MATEMATIKA, MJIM, Volume 39, Number 3, 213–226

Multi-Population O'Hare with ARIMA, ARIMA-GARCH and ANN in Forecasting Mortality Rate

Nurul Syuhada Samsudin and Siti Rohani Mohd Nor*
Department of Mathematical Sciences, Faculty of Science, Universiti Teknologi Malaysia
81310 Johor Bahru, Johor
*Corresponding author: sitirohani@utm.my

Article history
Received: 15 May 2023
Received in revised form: 15 October 2023
Accepted: 31 October 2023
Published on line: 1 December 2023

Abstract Multi-population mortality model has gained attention from prominent researchers of mortality due to its ability to provide biologically reasonable forecast. Previously, many researchers have proposed several multi-population stochastic mortality models that they considered adequate to produce accurate life expectancy. However, little have been addressed of the variability in full ages and time, which can contribute to an erroneous estimation of life expectancy. Therefore, this study proposed a new multi-population O'Hare with ARIMA, ARIMA-GARCH and ANN in forecasting the mortality rate for male and female in Malaysia, Taiwan, Japan, Hong Kong, Australia, USA, UK, Canada, and Switzerland. Multi-population O'Hare was used as a reference model, whilst ARIMA, ARIMA-GARCH and ANN were incorporated to the reference model to forecast the mortality rates. The adequacy of the proposed model was assessed by using measurement errors which were Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE). The results showed by multi-population O'Hare with ARIMA-GARCH gave the best forecasting performance for Taiwan, Japan, Australia, USA, UK, Canada, and Switzerland. On the other hand, multi-population O'Hare with ARIMA gave the best forecasting performance for Malaysia, whereas multi-population O'Hare with ANN gave the best forecasting performance for Hong Kong.

Keywords ARIMA; ARIMA-GARCH; ANN; Multi-Population O'Hare.

Mathematics Subject Classification 62M10

1 Introduction

Modelling and forecasting mortality rates have been popular issues in the research world to predict the accurate mortality projection of population between genders. Predicting future mortality rate is essential and difficult to achieve due to insufficient relevant knowledge needed to forecast the mortality rate. Thus, researchers need to find the best mortality model to have an accurate prediction of mortality. There were numerous models presented. The well-known multi-population mortality model which is the model provided by Li and Lee [1] is the most well-known for estimating the age-specific mortality rate and assimilating the relation between population. The model has received much interest and become a basis for mortality modelling. However, when compared to other mortality models [3; 4], the mortality modelling from Lee and Carter (LC) [2] model has been subjected to criticism due to limited control in adjusting age variations and historical fit. Wan and Bertschi [5] improved the Li and Lee model by incorporating the Plat [6] framework into the regression modelling structure, inspired by the constraints of the Li and Lee model.

Nevertheless, the model developed by Wan and Bertschi [5] has a drawback in which the model is unable to capture the nonlinear pattern found at the earlier ages of mortality [7]. This is mainly due to the integration of the Plat model that only fits death rates above the age of 20 [8, 9]. According to Hauser and Weir [10] the inclusion of the earlier ages of mortality is significant because it acts as a preliminary study for future life. Furthermore, mortality variations across all ages have a substantial influence on estimating the average life expectancy. As a result, the failure of the existing stochastic mortality model to incorporate in early mortality would eventually result in an incorrect estimation of annuity price [11, 12]. Therefore, the model of multi-population O'Hare was proposed in this study by expanding the Li and Lee [1] model. This model can be developed by integrating the quadratic impact parameter by capturing the variabilities between ages, period and population given by O'Hare and Li to fill in the model gap by Wan and Bertschi [5].

In the early development of the LC model in 1992, the model has two factors including two age-specific parameters for each group and a time-varying effect so that the tendency of all age-specific central death rates follows the same pattern of stochastic evolution over time. Using the mortality data from 1933 to 1987, Lee and Carter used the LC model to forecast death rates in the United States. At the moment, this approach is still being used to predict the death rate of different nations' populations all over the world [13]. For example, the LC model and Li-Lee model have been enforced to forecast the mortality rates in China and other 15 developed countries including Italy, Japan, Canada, Norway, France, Austria, Spain, Denmark and others [14]. The extension of LC model known as the functional data (FD) model proposed by Hyndman and Ullah [15] was performed on a research comparing the accuracy of forecasting mortality rate in ASEAN countries namely Indonesia, Singapore, Malaysia, and Thailand [16].

The ARIMA model has been used in many studies in forecasting mortality rates. The LC model has been applied to age-specific death rates by gender in Argentina [17]. In this paper, they used autoregressive moving average (ARIMA) and the space-state model to forecast the general index for the time period in order to project life expectancy at birth. Aside from that, Ngataman et al. [18] had used LC model on Malaysia's data to forecast mortality rates for time periods in the future using the ARIMA method. Hong *et al.* [13] used several LC-based models to forecast the mortality rates in a case study of Malaysian population. According to their study, they found that the LC-ARIMA model is the best at forecasting death rates in nations with higher life expectancy and a robust health care system.

