration. In the same figure, it is shown that k_t declined throughout the years indicating an improved mortality with a few spikes. In particular, there is a spike in 1998-1999. This could be caused by the emergence of death related disease in Malaysia. In 1998, there was an outbreak of Nipah virus disease and many deaths are recorded at the time. From the lower k_t values in Figure 4 compared to Figure 3, we can see that mortality improvement is much faster in female as compared to male.

The $k_t^{(1)}$ values for CBD in Figure 3 and 4 also shows the same declining trend. The same spike trending 1998-1999 can also be seen from the values $k_t^{(1)}$ from the CBD parameters. The parameter $k_t^{(2)}$indicates the steepness of the improvement at any age.

3.2 Analysis of Good Fit

3.2.1 Residuals Analysis

For the model to be good fit, the residuals should be distributed randomly across ages and years. This can be easily examined by the help of heatmap. A darker colored pixel on the heatmap represents residual that is further away from the fitted line while a lighter colored pixel indicates residual closer to the fitted line. A recognizable pattern of clump or diagonal pattern suggests that the model failed to account for certain age, period or cohort effect in the data. From the deviance residuals heatmap shown in Figure 5, it can be seen that LC and CBD model for male and female show some patterns. Diagonal pattern can be seen for all models except for CBD female. This indicate that both LC and CBD models fails to capture cohort effect for those population. For CBD female, the pattern does exist but only slightly. When examined across age, the heatmap shows that LC and CBD models cannot fit well for both sexes aged 75 and 80. This can be seen by the darker colour of the heatmap.

3.2.2 AIC and BIC Comparison

AIC and BIC values has been used fairly widely to determine model fit. It is suggested that the lower the value, the better is the model to fit the data. Table 1 shows the number of parameters involved in the model and also the values of log likelihood from each model to form the AIC and BIC value. From Table 2, it shows that for both sexes, the values of AIC and BIC for CBD model is lower. The AIC and BIC values are relative values compared between models.

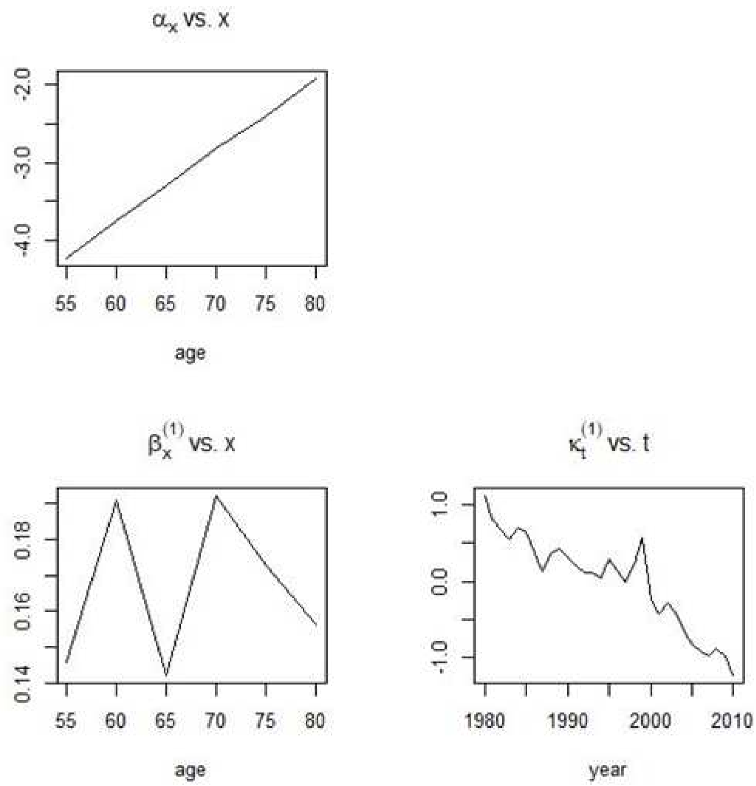

Figure 1: Parameters for the Lee-Carter (LC) Model Fitted to Malaysia Male Population for Ages 55–80 and the Period 1980–2010

