Sains Malaysiana 52(3)(2023): 1011-1021
http://doi.org/10.17576/jsm-2023-5203-23

Hybrid Lee-Carter Model with Adaptive Network of Fuzzy Inference System and Wavelet Functions

(Model Hibrid Lee-Carter dengan Rangkaian Adaptif Sistem Inferens Kabur dan Fungsi Gelombang Kecil)

Jamil J. Jaber[1] Nurul Aityqah Yaacob [2,3,*] & Sadam Alwadi[1]

[1]*Department of Finance, Faculty of Business, The University of Jordan /Aqaba branch, Aqaba, Jordan*
[2]*Institute of Mathematical Sciences, Faculty of Science, Universiti Malaya, 50603 Kuala Lumpur, Federal Territory, Malaysia*
[3]*Mathematical Sciences Studies, College of Computing, Informatics and Media, Universiti Teknologi MARA Cawangan Negeri Sembilan, Kampus Kuala Pilah, 72000 Kuala Pilah, Negeri Sembilan, Malaysia*

ABSTRACT

Mortality studies are essential in determining the health status and demographic composition of a population. The Lee–Carter (LC) modelling framework is extended to incorporate the macroeconomic variables that affect mortality, especially in forecasting. This paper makes several major contributions. First, a new model (LC-WT-ANFIS) employing the adaptive network-based fuzzy inference system (ANFIS) was proposed in conjunction with a nonlinear spectral model of maximum overlapping discrete wavelet transform (MODWT) that includes five mathematical functions, namely, Haar, Daubechies (d4), least square (la8), best localization (bl14), and Coiflet (c6) to enhance the forecasting accuracy of the LC model. Annual mortality data was collected from five countries (Australia, England, France, Japan, and the USA) from 1950 to 2016. Second, we selected gross domestic product (GDP), unemployment rate (UR), and inflation rate (IF) as input values according to correlation and multiple regressions. The input variables in this study were obtained from the World Bank and Datastream. The output variable was collected from the mortality rates in Human Mortality Database. Finally, the LC model's projected log of death rates was compared with wavelet filters and the traditional LC model. The performance of the proposed model (LC-WT-ANFIS) was evaluated based on mean absolute percentage error (MAPE) and mean error (ME). Results showed that the LC-WT-ANFIS model performed better than the traditional model. Therefore, the proposed forecasting model is capable of projecting mortality rates.

Keywords: ANFIS; forecast; macroeconomic; mortality; Lee–Carter model; wavelet

ABSTRAK

Kajian kematian adalah penting dalam menentukan status kesihatan dan komposisi demografi populasi. Rangka kerja pemodelan Lee–Carter (LC) diperluaskan untuk menggabungkan pemboleh ubah makroekonomi yang mempengaruhi kematian, terutamanya dalam peramalan. Sumbangan utama kertas ini adalah seperti berikut. Pertama, model baharu (LC-WT-ANFIS) yang menggunakan sistem inferens kabur berasaskan rangkaian adaptif (ANFIS) telah dicadangkan bersama dengan model spektrum tak linear bagi transformasi gelombang kecil diskret bertindih maksimum (MODWT) yang merangkumi lima fungsi matematik, iaitu, Haar, Daubechies (d4), least square (la8), best localization (bl14) dan Coiflet (c6) untuk meningkatkan ketepatan ramalan model LC. Data kematian tahunan telah dikumpulkan dari lima negara (Australia, England, Perancis, Jepun dan Amerika Syarikat) dari tahun 1950 hingga 2016. Kedua, keluaran dalam negara kasar (GDP), kadar pengangguran (UR) dan kadar inflasi (IF) dipilih sebagai nilai input mengikut korelasi dan regresi berganda. Pemboleh ubah input bagi kajian ini diperoleh dari World Bank dan Datastream, manakala pemboleh ubah output dikumpulkan daripada kadar kematian dalam *Human Mortality Database*. Akhir sekali, unjuran log kadar kematian model LC dibandingkan dengan penapis gelombang kecil dan model tradisional LC. Prestasi model yang dicadangkan (LC-WT-ANFIS) dinilai dari segi ralat peratusan mutlak min (MAPE) dan ralat min (ME). Keputusan kajian menunjukkan bahawa prestasi LC-WT-ANFIS adalah lebih baik daripada model tradisi. Oleh itu, model ramalan yang dicadangkan mampu mengunjurkan kadar kematian.

Kata kunci: ANFIS; gelombang kecil; kematian; makroekonomi; model Lee–Carter; ramalan

Introduction

Longevity risk and greater correlation amongst closely related populations such as gender, race, states, and countries are the key difficulties that industrialised countries face (Nor, Yusof & Norrulashikin 2021). The risk of death, particularly the longevity risk, has recently piqued the scientific community's interest, as the insurance industry struggles with product pricing. The risk of death is calculated using a future mortality rate (Brouhns, Denuit & Vermunt 2002). Insurance premiums are calculated using demographic and financial data. Forecasting future mortality trends is strongly determined by forecasting assumptions (Atsalakis et al. 2007). Actuaries have been developing enhanced methodologies for forecasting future mortality to manage longevity risk and obtain a more precise life expectancy projection (Nigri et al. 2019).