Although ARIMA method has become a popular technique for time series analysis, there are two flaws in the model. First, it assumes independent and dependent variables have a linear relationship. Meanwhile, real-world relationships are usually non-linear relationships. Thus, ARIMA model does not perform well when the data is complex. This causes heteroscedasticity to exist in the residuals of the ARIMA model. Thus, to address the heteroscedasticity issue, ARIMA model is combined with a Generalized Auto Regressive Conditional Heteroskedasticity (GARCH) model. Dritsaki [19] used ARIMA-GARCH method to forecast the volatility in the return of oil prices since ARIMA method is unable to cope with the volatility and nonlinearity of the data series provided. Other than that, ANN method can also overcome the limitation of ARIMA model because ANN model can capture the nonlinear trend to forecast the mortality rate and produce a mortality model that is more coherent and non-divergent [13]. Other than that, ANN can also predict accurate mortality rate by reducing complex and non-linear interaction between variables by decreasing the value of expected and actual outcomes [20].

This paper will arrange as follows; in section 2, an overview of the multi-population mortality model and time series models which are multi-population O'Hare model and forecasting model of ARIMA model, ARIMA-GARCH model and ANN model will be explained. The results of the mortality rate forecast for Malaysia, Taiwan, Japan, Hong Kong, Australia, USA, UK, Canada, and Switzerland using the multi-population O'Hare model in combination with ARIMA, ARIMA-GARCH, and ANN will then be compared to determine the most accurate model with the lowest margin of error. In conclusion, this study will summarize the results gathered from analyzing the forecasting models for each country.

2 Materials and Methods

In this section, the reference model which is multi-population O'Hare model is discussed briefly. A good mortality model can capture accurate observed mortality data for the whole age span. This model, which is multi-population O'Hare framework is intended to capture the central tendency of a set of populations, as well as population-specific volatility [7]. The model is developed from O'Hare single stochastic mortality model and Li Lee model from multi-population mortality model.

2.1 Multi-population O'Hare model

Li and Lee [1] had developed Li-Lee method which is $\ln(m_{x,t,i}) = a_{x,i} + \beta_x k_t$. Hence, inspired by this model and O'Hare model, Nor *et al.* [7] has developed a new model which is multi-population O'Hare model. The formula of multi-population O'Hare is stated as below:

$$\ln(m_{x,t,i}) = a_{x,i} + k_t^1 + k_t^2(\overline{x} - x) + k_t^3\left((\overline{x} - x) + \left[(\overline{x} - x)^+\right]^2\right) + a_{x,i} + \sum_{j-1}^{L} \beta_{x,i,j} k_{t,i,j} + \varepsilon_{x,t,i}, \qquad (1)$$

where $\ln(m_{x,t,i})$ is predicted mortality rate for age x in calendar year t for reference population total of male and female denoted by $i = 1, 2$. a_x is overall mortality rates across ages, k_t, k_t^1, k_t^2, k_t^3 and $k_{t,i}$ is time varying mortality index and β_x is the age-specific component that shows the speed with which the mortality rate responds to changes in the time-varying mortality index. The term $\overline{x}$ represents the sample average of age groups where the value of sample will be zero when $(\overline{x} - x)^+$ is negative and $(\overline{x} - x)^+$ is positive, it takes the value of $\overline{x} - x$

Geometric mean $(\prod_{i=1}^{M} m_{x,t,i})$, is used to calculate the weighted average of mortality rates where L is the component's value of j^{th} principal component and M is the population's value for i^{th} population. Maximum Likelihood Estimation (MLE) for technique of Poisson distribution is used to estimate $a_{x,i} + k_t^1 + k_t^2(\overline{x} - x) + k_t^3\left((\overline{x} - x) + \left[(\overline{x} - x)^+\right]^2\right)$. Next, estimation of mortality model is done by fitting the procedure and Singular Value Decomposition (SVD)'s technique is used to estimate the parameter of population-specific parameters $\beta_{x,i,j}$ and $k_{t,i,j}$.

2.2 Time series model

In this section, the time-series models which are Autoregressive Integrated Moving Average (ARIMA) model, Integrated Moving Average with Generalized Autoregressive Conditional Heteroscedasticity (ARIMA-GARCH) model and Artificial Neural Network (ANN) model are discussed briefly. The value of forecasting time-series parameters k_t, k_t^1, k_t^2, k_t^3 and $k_{t,i}$ of the model multi-population O'Hare model are predicted by using the method of ARIMA, ARIMA-GARCH and ANN. The process of forecasting is to estimate the values of out-sample error measurement for male and female.

2.2.1 ARIMA time series model

Parameter of the time series k_t, k_t^1, k_t^2, k_t^3 and $k_{t,i}$ of the model multi-population O'Hare model is selected and fitted by using ARIMA process. Then, ARIMA (p,d,q) is selected to forecast the time series parameter of multi-population O'Hare by following Box-Jenkins procedure. The Box–Jenkins method was chosen to develop a simulation and forecasting tool because of its: 1) ability to deal with complex situations; 2) adaptability in processing dependent time-series data; 3) advanced mathematical and statistical processes; 4) risk and uncertainty analysis functionality; and 5) ease of implementation. For any collection of data, Box–Jenkins technique frequently provide the most accurate forecasting models [21].

