UITMCNS
SEREMBAN

J E M U R

JOURNAL OF EXPLORATORY MATHEMATICAL UNDERGRADUATE RESEARCH

RECENT **MATHEMATICAL PROJECTS** AT UITM, *SEREMBAN CAMPUS*

Vol. 002 | PREPARED BY: KPPIM, UiTM N. SEMBILAN. | DATE: 05 OCT 2024

EXTREMAL PROPERTIES OF CERTAIN CLASS OF TILTED UNIVALENT ANALYTIC FUNCTIONS

MONTHLY EXPENDITURE OF PTPTN LOAN RECEIVER AMONG SHAH ALAM UNIVERSITY STUDENTS

LABOUR FORCE PARTICIPATION RATE AND UNEMPLOYMENT RATE: A MALAYSIAN PERSPECTIVE

CLASSIFICATION OF AIR QUALITY IN THE KLANG VALLEY USING K-MEANS CLUSTERING

FACTORS AFFECTING STUDENTS' ACADEMIC PERFORMANCE THROUGH ONLINE DISTANCE LEARNING

Photography courtesy of Porapak Apichodilok

IMPROVING MORTALITY FORECASTING: INTEGRATING THE LEE-CARTER MODEL WITH NEURAL NETWORKS

Hana Natasya Abd Hamid[1], Khairunnisa Khairul Hizam[2] and Nurul Aityqah Yaacob[3,*]
[1,2]Universiti Teknologi MARA Cawangan Negeri Sembilan, Kampus Seremban, Negeri Sembilan, Malaysia, [3]Universiti Teknologi MARA Cawangan Negeri Sembilan, Kampus Kuala Pilah, Negeri Sembilan, Malaysia
[*]corresponding author: aityqah@uitm.edu.my

Abstract

This research integrates the Lee-Carter (LC) model with neural network (NN) methods to enhance mortality forecasting. The LC model, widely used for demographic forecasting, has limitations in capturing complex and nonlinear mortality trends. To address these limitations, we incorporate NN methods, namely a multilayer feed-forward neural network (MFFNN), to identify patterns within mortality data. The study evaluates the performance of the LC and LC-NN models across five countries: Germany, Japan, Malaysia, South Korea, and the United States of America (USA). Findings indicate that the LC-NN model outperforms the LC model, as demonstrated by lower Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) values. This integration significantly improves forecasting accuracy, providing more reliable insights into future mortality trends. The results have significant implications for public health planning and policymaking, offering a robust tool for forecasting demographic changes and their impact on healthcare systems. This integration advances mortality forecasting, leading to better public health outcomes.

Keywords: Mortality forecasting, Lee-Carter Model, Neural network, Multilayer feed-forward network, Public health

1. Introduction

Over the past two centuries, there has been a notable and constant increase in the average life expectancy, with various countries exhibiting consistent improvements in this metric (Torri and Vaupel, 2012). The evaluation of life expectancy and mortality has been a focal point of interest in actuarial and social sciences, leading to the development of numerous models to accurately depict mortality rates (Alho, 1990). Mortality rates in a country are significantly influenced by its growth rate, as well as factors such as medical advancements and social conditions (Baran et al., 2007). These rates serve as reliable indicators of quality of life and provide indirect insights into the prevailing socioeconomic conditions. Moreover, a comprehensive analysis of past mortality trends, encompassing many causes of death, offers significant insight into the determinants contributing to the progressive escalation of mortality rates over successive periods (Zhang et al., 2023). Forecasting future mortality rates is crucial for policy formulation and planning, with the Lee-Carter (LC) model emerging as a prominent tool due to its stochastic approach and widespread applicability (Baran et al., 2007). Introduced in 1992 by Ronald Lee and Lawrence Carter, the LC model used a base model of age-specific death rates with a time series model, known as autoregressive integrated moving average (ARIMA), to effectively forecast mortality rates (Lee and Carter, 1992). The model parameters are computed using Singular Value Decomposition (SVD) and applied to the initial LC model established by Lee and Carter. After obtaining the parameters, the variable k_t will be forecasted with the help of ARIMA. The model includes two sets of age-specific constants: b_x, which reflect age-specific constants, and a_x, which characterise the general mortality pattern by age. In the original LC model, Lee and Carter (1992) proposed the standard univariate ARIMA(0,1,0) time series model which is a random walk with drift model for forecasting future mortality rates. Despite the robustness and widespread applicability of the LC model, it faces criticisms regarding its long-term stability and sensitivity to structural changes over time, which can lead to significant forecast errors if not accurately accounted for (Lee and Miller, 2001; Koissi et al., 2006). ARIMA is adept at capturing temporal dependencies and trends in mortality data, enhancing the LC model's capacity for short-term mortality forecasting (Hong et al., 2021). However, challenges such as overfitting and selecting appropriate ARIMA parameters remain, particularly in cases of sparse data or unstable mortality trends (Nigri

et al., 2019).