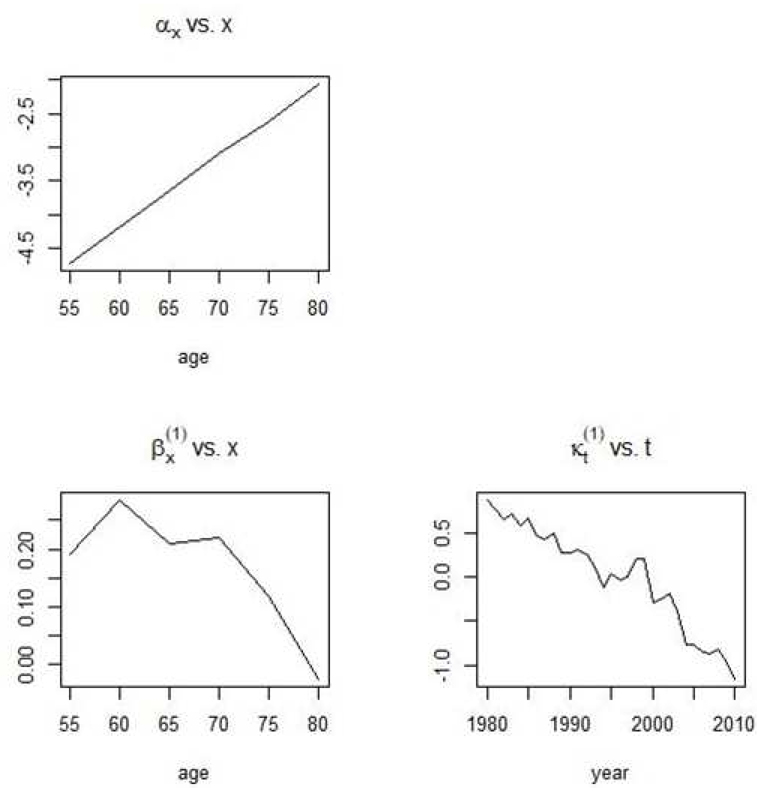

Figure 2: Parameters for the Lee-Carter (LC) Model Fitted to Malaysia Female Population for Ages 55–80 and the Period 1980–2010

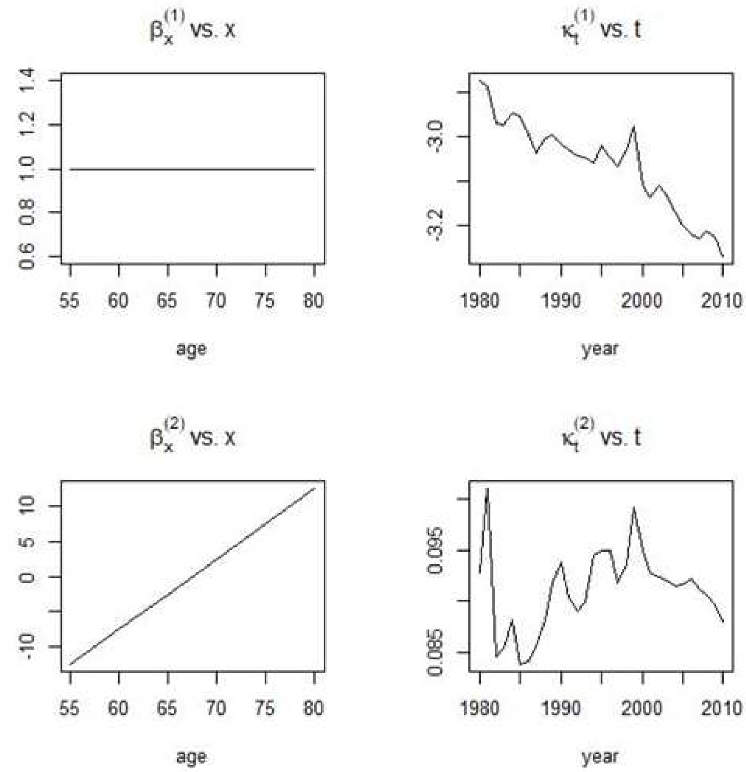

Figure 3: Parameters for the Cairns, Blake and Dowd (CBD) Model Fitted to Malaysia Male Population for Ages 55–80 and the Period 1980–2010