There are four steps needed to apply ARIMA (p,d,q) time series before the model of multi-population O'Hare is forecasted. The four steps are:

Step 1: It is necessary to test on the stationarity of the data. If the data from time-varying parameters estimated from observed mortality rates are non-stationary, take the first differences of the data until the data is stationary.

Step 2: Choose the appropriate ARIMA (p,d,q) by testing on ACF and PACF. ARIMA (p,d, q) is estimated using Maximum Likelihood estimation.

Step 3: Once the values for p, d, and q are known, try applying the ARIMA (p, d, q) model to the period-mortality index. Then, verify the normality assumption in the residuals from the selected ARIMA (p, d, q) model.

Step 4: Calculate the forecast value in the next h year once the residuals look like white noise.

2.2.2 ARIMA-GARCH time series model

There are two steps needed to analyze and forecast time series parameters of multi-population O'Hare with ARIMA-GARCH model.

Step 1: ARIMA model was used to fit on stationary and linear time series data, with the non-linear element of the data being contained in the residuals of the linear model.

Step 2: To get non-linear residuals pattern, GARCH model was used.

Therefore, the combination of ARIMA model and conditional variance of GARCH model will produce a multi-population mortality model that contains non-linear residual pattern, and the procedure was estimated based on maximum likelihood estimation.

Next, residuals were used to do a diagnostic test on ARIMA-GARCH model. Then, autocorrelation function squared residuals were used to test on conditional heteroscedasticity and serial correlation test was applied by using Ljung and Box [22] for all autocorrelation time delays. Forecast evaluation on ARIMA-GARCH consists of two types of forecasts which are static and dynamic. Static forecast is also known as one-step forecasting while dynamic forecast is known as n-step forecasting.

2.2.3 ANN time-series model

ANN is the method proposed by McCulloch and Pitts [23] to model an activity of neuron in the brain. Hebb [24] developed a reinforcement – based learning centre to understand more on human brain. The research of ANN time series model started in 1958 by Rosenblatt to create a machine learning system based on the biological structure of brain, especially the bioelectrical activity of brain neurons [25].

There were four steps that had to be applied to re-estimate the values of k_t from multi-population O'Hare as follows:

Step 1: Standardise the value of k_t.
The data were standardized and the value of k_t was produced from the observed value minus the lowest value of k_t and divided by the difference of the highest and lowest value of k_t.

Step 2: Preparing data for training and testing.

The first three components of the data correspond to the input layers and the fourth component defines the desired value for the output layer to train the ANN model for time series forecasting.

Framework of multi-population O'Hare with ANN model is discussed as below:

Normalisation of Data

- The equation of normalization was used to the parameter of k_t for each year.

Separation of Data

- The normalized data of k_t was then divided into two groups: training and testing. Then, the pattern in Step 2 was followed for the training pattern of the dataset.

Training Set

- The trained data with the pattern indicated in Step 2 was transferred. Then, to find the ideal combination of hyperparameters for the ANN algorithm, a preliminary round of trial and error was carried out to select the appropriate number of hidden layers and neurons based on the error measurements.

Validation Set

- The assumption of k_t values can be carried out after the training process is completed. Then, the process of 'denormalized' the k_t values will be done if the prediction of k_t values is normalized by multiplying the differences of highest and lowest k_t values of data and adding up the lowest values of k_t. Then, multi-population O'Hare in equation 1 is used into 'denormalization' of k_t to calculate the values of mortality rates.

Step 3: Data Fitting

Some rules of thumb were applied to recognise the right number of neurons needed in hidden data such as the following:

- Input nodes contribute about 70-90% of neurons' quantity in hidden nodes.
- The number of neurons in hidden nodes should be less than twice the number of neurons in input nodes.
- The number of neurons in hidden nodes is in the middle of the quantity of neurons in input and output nodes.

A good combination of hyperparameter for ANN algorithm was fulfilled by calculating the number of hidden nodes and neurons from trial-and-error measurement and later was used for forecasting procedure.

Step 4: Estimation of the k_t values
In order to estimate the values of k_t from 2016 to 2018, the ANN model was developed. The normalized k_t values were multiplied with the highest and lowest values of k_t and adding up the minimum amount of k_t values was applied when the k_t values were normalized.

2.2.4 Measurement errors

To compare the model performance between nine countries, Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE) were used as measurement errors. The equation of MAPE and RMSE are denoted as below:

$$\text{MAPE} = \frac{1}{NT}\sum_{x}\sum_{t}\frac{|\hat{m}_{x,t} - m_{x,t}|}{m_{x,t}} \times 100 \tag{2}$$

$$\text{RMSE} = \left(\sqrt{\frac{1}{NT}\sum_{x}\sum_{t}\frac{\hat{m}_{x,t} - m_{x,t}}{m_{x,t}}}\right) \times 100 \tag{3}$$

where N is the length of age-group, T is length of period year t, $\hat{m}_{xt}$ is anticipated mortality rate values and $m_{x,t}$ is the actual mortality rates. The lowest value of MAPE and RMSE for both gender, that is male, and female shows lower prejudice, less variance magnitude and less size of variance [7].