This research, therefore, seeks to fill this gap by combining the original LC model with NNs to forecast future changes in the parameter k_t, addressing the limitations observed in ARIMA processes. By leveraging NNs to capture complex mortality trends while utilising the LC model's temporal trend modelling capability, this study aims to create a more accurate and interpretable mortality forecasting model. This integration promises to enhance both the LC model's performance and the precision of mortality forecasts by mitigating biases and incorporating structural modifications. In recent years, various researchers have proposed using machine learning (ML) algorithms to forecast death rates for populations worldwide (Hong et al., 2021). Neural Networks (NNs), in particular, have been suggested for forecasting and simulating death rates due to their ability to identify latent time frame processes and directly forecast mortality (Hainaut, 2018). ML approaches have been utilised to improve the estimation of log mortality rates within the LC framework, contributing to more robust mortality forecasts (Deprez et al., 2017; Levantesi and Pizzorusso, 2019; Richman and Wüthrich, 2018). Despite scepticism in the demographic field, recent research demonstrates the potential of ML and NNs in mortality forecasting. Hainaut (2018) employed NNs to identify hidden time processes and forecast mortality directly, while Deprez et al. (2017) enhanced mortality rate estimation accuracy using ML techniques. Richman and Wuthrich (2019) automated the structure determination of the LC model using NNs, outperforming conventional models in out-of-sample forecasts. Additionally, Hong et al. (2021) proposed a hybrid model combining the LC model, random forest (RF), and artificial neural networks (ANN), demonstrating superior mortality rate forecasting performance.

The paper is organized as follows: Section 2 presents the data and methods used in the paper. Section 3 discusses the results. Finally, Section 4 provides a summary of the conclusions drawn and outlines potential future research directions.

2. Research Method

The mortality data for five countries, which are Germany, Japan, Malaysia, South Korea, and the United States of America (USA) were sourced from the Human Mortality Database (HMD), while data for Malaysia were obtained from the Department of Statistics Malaysia (DOSM) for use in this study. The dataset comprises mortality rates for the general populace, segmented by single years of age up to 95 years for Germany, Japan, South Korea, and the USA, with Malaysia extending up to 85 years and grouped by age. Lee and Carter (1992) computed the average of $\ln(m_{x,t})$ over time to estimate a_x, then used Singular Value Decomposition (SVD) on the centred data to estimate parameters b_x, and k_t. The study commenced with a time series extracted from k_t. The data sets were split into training and testing periods using an 80:20 train-test ratio. Table 1 details the total years covered in each country's data along with the respective years allocated for training and testing sets during the analysis.

Table 1: Total, training and testing set years by country

Country	Total years	Training set years	Testing set years
Germany	1990 – 2020	1990 – 2014	2015 – 2020
Japan	1947 – 2022	1947 – 2007	2008 – 2022
Malaysia	1991 – 2015	1991 – 2010	2011 – 2015
South Korea	2003 – 2020	2003 – 2017	2018 – 2020
USA	1950 – 2021	1950 – 2006	2007 – 2021

2.1. *Lee-Carter Model*

The LC model, introduced by Lee and Carter (1992), is a statistical time series model designed to extrapolate mortality trends and age patterns. The model is formulated using the overall mortality rate (denoted

as $(m_{x,t})$). The LC model is articulated as follows:

$$\ln(m_{x,t}) = a_x + b_x k_t + \varepsilon_{x,t}; x = 1, \ldots, \omega \tag{2.1}$$

where a_x:

$$a_x = \frac{\sum_{t=1}^{n} \ln m_{x,t}}{n} \tag{2.2}$$

is a set of age-specific constants describing the general pattern of mortality by age (Hong et al., 2021). b_x is a set of age-specific constants describing the rates that decline rapidly and which rates decline slowly in response to changes in k, $\varepsilon(x,t)$ is the residual at age x in year t; and $m_{x,t}$ is the central death rate at age x in year t (Booth et al., 2002). The age-specific historical influences are reflected in $\varepsilon(x,t)$, which is not entirely represented by the $N(0,\sigma^2)$ distribution-following, independently and identically distributed model. The parameters in equation (2.1) are estimated using a two-stage method by imposing the following restrictions:

$$\sum_t k_t = 0 \quad \text{and} \quad \sum_x b_x = 1, \tag{2.3}$$

to identify a single solution to the model's set of equations, use the matrix of centered age profiles, $\ln(m_{x,t}) - a_x$, which subjected to the Singular Value Decomposition (SVD) technique for initial calculation of the parameters b_x and k_t (Yaacob et al., 2021).