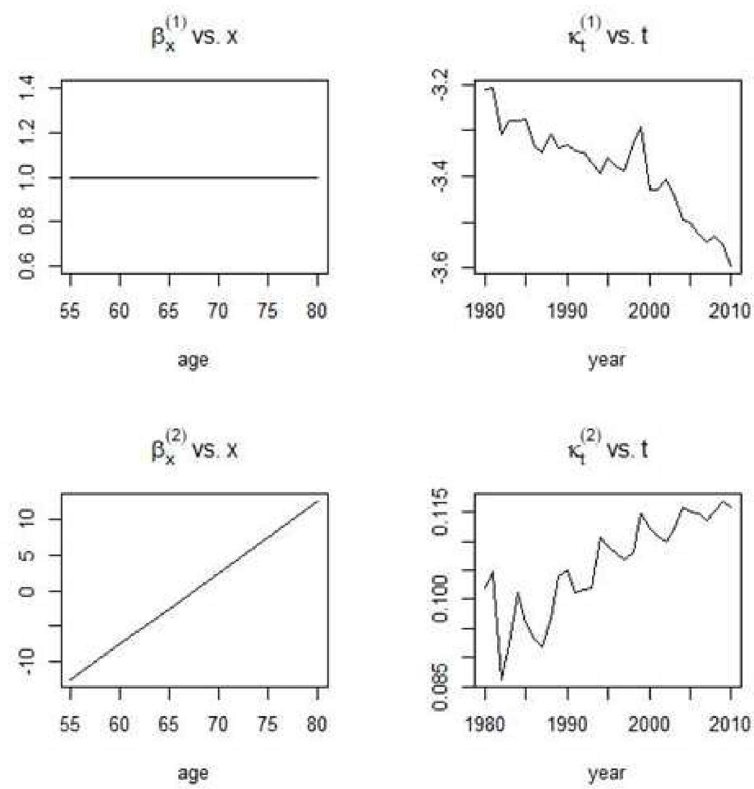

Figure 4: Parameters for the Cairns, Blake and Dowd (CBD) Model Fitted to Malaysia Female Population for Ages 55–80 and the Period 1980–2010