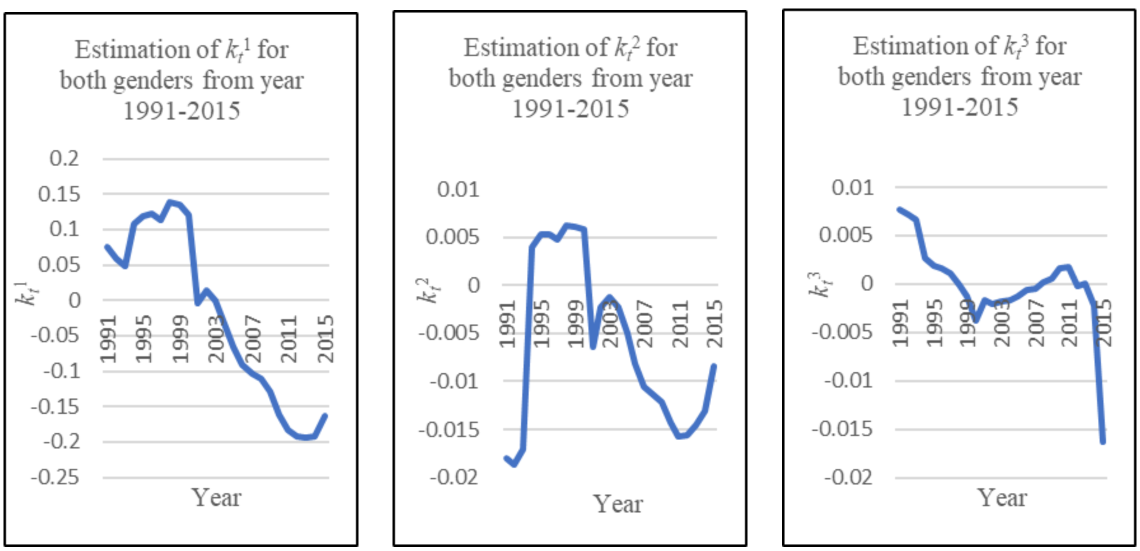
3 Results and Discussions

The multi-population O'Hare model was applied from the period of 1991 to 2015 as in-sample evaluation while multi-population O'Hare with ARIMA, multi-population O'Hare with ANN and multi-population O'Hare with ARIMA-GARCH model were used for estimating the out-sample values from the period of 2016 to 2018. The models were fitted to the mortality datasets of 9 countries which were Taiwan, Japan, Hong Kong, Australia, USA, UK, Canada, Switzerland, and Malaysia from the year 1991-2018 collected from the website of Human Mortality Database from mortality.org and Department of Statistics Malaysia (DOSM).

3.1 Application to mortality datasets for Malaysia

Figure 1 shows the graph of period effect parameters of time varying mortality index for multi-population O'Hare model on Malaysian mortality rates for male and female total population. k_t^1 represents changes in mortality rate for people of all ages, k_t^2 represents the mortality rate for different ages, reflecting that historical fitting might be different among ages and the parameter of k_t^3 represents the mortality rate of young ages. k_t^3 is used to capture nonlinear incidents that happen during young ages. The estimation of k_t is important to ensure that the historical data can be fitted well with the mortality model and the time series k_t is used to forecast accurate death rates using traditional time series forecasting techniques [9].

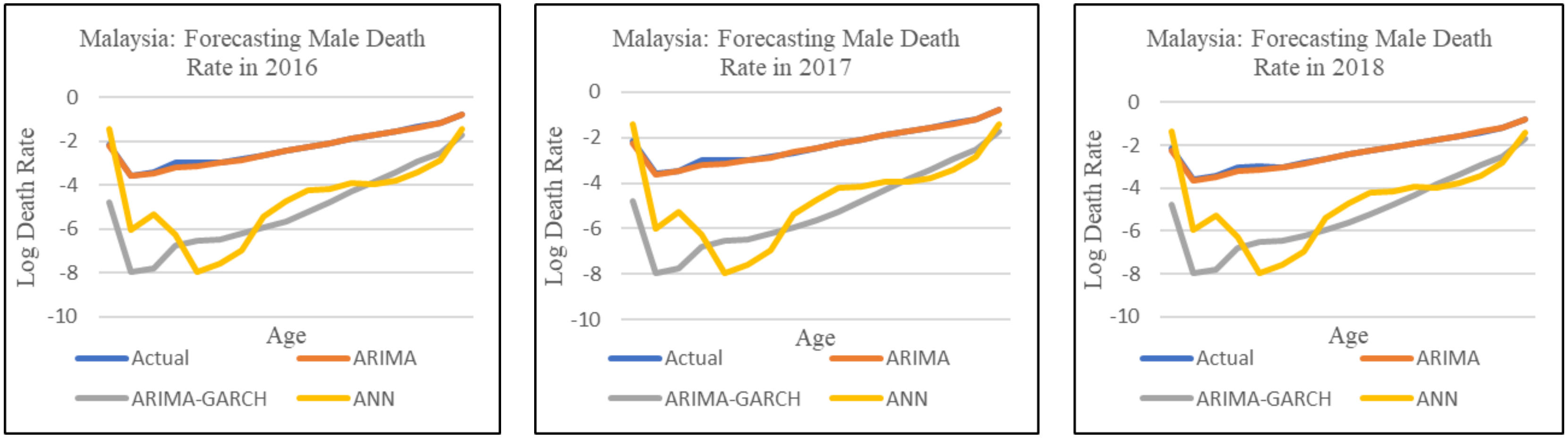
Figure 1(a) shows that the value k_t^1 for both Malaysian gender of male and female decreases as the year increases from 1991 to 2015. Figure 1(b) shows that the value of k_t^2 for both Malaysian gender of male and female increases from the year 1991 to 2000 and decreases from the year 2001 to 2015. Figure 1(c) shows the value of k_t^3 for both Malaysian gender of male and female decreases as the year increases from year 1991 to 2015. Although the value of k_t is decreasing for k_t^1, k_t^2 and k_t^3 for Malaysian gender of male and female, but the trend of decreasing shows linear decreasing and does not have much fluctuation for all k_t value. The linear decreasing explains the reason that ARIMA time-series model outperformed ARIMA-GARCH and ANN model in forecasting Malaysian mortality rates.