$$b_x = \frac{1}{\sum_x u_{x,1}} (u_{1,1}, u_{2,1}, \ldots, u_{x,1})^n \tag{2.4}$$

$$k_t = \sigma_1 \times (v_{1,1}, v_{2,1}, \ldots, v_{t,1}) \tag{2.5}$$

The basic mortality rates, $\hat{m}_{x,t} = \frac{D_{x,t}}{E_{x,t}}$ are fitted to the model in equation (2.1), where $D_{x,t} > 0$ indicates the number of deaths of age x at time t, and $E_{x,t}$ is the corresponding central exposure of age at time t. In order to compare actual and forecast mortality, a second stage estimate of k_t is found once b_x and k_t are estimated by satisfying equation (2.2). This ensures that the total expected deaths for each t and the total actual deaths are equal. Thus, the estimates of the parameters meet:

$$\sum_{x=x_1}^{\omega} D_{x,t} = \sum_{x=x_1}^{\omega} E_{x,t} \left[\exp\left(\hat{a}_x + \hat{b}_x \hat{k}_t\right)\right] \forall t \tag{2.6}$$

By counterbalancing the effect of adopting a log transformation of the mortality rates, this adjustment provides more weight to high rates. Using future extrapolation values, i.e., k_{t+n}, an appropriate time series was fitted to k_t to forecast mortality rates. Therefore, the forecasted mortality rate would be:

$$\ln(m_{x,t}) = a_x + b_x k_{t+n} \tag{2.7}$$

where a_x and b_x are fixed (Yaacob et al., 2021). The ARIMA time series models are used to extrapolate the adjusted k_t. Lee and Carter (1992) used a random walk with drift model, which can be expressed as:

$$k_t = k_{t-1} + d + e_t \tag{2.8}$$

where d represents the annual change in k that is constant, and e denotes uncorrelated errors. The uncertainty correlated with a one-year forecast is approximated by the sum of the standard errors in d and e_t. This is used to generate forecasted intervals based on probabilities for the forecasted values of k_t. Age-specific mortality rates are forecasted through the utilisation of extrapolated k_t and fixed a_x and b_x. The fitted rates in this instance are the jump-off rates (i.e., the rates in the final year of the fitting period or jump-off year).

2.2. *Neural Network*

The NN becomes non-linear once we include an intermediate layer that contains hidden neurons. A simple example is shown in Figure 1. The parameters b_1, b_2, b_3 and $w_{1,1}, \ldots, w_{4,3}$ from the data are "learned" (or estimated); the objective is to minimize a function in a high-dimensional space that calculates the difference between the anticipated values ϕy and the actual values ŷ. The value of ŷ is influenced by the matrices of the weights. Weight values are frequently limited to prevent them from becoming extremely huge. The "decay parameter" is a limit on the weights; it is frequently designed to have a value of 0.1. Starting with random values, the weights are afterward modified based on the observed data. Considering this, the forecasts generated by an NN contain an amount of randomness. Therefore, the network is typically trained on multiple occasions from separate random initial states, with the average outcomes being computed.

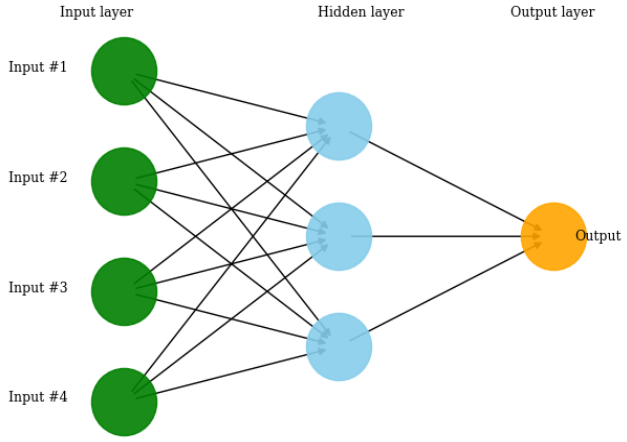

Figure 1: A NN with four inputs and one hidden layer with three hidden neurons (Hyndman and Athanasopoulos, 2021)