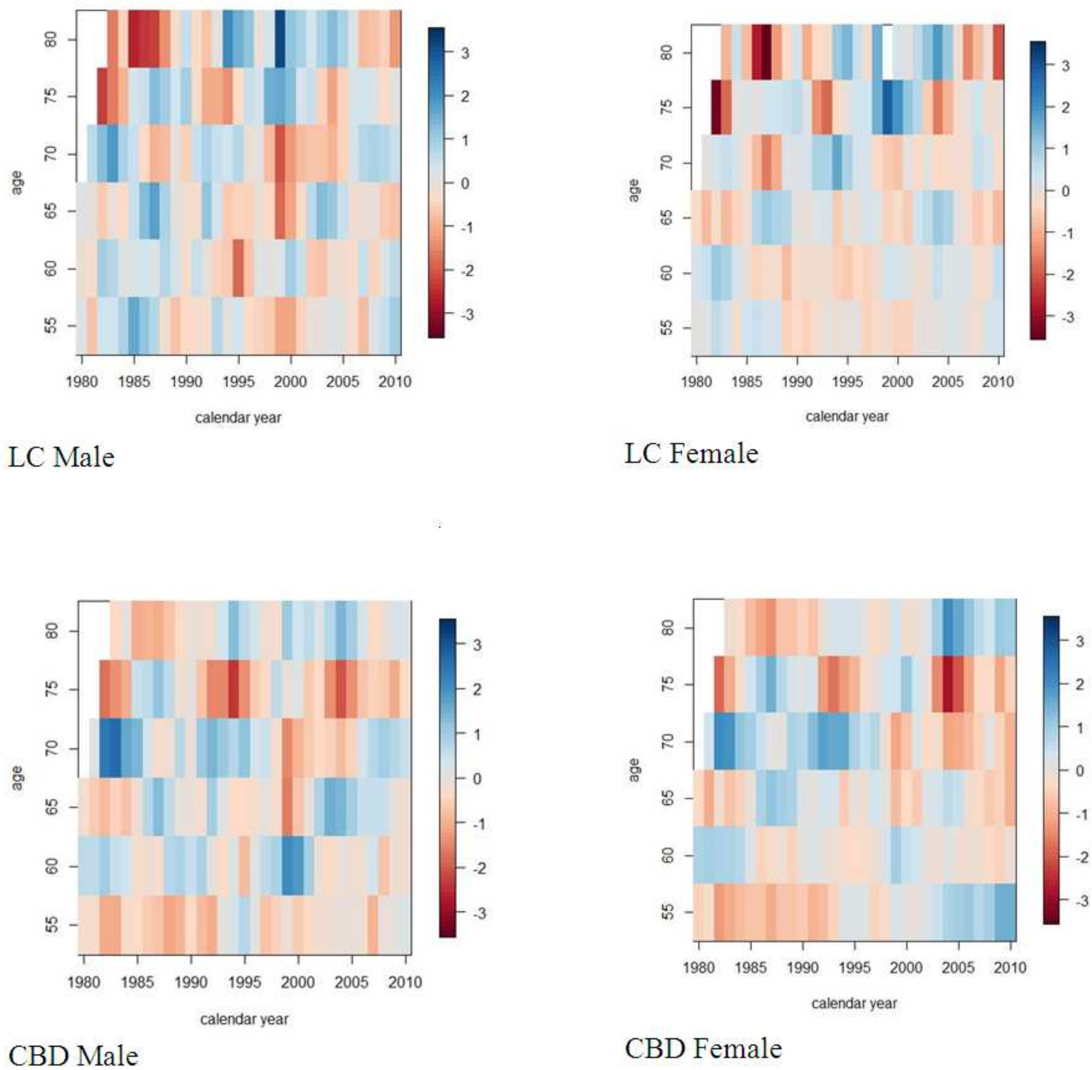

Figure 5: Heatmaps of Deviance Residuals for Different Models Fitted to Malaysia Population for Ages 55–80 and the Period 1980–2010. The White Cells Indicate the Cell is Not Included in Analysis

It does not simply mean one model is superior than the other, but rather it means, one model is more likely to be better model than the other. From this analysis, CBD model is seen to be the better model for this data. However, there will be no model that perfectly fit a dataset, so we have to look at other goodness of fit technique as well.

3.2.3 RMSE and MAPE Comparison

From Table 3, examining through the errors from the model to the fitted data, LC model shows lower error for both sexes as compared to CBD model be it RMSE or MAPE. The lower values come from the difference between the fitted data to the actual data. A better model can potentially fit a data that is really close to the actual data. From this RMSE and MAPE analysis, it indicates that LC model fit the data better than CBD model.

Table 1: Number of Parameters Estimated and Log Likelihood for LC and CBD Model Fitted to Malaysia Population for Ages 55-80 and Period 1980-2010

	Number of parameters		Log likelihood	
Model	Male	Female	Male	Female
Lee Carter	41	41	-1837.172	-2268.053
Cairns, Blake and Dowd	62	62	-1731.569	-2159.213

Table 2: AIC and BIC for LC and CBD Model Fitted to Malaysia Population for Ages 55-80 and Period 1980-2010

	Male		Female	
Model	AIC	BIC	AIC	BIC
Lee Carter	3756.344	3887.255	4618.106	4749.018
Cairns, Blake and Dowd	**3587.138**	**3785.101**	**4442.426**	**4640.389**

Table 3: RMSE and MAPE for LC and CBD Model Fitted to Malaysia Population for Ages 55-80 and Period 1980-2010

	Male		Female	
Model	RMSE (%)	MAPE (%)	RMSE (%)	MAPE (%)
Lee Carter	**7.2**	**0.019**	**7.3**	**0.017**
Cairns, Blake and Dowd	7.9	0.021	8.2	0.020

3.3 Out of Sample Forecast for Lee Carter (LC) and Cairns, Blake and Dowd (CBD) Comparison

We projected mortality for both model to the year 2020, 2030 and 2040 to see the improvements over the years. In Figure 6, for male, LC show more mortality improvements for all ages except age 65 and 80. The lines show convergence for both models with a slight divergence at age 60,65 and 80.

In Figure 7, for female, LC model projected better improvements as well across the ages. The lines show divergence for most part except for the age 65 and 75. CBD model only predicted better mortality improvement at ages 55 and 80.

Another way to compare models is using out of sample forecast or projection fit of the models. D'amato *et al.* [10] suggested that it is possible that a model fits well with historical data, but produce inaccurate forecasts. For the out of sample comparison, Malaysia mortality data from year 2011 to 2017 was used. The data was compared to seven years projected male and female LC and CBD model result. The RMSE and MAPE for the projection is shown in Table 4. We can see that RMSE and MAPE for CBD is lower for male as compare to LC but the result for female shows otherwise. This result is inconsistent with results obtained in the in sample fit analysis in terms of RMSE and MAPE. This case is similar to what D'amato *et al.* [10] had suggested.