(a) k_t^1 from 1991 to 2015 (b) k_t^2 from 1991 to 2015 (c) k_t^3 from 1991 to 2015

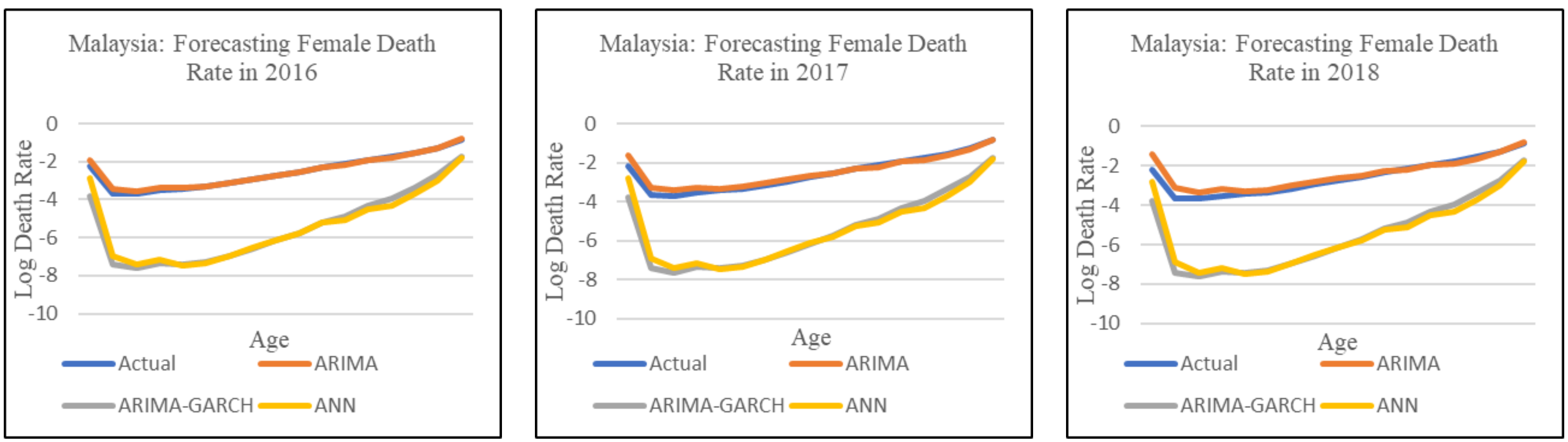
Figure 1: Comparison of the values of k_t from 1991 to 2015

Figure 2 and Figure 3 show the graph comparison of the forecasting of male and female mortality rates for the year 2016, 2017 and 2018 using the actual data collected from Department of Statistics Malaysia (DOSM). The Malaysia mortality rates are forecasted using multi-population O'Hare with ARIMA-GARCH, multi-population O'Hare with ARIMA and multi-population O'Hare with ANN.



(a) 2016 (b) 2017 (c) 2018

Figure 2: Forecasting male death rate from 2016 to 2018



(a) 2016 (b) 2017 (c) 2018

Figure 3: Forecasting female death rate from 2016 to 2018

The purpose of choosing ARIMA, ARIMA-GARCH and ANN to forecast the mortality rates is because ARIMA is better in capturing linear trend of time series while ARIMA-GARCH is used to predict volatile changes in mortality rate and ANN is used to capture nonlinear trend on mortality rates due to its capabilities to capture nonlinear behaviours [11; 13]. The figures show that multi-population O'Hare with ARIMA is the most accurate to the actual data compared to multi-population O'Hare with ARIMA-GARCH and multi-population O'Hare with ANN. Multi-population mortality O'Hare model with ARIMA is accurate to capture the mortality rate of Malaysian mortality data with 16 age length from 0 to 80 years old because ARIMA model is suitable to detect a small amount of data in a short term. The trend of Malaysian mortality data is a linear trend where the oldest age group has the highest death risk whilst the age of 0 to 20 years old has the lowest mortality risk [16; 26; 27].