In Figure 1 the diagram refers to a multilayer feed-forward network in which each layer of nodes grows inputs from the past layers. The outputs of the nodes in a given layer serve as the inputs for the following layer. The inputs to each node are combined by a weighted linear combination. Afterward, the outcome is altered by a non-linear function before it is delivered as an output. This paper will employ a feedforward neural network with a multi-layer net architecture, including a single layer in the hidden layer. Following this, the backpropagation algorithm is used for training, and the sigmoid function is used for activation. A synopsis of the neural network architecture that will be implemented in this paper. The backpropagation algorithm comprised the subsequent steps, as described by Safitri et al. (2018):

i. The determination of the input nodes' quantity will be accomplished through one of the following methods: evaluation of the correlation value between variables based on relationship proximity; device-generated description of the correlogram R (Hong et al., 2021).
ii. The dataset will be split into two distinct sets: training data and testing data. The training data will be used for the purpose of learning, whereas the testing data will be employed to validate the model that was acquired, specifically to determine whether it adequately describes the existing data. Following that, normalisation of the data will occur. Following that, early initialization will be performed on all weights and bias values using a random number range of 0 to 1, and the activation function will be determined.
iii. Subsequently, input values are entered, and each input is multiplied by its corresponding weight and added together; this process advances the input to the concealed layer. The formula for this operation is as follows:

$$z_j = b_j + \sum_{i=1}^{4} w_{i,j} x_i \tag{2.9}$$

After obtaining the z_j value for each node in the hidden layer, compute the hidden layer's output signal as follows: The expression $z_j = f(z_{\text{net}j})$ denotes f, the predefined activation function.

iv. Execute step 3 in proportion to the quantity of hidden layers employed.

v. Following this, the value of the concealed layer's output will be obtained. The output of the concealed layer is transmitted to the output layer via the formula below:

$$Y_{\text{net}_k} = w_{k0} + \sum_{j=1}^{p} z_j w_{kj} \tag{2.10}$$

vi. Both the actual output and the expected output will subsequently be compared. If deemed unsuitable or if the error is substantial, weight adjustments will be applied to each layer. Weight correction may be performed using either approach. It may begin with the weight between the hidden output layers, or it may be executed from behind. The formula for modifying the weights and biases between the hidden and output layers is as follows:

$$w_{kj}^{(\text{new})} = w_{kj}^{(\text{existing})} + \Delta w_{kj} \tag{2.11}$$

$\Delta w_{kj} = \alpha \delta_k z_j$, $j \neq 0$ is correction formula for weight value and $\Delta w_{k0} = \alpha \delta_k$ is correction formula for bias value, where δ_k was obtained from formula presented below:

$$\delta_k = (t_k - Y_k) \cdot f'(Y_{\text{net}_k}) \tag{2.12}$$

vii. Then value of weight and bias between the input layer and hidden layer will be renewed with formula as follows:

$$v_{ji}^{(\text{new})} = v_{ji}^{(\text{existing})} + \Delta v_{ji} \tag{2.13}$$

$\Delta v_{ji} = \alpha \delta_j x_i$, $i \neq 0$ is correction formula for weight value, and $\Delta v_{j0} = \alpha \delta_j$ is correction formula for bias value, where δ_j was obtained from formula presented below:

$$\delta_j = (\delta_{\text{net}_j} \cdot f')(Z_{\text{net}_j}) \tag{2.14}$$

$$\delta_{\text{net}_j} = \sum_{k=1}^{m} \delta_k w_{kj} \tag{2.15}$$

This procedure is iterated until the error value is minimised to its minimum. Following this, the model with the most accurate estimation weight will be obtained for subsequent forecasting attempts (Safitri et al., 2018).

Once the training procedure is concluded, it is possible to proceed with the forecasting of the k_t values. To renormalize the forecasted k_t values, which have been normalised, the product of the differences between the minimum and maximum k_t values of the data must be applied to the k_t values. The mortality rates are then calculated using the denormalized k_t values and equation (2.1) (Hong et al., 2021).

2.3. *Evaluation Metrics*

The evaluation of forecast accuracy involves comparing forecasted values with actual values, typically done by calculating the forecast error as the difference between the forecasted and actual values. To gauge the accuracy of forecasted mortality rates, commonly used metrics such as MAPE and RMSE are employed (Ibrahim et al., 2021). In this paper, we have computed the MAPE and RMSE for each model $m_{x,t}$ to determine and compare their accuracy in forecasting mortality rates. As MAPE and RMSE decrease, the accuracy of the constructed model increases. The degree of similarity between a measured or predicted value and the genuine or accepted value of the quantity being measured or predicted is the accuracy measure. Its significance in forecasting is that it has an immediate influence on the dependability, credibility, and efficacy of assessments or prognostications. By assessing the degree to which a measurement corresponds to the actual quantity it intends to represent, accuracy measures

guarantee the reliability of experimental, observational, or predicted data for the purposes of analysis, decision-making, and comparison with other datasets.

$$\text{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\frac{|Y_i - P_i|}{Y_i} \times 100 \tag{2.16}$$

$$\text{RMSE} = \left[\frac{1}{N}\sum_{i=1}^{N}(P_i - Y_i)^2\right]^{\frac{1}{2}} \tag{2.17}$$

where N is the number of data, Y_i is an actual value, P_i is the forecasted value of the i^{th} data obtained.