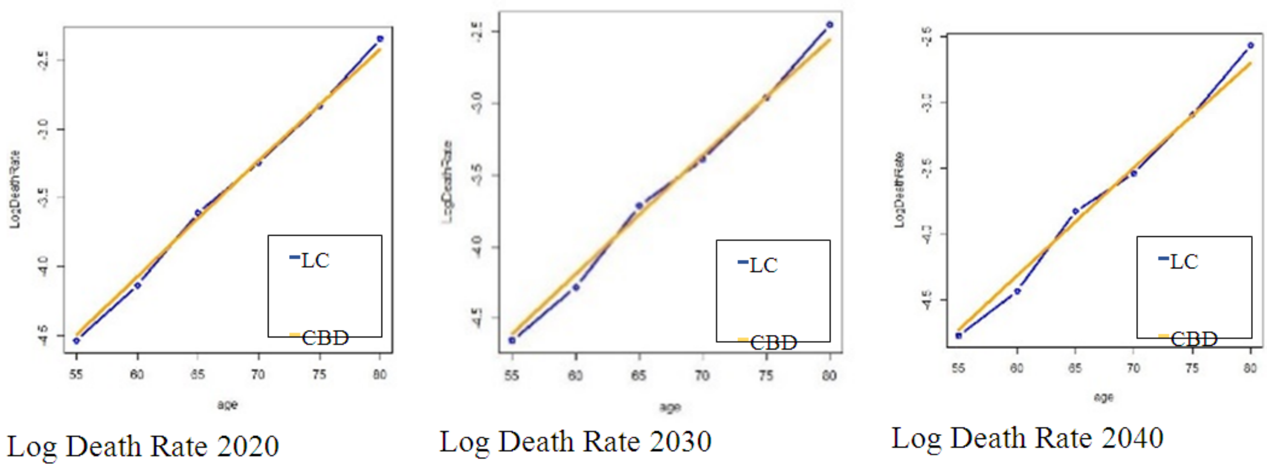

Figure 6: Comparison of Mortality Trend LC vs CBD for Male. Projection was done to 2020, 2030 and 2040. Blue line represents LC and yellow represents CBD mortality

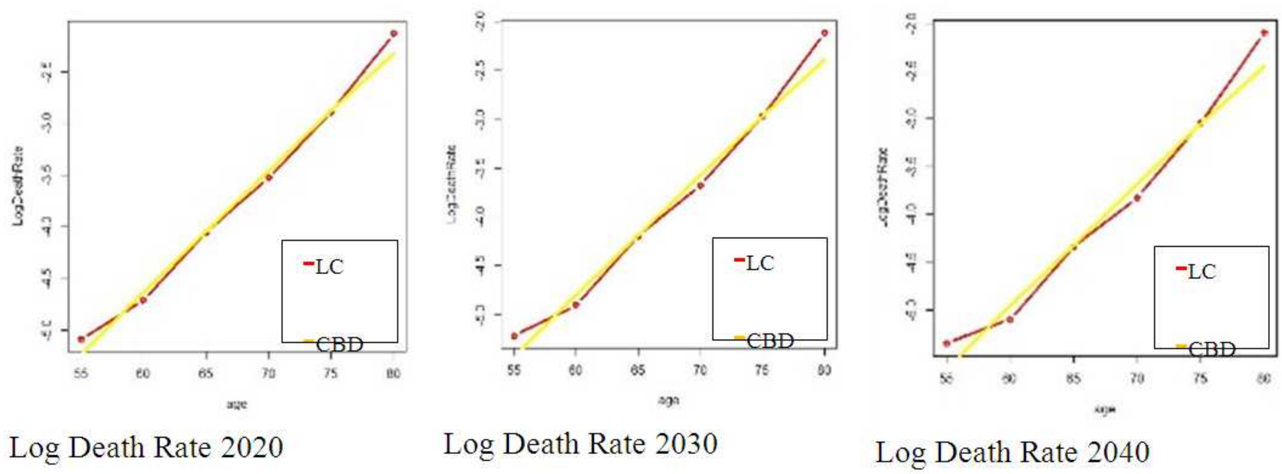

Figure 7: Comparison of Mortality Trend LC vs CBD for Female. Projection was done to 2020, 2030 and 2040. Red line represents LC and orange line represents CBD mortality

Table 4: RMSE and MAPE for LC and CBD Model 6 Years' Projection for Malaysia Population for Ages 55-80 and Period 2011-2017

	Male		Female	
Model	RMSE (%)	MAPE (%)	RMSE (%)	MAPE (%)
Lee Carter	0.438	0.072	**0.441**	**0.084**
Cairns, Blake and Dowd	**0.434**	**0.066**	0.910	0.129

4 Conclusion

This paper compares the mortality model of Lee Carter and Cairns, Blake and Dowd model to Malaysian mortality. The focus is to get the best model to fit and project mortality of the higher ages for male and female, in particular ages 55 and above, up to 80. The parameters estimate was discussed and each model was fitted to the specified data set. There were some techniques used to determine and compare the goodness of fit and also projection accuracy.