3.2 Analysis of the best model for male and female

Table 1 shows the result of in-sample measurement for countries of Taiwan, Japan, Hong Kong, Australia, USA, UK, Canada, Switzerland and Malaysia. The result of Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE) is based on the basic model which is multi-population O'Hare mortality model with the age of 0 – 80 from the year 1991 to 2015. Result of in-sample from Table 1 shows that Malaysia has the lowest value of MAPE and RMSE for male and female population compared to other countries.

Table 1: In-Sample Fit Error Measurements for Male and Female

	Male		**Female**	
Model	**MAPE**	**RMSE**	**MAPE**	**RMSE**
Taiwan	0.049	0.070	0.060	0.087
Japan	0.037	0.049	0.043	0.058
Hong Kong	0.155	0.371	0.267	2.204
Australia	0.071	0.096	0.082	0.115
USA	0.028	0.035	0.030	0.037
UK	0.053	0.068	0.058	0.076
Canada	0.050	0.072	0.059	0.087
Switzerland	0.110	0.183	0.146	0.291
Malaysia	0.021	0.027	0.024	0.031

Table 2: Out-Sample Fit Error Measurements for Male

Model	**ARIMA**	**Hybrid ARIMA-GARCH**	**ANN**	**ARIMA**	**Hybrid ARIMA-GARCH**	**ANN**	**ARIMA**	**Hybrid ARIMA-GARCH**	**ANN**
Forecast Horizon	**2016**	**2016**	**2016**	**2017**	**2017**	**2017**	**2018**	**2018**	**2018**
	MAPE	**MAPE**	**MAPE**	**MAPE**	**MAPE**	**MAPE**	**MAPE**	**MAPE**	**MAPE**
Taiwan	0.413	**0.172**	0.486	0.820	**0.338**	0.978	1.210	**0.481**	1.453
Japan	0.633	**0.606**	0.677	1.268	**1.141**	1.352	1.896	**1.595**	2.015
Hong Kong	3.812	8.567	**1.001**	4.052	8.862	**1.398**	4.267	9.204	**1.869**
Australia	0.696	**0.542**	0.761	1.391	**1.071**	1.509	2.101	**1.576**	2.250
USA	0.471	**0.448**	0.483	0.941	**0.891**	0.957	1.418	**1.337**	1.434
UK	0.698	**0.569**	0.731	1.363	**1.081**	1.436	2.038	**1.602**	2.150
Canada	0.672	**0.590**	0.740	1.351	**1.179**	1.469	2.034	**1.762**	2.192
Switzerland	0.853	**0.782**	0.904	1.703	**1.578**	1.806	2.561	**2.380**	2.710
Malaysia	**0.069**	0.187	1.263	**0.149**	0.380	1.950	**0.235**	0.593	3.356
	RMSE	**RMSE**	**RMSE**	**RMSE**	**RMSE**	**RMSE**	**RMSE**	**RMSE**	**RMSE**
Taiwan	0.447	**0.215**	0.505	0.625	**0.302**	0.719	0.753	**0.355**	0.877
Japan	0.640	**0.614**	0.679	0.906	**0.821**	0.960	1.106	**0.944**	1.168
Hong Kong	34.11	79.12	**4.875**	34.11	79.12	**4.874**	34.12	79.12	**5.030**
Australia	0.705	**0.548**	0.767	0.997	**0.767**	1.075	1.228	**0.924**	1.308
USA	0.506	**0.481**	0.499	0.715	**0.678**	0.700	0.879	**0.831**	0.857
UK	0.708	**0.585**	0.742	0.980	**0.793**	1.033	1.196	**0.959**	1.261
Canada	0.685	**0.612**	0.748	0.973	**0.864**	1.050	1.194	**1.054**	1.278
Switzerland	0.862	**0.797**	0.909	1.218	**1.133**	1.285	1.495	**1.393**	1.574
Malaysia	**0.119**	0.236	2.805	**0.175**	0.340	2.903	**0.220**	0.430	4.163