3. Result and Discussion

This section presents the results and discussions derived from the death rate forecasting analysis using two distinct models: the LC model and the integrated LC-NN model. This chapter aims to evaluate the performance and effectiveness of these models in forecasting death rates across the five countries studied. The analysis is based on a thorough examination of observed death rates, forecasted values, and performance measures such as mean absolute percentage error (MAPE) and root mean square error (RMSE). At the outset, the chapter reviews the trends in the log death rates during different periods and among several age groups. These are critical in the basic comprehension of the essentials of the dataset before examining the difference between the forecasted and the actual death rates. After the forecasted and actual log death rates are presented, the chapter then measures the performance of the LC model and the integrated LC-NN model.

3.1. *Comparison of Forecasted and Actual Death Rates*

Figures 2–6 were derived from observed using R-Studio, showcase the comparison between actual and forecasted death rates for total populations in five countries. Circles denote observed rates, while red (NN - Neural Network) and blue (LC – Lee-Carter) lines represent forecasted trends. These visualisations provide crucial insights into the predictive performance of these models across varied demographic scenarios.

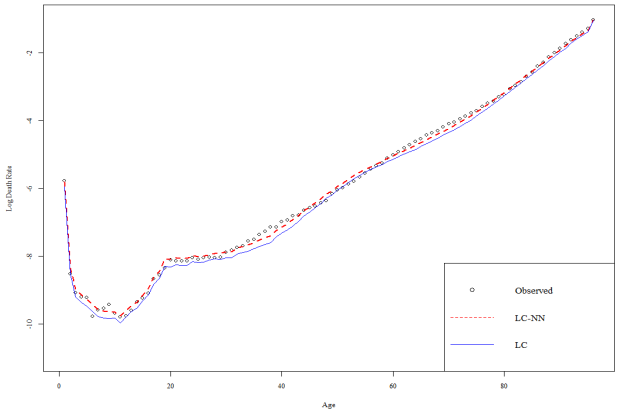

Figure 2: Observed and Forecasted Log of Death Rates for Germany in 2020

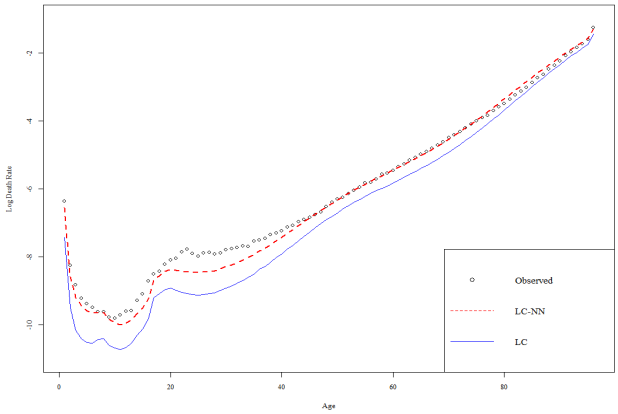

Figure 3: Observed and Forecasted Log of Death Rates for Japan in 2022

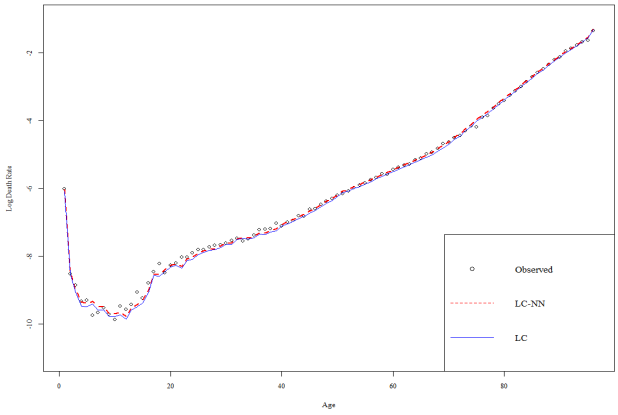

Figure 4: Observed and Forecasted Log of Death Rates for South Korea in 2020

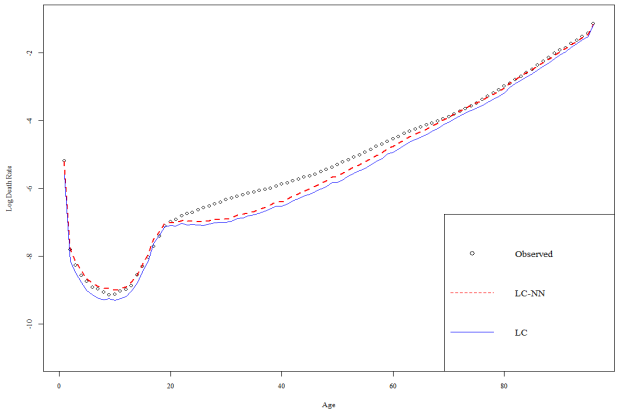

Figure 5: Observed and Forecasted Log of Death Rates for USA in 2021

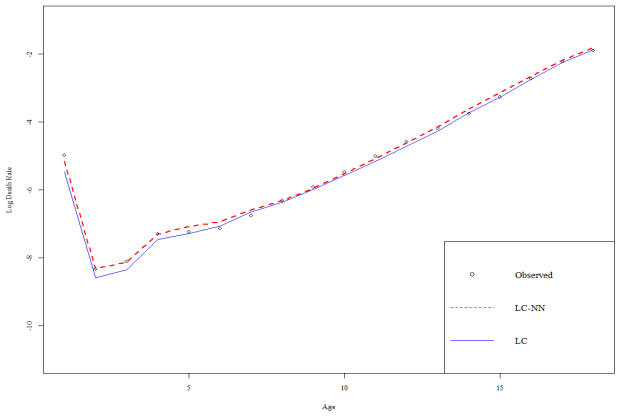

Figure 6: Observed and Forecasted Log of Death Rates for Malaysia in 2015

In short, Figures 2-6 depict the actual and forecasted death rates for different countries and years, demonstrating an ongoing trend of the NN model outperforming the LC model in forecasting accuracy. In each case, both models capture a forecast upward trend in mortality rates with age, but the NN model consistently aligns better with the actual data points. This shows the LC-NN model can identify fluctuations and nonlinearities in the data, resulting in smoother forecasts and a more accurate fit to the actual data. Whether in Germany in 2020, Japan in 2022, South Korea