The results of AIC and BIC showed that Cairns, Blake and Dowd model relatively did fit data better than Lee Carter model in both sexes, but produces slightly larger error. There were cohort effects noted through heatmap of residual in all dataset except female CBD model. The residuals also show that both LC and CBD models cannot fit well the data for top higher ages especially 70-80 years.

The forecasting results showed interesting findings. First, LC models predict faster mortality improvements than CBD across all sexes and projection point. Second, LC model projected mortality more accurately for female and CBD did projected accurately for male although LC does seem fit the data better historically through the in sample analysis. This is in line with few suggestions from other researcher such as Armstrong [11] that a model that fits a data might not produce more accurate forecast. Final conclusion is that by looking at several techniques applied, not one model performs better than the other in every aspect. Both model show good fit and forecast accuracy over different section of the sample period and age range. However, since we note that there is a cohort and age effect, future analysis could consider using modified model that incorporates cohort and age effect. The analysis also shows that in future research, models must be chosen wisely and according to data available, in order to have good fit and accurate forecast.

5 Acknowledgments

The authors would like to Ministry of Higher Education Malaysia under Fundamental Research Grant Scheme (FRGS/1/2018/STG06/UTM/02/2), with vot R.J130000.7854.5F015. The authors also would like to thank all parties that contributed to this paper especially the Department of Statistics Malaysia (DOSM) for all the data obtained.

References

[1] Zakiyatussariroh, W. W., Said, Z. M., & Norazan, M. R. Evaluating the performance of the Lee-Carter method and its variants in modelling and forecasting Malaysian mortality. In *AIP Conference Proceedings (Vol. 1635, No. 1, pp. 762-769). AIP.* 2014, December.

[2] Shair, S. N., Rosmizan, M. A. S., Ting, M. J. M. S., & Zaini, M. A. A. Projected Malaysian Lifetable: Evaluations of The Lee-Carter and Poisson Log-Bilinear Models. *International Journal.* 2018. 1(4): 60-72.

[3] Nocito, S. *Stochastic Mortality Projections: A Comparison of the Lee-Carter and the Cairns-Blake-Dowd models Using Italian Data. University of Studies of Turin.* 2015.

[4] Maccheroni, C., & Nocito, S. Back testing the Lee–Carter and the Cairns–Blake–Dowd stochastic mortality models on Italian death rates. *Risks.* 2017. 5(3): 34.

[5] Cairns, A. J., Blake, D., & Dowd, K. A two-factor model for stochastic mortality with parameter uncertainty: theory and calibration. *Journal of Risk and Insurance.* 2006. 73(4): 687-718.

[6] Cairns, A. J., Blake, D., Dowd, K., Coughlan, G. D., Epstein, D., Ong, A., & Balevich, I. A quantitative comparison of stochastic mortality models using data from England and Wales and the United States. *North American Actuarial Journal.* 2009. 13(1): 1-35.

[7] Villegas, A., Kaishev, V. K. & Millossovich, P. StMoMo: An R package for stochastic mortality modeling. *Journal of Statistical Software* 2018. 84(3): 1–38.

[8] Kisslinger, C. After shocks and fault-zone properties. *Advances in Geophysics.* 1996. 38: 1-36.

[9] Farina, R., Di Bene, C., Piccini, C., Marchetti, A., Troccoli, A., & Francaviglia, R. Do Crop Rotations Improve the Adaptation of Agricultural Systems to Climate Change? A Modeling Approach to Predict the Effect of Durum Wheat-Based Rotations on Soil Organic Carbon and Nitrogen. In *Soil Management and Climate Change*. Academic Press. 2018. 221-236.

[10] D'amato, V., Gabriella, P. and Maria, R. The Mortality of the Italian Population: Smoothing Techniques On the Lee–Carter Model. *The Annals of Applied Statistics.* 2011. 5(2A): 705–724.

[11] Armstrong, J. S. (Ed.). *Principles of Forecasting: A Handbook for Researchers and Practitioners (Vol. 30).* Springer Science & Business Media. 2001.