Table 3: Out-Sample Fit Error Measurements for Female

Model	**ARIMA**	**Hybrid ARIMA-GARCH**	**ANN**	**ARIMA**	**Hybrid ARIMA-GARCH**	**ANN**	**ARIMA**	**Hybrid ARIMA-GARCH**	**ANN**
Forecast Horizon	**2016**	**2016**	**2016**	**2017**	**2017**	**2017**	**2018**	**2018**	**2018**
	MAPE	**MAPE**	**MAPE**	**MAPE**	**MAPE**	**MAPE**	**MAPE**	**MAPE**	**MAPE**
Taiwan	0.488	**0.221**	0.486	0.981	**0.428**	0.978	1.463	**0.598**	1.453
Japan	0.682	**0.658**	0.677	1.368	**1.256**	1.352	2.048	**1.786**	2.015
Hong Kong	4.690	11.68	**1.001**	7.374	17.92	**1.398**	10.69	25.84	**1.869**
Australia	0.771	**0.527**	0.761	1.532	**1.043**	1.509	2.293	**1.535**	2.250
USA	0.477	**0.471**	0.483	0.944	**0.933**	0.957	1.412	**1.394**	1.434
UK	0.733	**0.627**	0.731	1.438	**1.214**	1.436	2.156	**1.815**	2.150
Canada	0.743	**0.687**	0.740	1.481	**1.362**	1.469	2.218	**2.029**	2.192
Switzerland	0.906	**0.867**	0.904	1.809	**1.739**	1.806	2.718	**2.617**	2.710
Malaysia	**0.208**	0.595	1.263	**0.664**	1.191	1.950	**1.409**	1.800	3.356
	RMSE	**RMSE**	**RMSE**	**RMSE**	**RMSE**	**RMSE**	**RMSE**	**RMSE**	**RMSE**
Taiwan	0.505	**0.233**	0.505	0.719	**0.321**	0.719	0.877	**0.372**	0.877
Japan	0.684	**0.661**	0.679	0.970	**0.894**	0.960	1.186	**1.042**	1.168
Hong Kong	30.92	79.51	**4.875**	38.63	96.89	**4.874**	48.49	120.4	**5.030**
Australia	0.776	**0.532**	0.767	1.090	**0.745**	1.075	1.331	**0.896**	1.308
USA	0.494	**0.482**	0.499	0.692	**0.675**	0.700	0.846	**0.824**	0.857
UK	0.744	**0.647**	0.742	1.037	**0.892**	1.033	1.268	**1.085**	1.261
Canada	0.751	**0.699**	0.748	1.058	**0.980**	1.050	1.293	**1.191**	1.278
Switzerland	0.911	**0.874**	0.909	1.286	**1.239**	1.285	1.577	**1.520**	1.574
Malaysia	**0.346**	0.968	2.805	**0.863**	1.351	2.903	**1.639**	1.645	4.163

Table 2 and Table 3 show the result of out-sample error measurements for male and female in three models, which are multi-population O'Hare with ARIMA, multi-population O'Hare with ARIMA-GARCH and multi-population O'Hare with ANN model. In order to assess the goodness of fit test, the forecast mortality was projected for 2016, 2017 and 2018 years ahead and measured using MAPE and RMSE. The three models are used to compare which country gives the best projection for estimation of mortality rates based on the lowest value of MAPE and RMSE.

The error of out-sample fit error measurement for male from Table 2 and error of out-sample fit error measurement for female in Table 3 shows that when the three models are forecasting for year 2016, 2017 and 2018, multi-population O'Hare with ARIMA-GARCH shows the lowest value of MAPE for countries of Taiwan, Japan, Australia, USA, UK, Canada and Switzerland. Multi-population O'Hare with ANN has the lowest MAPE value when forecasting Hong Kong mortality rate and multi-population O'Hare with ARIMA has the lowest value of MAPE when forecasting Malaysia mortality rate.

Furthermore, result of RMSE from Table 2 and Table 3 shows that when the three models are forecasting for year 2016, 2017 and 2018, the lowest value of RMSE can be seen from multi-population O'Hare with ARIMA-GARCH for the countries of Taiwan, Japan, Australia, USA, UK, Canada and Switzerland. Multi-population O'Hare with ANN has the lowest RMSE value when forecasting Hong Kong mortality rate and multi-population O'Hare with ARIMA has the lowest value of RMSE when forecasting Malaysia mortality rate.

Generally, most of the countries selected in this study showcased that multi-population O'Hare with ARIMA-GARCH shows the lowest value of MAPE and RMSE for both male and female. Only Hong Kong has the lowest value of MAPE and RMSE for male and female when forecasting using the model of multi-population O'Hare with ANN and Malaysia has the lowest value of MAPE and RMSE for male and female using the model of multi-population O'Hare with ARIMA. The result indicates that multi-population O'Hare with ANN outperformed other model when forecasting using Hong Kong mortality rates because ANN is good in solving non-linear problems. Meanwhile, multi-population O'Hare with ARIMA shows the lowest value of MAPE and RMSE for Malaysia due to the advantage of ARIMA method in solving time-series with linear trend. Lastly, multi-population O'Hare with ARIMA-GARCH shows the lowest value of MAPE and RMSE for countries except Hong Kong and Malaysia due to the advantage of ARIMA-GARCH to capture linear and non-linear pattern of mortality data.

4 Conclusion

In conclusion, this study proposed multi-population O'Hare with ARIMA, multi-population O'Hare with ARIMA-GARCH and multi-population O'Hare with ANN to forecast the mortality rate for 9 countries which are Malaysia, Taiwan, Japan, Hong Kong, Australia, USA, UK, Canada and Switzerland. The performance to forecast on mortality rate for each country was then assessed by using measurement errors denoted as AE, MAPE and RMSE. Based on the value of measurement errors, multi-population O'Hare with ARIMA performed better than the others at forecasting mortality rates for male and female in Malaysia. Multi-population O'Hare with ANN performed better at forecasting mortality rates for male and female in Hong Kong, and multi-population O'Hare with ARIMA-GARCH performed better at forecasting mortality rates for male and female in Taiwan, Japan, Australia, the USA, the UK, Canada, and Switzerland.

Acknowledgment

This work was supported by the Ministry of Higher Education under Fundamental Research Grant Scheme (FRGS/1/2020/STG06/UTM/02/6), with vot R.J130000.7854.5F370 and Universiti Teknologi Malaysia (UTM) for the UTM Fundamental Research (UTMFR) Grant with vot number Q.J130000.2554.21H65.