in 2020, the USA in 2021, or Malaysia in 2015, the NN model constantly exceeds the LC model. This indicates that the NN model's ability to analyse complex patterns and correlations in data allows it to produce accurate forecasts of mortality rates across countries and years. Thus, for forecasting mortality rates, the NN model appears to be the most appropriate approach due to its deeper forecasting accuracy and ability to efficiently capture underlying data patterns. Nonetheless, a critical interpretation of model results is required, noting inherent limits and the possible effect of endogenous incidents, such as the continuing COVID-19, on death patterns (Iwamoto, 2023).

3.2. *Performance Evaluation of LC-Model and LC-NN Model*

Table 2 is obtained by data analysis using R-Studio and clarifies the performance study of overall population trends using the LC model and integrated LC-NN model. A thorough examination of the comparison between actual and forecasted death rates is outlined, considering demographic differences across many countries. The forecasting power of the LC and integrated LC-NN model is greatly enhanced by these analytical representations, which show how well it can anticipate demographic changes in the complexity of different national populations.

Table 2: Accuracy measures based on RMSE and MAPE

Country	**GERMAN**		**JAPAN**		**MALAYSIA**	
	LC	LC-NN	LC	LC-NN	LC	LC-NN
RMSE	0.1426	**0.0955**	0.5191	**0.2945**	0.1259	**0.0960**
MAPE	0.0210	**0.0140**	0.0577	**0.0350**	**0.0161**	0.0164

Country	**SOUTH KOREA**		**USA**	
	LC	LC-NN	LC	LC-NN
RMSE	0.1135	**0.1026**	0.1651	**0.1494**
MAPE	0.0119	**0.0107**	0.0217	**0.0206**

*Smallest value are in bold

The RMSE and MAPE values offer a measure of the accuracy of the LC and LC-NN models in forecasting mortality rates for different countries. Germany demonstrates a comparatively low RMSE of 0.1426 in the LC model and an even lower 0.0955 in the LC-NN model, indicating a high level of accuracy for forecasting mortality patterns. The MAPE values for Germany likewise demonstrate this pattern, with the LC model at 0.0210 and the LC-NN model at 0.0140, further confirming enhanced precision with the LC-NN model. Japan exhibits a distinct trend, as seen by an RMSE of 0.5191 in the LC model and a substantially reduced 0.2945 in the LC-NN model. This demonstrates that the LC-NN model significantly enhances the accuracy of predictions compared to the LC model. Japan's MAPE values are 0.0577 for the LC model and 0.035 for the LC-NN model, which confirms the improved accuracy achieved with the LC-NN method.

Malaysia presents a fascinating case, as it exhibits a moderate RMSE of 0.1259 in the LC model and a further improved value of 0.0960 in the LC-NN model. However, the MAPE values show no significant improvement, with the LC model at 0.0161 and the LC-NN model at 0.0164. This implies that the accuracy of the LC-NN model is only slightly impacted. This improvement may be influenced by the way the data is divided or provided in age-grouping forms. The segmentation could have varying effects on the predictive models in comparison to countries that gather data differently.

The RMSE values for South Korea are 0.1135 in the LC model and 0.1026 in the LC-NN model, indicating a moderate level of accuracy. The MAPE values for South Korea are 0.0119 for the LC model and 0.0107 for the LC-NN model. The United States has an RMSE of 0.1651 in the LC model and 0.1494 in the LC-NN model, suggesting a somewhat lower level of accuracy in its forecast compared to Germany and South Korea, but still within acceptable limits. The MAPE values for the United States are 0.0217 for the LC model and 0.0206 for the LC-NN model, indicating a little enhancement with the LC-NN model.

Furthermore, by comparing the RMSE values of the LC and LC-NN models within each country, we may gain insights into the efficacy of LC-NN modelling in improving forecast accuracy. In nations like Germany as well as South Korea, where both models provide low RMSE values, the LC-NN model demonstrates a significantly lower error rate, highlighting its effectiveness in enhancing forecast accuracy. In contrast, in nations such as Malaysia, although the RMSE values are greater, the LC-NN model still performs better than the LC model, although by a small margin.

This paper demonstrates differences in forecasting accuracy among different countries, underscoring the need for customised modelling approaches that consider the distinct demographic and healthcare characteristics of each region. Also, the findings highlight the superior performance of the LC-NN model in improving forecast accuracy, especially in Germany and South Korea, where it outperforms the LC model. These insights show that combining the NN approach with LC coincides with our objectives, as it enables educated, data-driven decision-making processes and improves mortality forecasting methods to effectively address global public health concerns.

4. Conclusion

In the area of mortality studies, numerous hypotheses have emphasised the continuing rise in life expectancy, confirming that advancements in mortality are truly developing. Nevertheless, there is ongoing controversy regarding the pace at which these enhancements happen. This topic is of critical importance as it directly influences our ability to accurately forecast future trends in life expectancy. Therefore, it is essential to focus on the non-linear estimation of the time-dependent parameters within the LC model. Unfortunately, researchers have mostly focused on enhancing the model's accuracy using historical data, disregarding the equally crucial task of forecasting future trends (Nigri et al., 2019). The LC model has been widely used for estimating death rates. This research has significance for examining the complex and extensive consequences of changing mortality patterns, considering the widespread use of the LC technique and other forecasting approaches among actuaries. The integration of NN with the LC model for mortality forecasting has numerous noteworthy implications. This technique improves the model's capacity to detect complex patterns and trends in mortality data that may have been overlooked by the original LC model. Significant progress has been made in mortality modelling to identify patterns in mortality data and provide precise forecasting of mortality rates in the future. Recently, there have been significant changes in forecasting mortality rates. This model uses a time series approach (ARIMA)

to clarify changes in age-specific mortality based on an underlying measure of mortality known as the overall mortality index. The research's significant findings are summarised below:

i. The study compares the performance of the LC and LC-NN models in forecasting mortality rates across Germany, Japan, Malaysia, South Korea, and the United States. It consistently demonstrates that the LC-NN model outperforms the traditional LC model in terms of forecast accuracy, as indicated by lower Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) values across most countries.
ii. The research highlights the efficiency of neural network models, specifically a multilayer feed-forward network, in capturing patterns and nonlinearities in mortality data. This capability enables more precise forecasts, particularly in countries with diverse demographic landscapes and varying mortality trends.
iii. By analysing mortality data from multiple countries, the research provides insights into global mortality trends and the efficacy of different forecasting methodologies. This global perspective is crucial for understanding and addressing public health challenges associated with ageing populations, declining birth rates, and other demographic shifts.
iv. The findings have significant policy implications for healthcare planning and resource allocation. Accurate mortality forecasts can aid governments and healthcare providers in anticipating future healthcare needs, developing effective policies, and ensuring sustainable healthcare systems.