References

[1] Li, N., & Lee, R. Coherent mortality forecasts for a group of populations: An extension of the Lee-Carter method. *Demography*. 2005. 42(3): 575-594.

[2] Lee, R. D., & Carter, L. R. Modeling and forecasting US mortality. *Journal of the American Statistical Association*. 1992. 87(419): 659-671.

[3] Basellini, U., Camarda, C. G., & Booth, H. Thirty years on: A review of the Lee–Carter method for forecasting mortality. *International Journal of Forecasting.* 2022.

[4] Shair, S. N., Zolkifi, N. A., Zulkefi, N. F., & Murad, A. A Functional Data Approach to the Estimation of Mortality and Life Expectancy at Birth in Developing Countries. *Pertanika Journal of Science & Technology.* 2019. 27(2).

[5] Wan, C., & Bertschi, L. Swiss coherent mortality model as a basis for developing longevity derisking solutions for Swiss pension funds: A practical approach. *Insurance: Mathematics and Economics*. 2015. 63. https://doi.org/10.1016/j.insmatheco.2015.03.025.

[6] Plat, R. On stochastic mortality modeling. *Insurance: Mathematics and Economics*. 2009. 45(3): 393-404.

[7] Nor, S. R. M., Yusof, F., & Norrulashikin, S. M. Coherent Mortality Model in A State-Space Approach. *Sains Malaysiana*. 2021. 50(4): 1101-1111.

[8] Babu, C. N., & Reddy, B. E. Prediction of selected Indian stock using a partitioning–interpolation based ARIMA–GARCH model. *Applied Computing and Informatics*. 2015. 11(2): 130-143.

[9] O'Hare, C., & Li, Y. Explaining young mortality. *Insurance: Mathematics and Economics*. 2012. 50(1): 12-25.

[10] Hauser, R. M., & Weir, D. Recent developments in longitudinal studies of aging in the United States. *Demography*. 2010. 47(SUPPL.1)

[11] O'Hare, C. Essays in modelling mortality rates. *Available at SSRN 2223749*. 2012.

[12] Seklecka, M., Pantelous, A. A., & O'Hare, C. (2019). The impact of parameter uncertainty in insurance pricing and reserve with the temperature reserve with the temperature related mortality model. *Journal of Forecasting.* 2019. 38(4): 327-345.

[13] Hong, W. H., Yap, J. H., Selvachandran, G., Thong, P. H., & Son, L. H. Forecasting mortality rates using hybrid Lee–Carter model, artificial neural network and random forest. *Complex & Intelligent Systems*. 2021. 7(1): 163-189.

[14] Li, H., Lu, Y., & Lyu, P. Modeling and forecasting Chinese population dynamics in a multi-population context. *SOA Research Reports. Schaumburg: Society of Actuaries*. 2018.

[15] Hyndman, R. J., & Ullah, M. S. Robust forecasting of mortality and fertility rates: a functional data approach. *Computational Statistics & Data Analysis*. 2007. 51(10): 4942-4956.

[16] Shair, S. N., Zolkifli, N. A., Zulkefli, N. F., & Mur, A. A functional data approach to the estimation of mortality and life expectancy at birth in developing countries. *Pertanika Journal of Science and Technology*. 2019. 27(2).

[17] Andreozzi, L., Blacona, M. T., & Arnesi, N. The Lee Carter method, for estimating and forecasting mortality: an application for Argentina. *Documento presentado al ISF*. 2011.

[18] Ngataman, N., Ibrahim, R. I., & Yusuf, M. M. Forecasting the mortality rates of Malaysian population using Lee-Carter method. In *AIP Conference Proceedings* (Vol. 1750, No. 1, p. 020009). (2016, June). AIP Publishing LLC.

[19] Dritsaki, C. The performance of hybrid ARIMA-GARCH modeling and forecasting oil price. 2018.

[20] Weng, S. F., Vaz, L., Qureshi, N., & Kai, J. Prediction of premature all-cause mortality: A prospective general population cohort study comparing machine-learning and standard epidemiological approaches. *PloS one*. 2019. 14(3). e0214365.

[21] Gaynor, P. E., & Kirkpatrick, R. C. *Introduction to time-series modeling and forecasting in business and economics* (No. QA279. 2 G28). 1994.

[22] Ljung, G. M., & Box, G. E. (1978). On a measure of lack of fit in time series models. *Biometrika*. 65(2): 297-303.

[23] Mcculloch, W.S. and Pitts, W. A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophy.* 1943. 5: 115-133.

[24] Hebb, D. *The Organization of Behavior.* New York: Wiley. 1949.

[25] Walczak, S. Artificial neural networks. In *Encyclopedia of Information Science and Technology, Fourth Edition* (pp. 120-131). IGI Global. 2018.

[26] Kamaruddin, H. S., & Ismail, N. Statistical comparison of projection Malaysia mortality rate by using Lee-Carter model and Lee-Carter extension of Hyndman-Ullah. In *AIP Conference Proceedings* (Vol. 2111, No. 1). (2019, June). AIP Publishing.

[27] Nor, S. R. M., Yusof, F., & Bahar, A. Multi-population mortality model: A practical approach. *Sains Malaysiana*. 2018. 47(6): 1337-1347.