Integrating the LC model with a multilayer feed-forward network enhances forecasting accuracy by leveraging the strengths of both methods. However, challenges may arise when addressing the complex temporal dependencies and inherent non-linearities present in mortality data. Recent advancements, such as the integration of recurrent neural networks (RNN) and long short-term memory (LSTM) networks with the LC framework, have demonstrated the efficacy of machine learning in improving long-term mortality forecasts (Marino et al., 2023). Future research could benefit from incorporating LSTM networks into the LC-NN approach for mortality modeling and forecasting, which is expected to significantly enhance the accuracy and robustness of forecasts. It should be noted that while the LC model is widely utilized for mortality forecasting, it is limited by its reliance on prior trend patterns and lack of consideration for external factors such as environmental and health impacts (Hartawan et al., 2023). Including mortality determinants in forecasting future mortality might lead to a more precise model.

Acknowledgment

The authors express their gratitude to the staff of Universiti Teknologi MARA for their support throughout this research. We also acknowledge the Human Mortality Database (HMD) and the Department of Statistics, Malaysia (DOSM), for providing the necessary data.

References

Alho, J. M. (1990). Stochastic methods in population forecasting. *International Journal of Forecasting*, 6(4):521–530.

Baran, S., Gall, J., Ispany, M., and Pap, G. (2007). Forecasting hungarian mortality rates using the lee-carter method. *Acta Oeconomica*, 57(1):21–34.

Booth, H., Maindonald, J., and Smith, L. (2002). Applying lee-carter under conditions of variable mortality decline. *Population studies*, 56(3):325–336.

Deprez, P., Shevchenko, P. V., and Wüthrich, M. V. (2017). Machine learning techniques for mortality modeling. *European Actuarial Journal*, 7:337–352.

Hainaut, D. (2018). A neural network analyzer for mortality forecast. *ASTIN Bulletin*, 48(2):481–508.

Hartawan, I. G. N. Y., Pujawan, I. G. N., Pranata, K. M., and Jayanta, K. (2023). Forecasting population mortality rates using generalized lee-carter model. *Enthusiastic: International Journal of Applied Statistics and Data Science*, pages 16–24.

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

Hyndman, R. and Athanasopoulos, G. (2021). *Forecasting: Principles and Practice*. OTexts, 3rd edition.

Ibrahim, N. S. M., Lazam, N. M., and Shair, S. N. (2021). Forecasting malaysian mortality rates using the lee-carter model with fitting period variants. In *Journal of Physics: Conference Series*, volume 1988, page 012103. IOP Publishing.

Iwamoto, K. (2023). Deaths jumped 8.9% in japan in 2022 to almost double birth total.

Koissi, M.-C., Shapiro, A. F., and Högnäs, G. (2006). Evaluating and extending the lee-carter model for mortality forecasting: Bootstrap confidence interval. *Insurance: Mathematics and Economics*, 38(1):1–20.

Lee, R. and Miller, T. (2001). Evaluating the performance of the lee-carter method for forecasting mortality. *Demography*, 38(4):537–549.

Lee, R. D. and Carter, L. R. (1992). Modeling and forecasting u. s. mortality. *Journal of the American Statistical Association*, 87(419):659–671.

Levantesi, S. and Pizzorusso, V. (2019). Application of machine learning to mortality modeling and forecasting. *Risks*, 7(1).

Marino, M., Levantesi, S., and Nigri, A. (2023). A neural approach to improve the lee-carter mortality density forecasts. *North American Actuarial Journal*, 27(1):148–165.

Nigri, A., Levantesi, S., Marino, M., Scognamiglio, S., and Perla, F. (2019). A deep learning integrated lee-carter model. *Risks*, 7(1):33.

Richman, R. and Wuthrich, M. V. (2019). Lee and carter go machine learning: Recurrent neural networks. *Available at SSRN 3441030*.

Richman, R. and Wüthrich, M. V. (2018). A neural network extension of the lee-carter model to multiple populations. *Annals of Actuarial Science*, 15(2):346–366.

Safitri, L., Mardiyati, S., and Rahim, H. (2018). Forecasting the mortality rates of indonesian population by using neural network. *Journal of Physics: Conference Series*, 974(1):012030.

Torri, T. and Vaupel, J. W. (2012). Forecasting life expectancy in an international context. *International Journal of Forecasting*, 28(2):519–531.

Yaacob, N. A., Jaber, J. J., Pathmanathan, D., Alwadi, S., and Mohamed, I. (2021). Hybrid of the lee-carter model with maximum overlap discrete wavelet transform filters in forecasting mortality rates. *Mathematics*, 9(18):2295.

Zhang, X., Huang, F., Hui, F. K., and Haberman, S. (2023). Cause-of-death mortality forecasting using adaptive penalized tensor decompositions. *Insurance: Mathematics and Economics*, 111:193–213.