The extrapolative stochastic model developed by Lee and Carter (1992), known as the LC model, has been widely used due to its simplicity and robustness. The LC model was developed to model mortality data in the USA from 1933 to 1987, and it has since been used as a benchmark for analysing all-cause and cause-specific mortality data from a wide variety of countries and periods (Girosi & King 2007). There are a few deficiencies in the LC model, which prompted numerous modifications and extensions. Some well-known extensions to the LC model, include Booth–Maindonald–Smith (2002), Li–Lee (2005), Renshaw–Haberman (2006) and Hyndman–Ullah (2007) models. It is necessary to forecast the time-dependent mortality index in the LC model, and this is achieved by using the autoregressive integrated moving average (ARIMA). Lee and Carter (1992) proposed a random walk with drift in a univariate ARIMA (p,d,q) time series model. However, the LC model may produce unreliable estimates for certain age groups (Winiowski et al. 2015). Nigri et al. (2019) asserted that the standard ARIMA has a limited capacity for detecting unknown and unrecognised patterns in future mortality trends over time.

Hainaut and Denuit (2020) suggested using wavelets to address the LC model's limitations. The wavelet transform (WT) is an effective tool for describing death rate trends over time, and it is frequently used in time series analysis since it can capture period dynamics and decompose time series into a linear combination of different frequencies. WT is also capable of decomposing signals into wavelets. Daubechies (1992) and Mallat (1989) were the first to apply the discrete WT (DWT) to statistically analyse nonstationary and nonlinear time series. The maximal overlap DWT (MODWT) is capable of processing a wide variety of wavelet filters, including Haar, Daubechies, Coiflets, least asymmetric and best-localised wavelets filters. The MODWT is a DWT modification that omits the subsampling step, resulting in more information in the resulting wavelet and scaling coefficients. The MODWT was chosen for this study because it allows for the retention of downsampled values at each level of decomposition and is well defined for all sample sizes (Cornish, Bretherton & Percival 2006). Yaacob et al. (2021) used a variety of MODWT filters to model and forecast the LC model's time-dependent mortality index, including least squares (la8), best localised (bl14) and Coiflet (c6).

According to Zhang et al. (1998), a network's forecasting ability is affected by network structure, training method, and sample data. Therefore, incorporating wavelet, fuzzy and artificial neural network (ANN) techniques will aid in forecast accuracy. The neural network model has recently gained enormous popularity in time series and forecasting problems. When used for time series forecasting, the neural networks become nonlinear, data-driven, nonparametric and flexible (Yan 2012). Many studies have been conducted on the effect of ANN on mortality prediction. Atsalakis et al. (2007), for example, used the ANFIS model to forecast mortality rates. Hong et al. (2021) used the LC model and two machine learning approaches, random forest and ANN, to forecast death rates in Malaysia. Nigri et al. (2019) enhanced the predictive capacity of the LC model by incorporating it into a recurrent neural network architecture with long short-term memory. The MODWT nonlinear spectrum model was proven to be effective in modelling and improving the forecasting accuracy of data patterns from the Saudi stock market by combining it with ANFIS (Alenezy et al. 2021).

According to the literature study, no research has been conducted to date on the usefulness of MODWT wavelet filters and ANFIS when mortality-related parameters are included in the LC model incorporating macroeconomics variables with the goal of modelling and improving forecast accuracy of mortality trends. Further research is required into the comparative applications of MODWT functions such as Haar, db, la8, bl14, and c6 in conjunction with the fitting of the ANFIS model in modelling and forecasting mortality rates with macroeconomic variables affecting mortality rates. Hanewald, Post and Gründl (2011) developed a dynamic asset-liability model by connecting macroeconomic fluctuations to the LC model's mortality index and

Hanewald (2011) examined the LC model's response to macroeconomic volatility. According to French (2014), mortality rates among populations may be related, citing economic literature on technology and knowledge diffusion. Prior research has shown that accounting for real-world variations in variables such as gross domestic product (GDP), health expenditure and lifestyle helps explain mortality declines and improves mortality rate forecasting (French & O'hare 2014). Boonen and Li (2017) improved the Li and Lee (2005) model's mortality forecast by incorporating the principal components of real GDP per capita.

The purpose of this study was to improve the forecasting accuracy of the LC model using macroeconomic data on mortality and forecasting future changes in the time-dependent mortality index. The hybrid model LC-WT-ANFIS is proposed to project log death rates. It was then compared to the accuracy of the original LC model using root mean error (ME) and mean absolute percentage error (MAPE). This paper is structured as follows: The next section discusses the methods, the subsequent section summarises the findings and the last section concludes the study.

Materials and Methods

THE DATA SET

In this study, the mortality data for Australia, England, France, Japan and the USA were obtained from the Human Mortality Database (HMD 2020). The data includes central mortality rates and mid-year populations by individual years up to 110 years of age, which spans the years 1950 to 2016. Individuals over the age of 85 were grouped as 95+ to avoid erratic rates for this age group (Booth et al. 2006). The factors that affect the mortality of each country, such as GDP, inflation rate and unemployment rate, were obtained from the World Bank (2021) and Datastream (2020). The choice of factors that affect a country's mortality rate in this study is also based on the availability of such data. Table 1 shows the periods and factors used to study mortality for different countries.

TABLE 1. Data on total period mortality in each country and macroeconomic variables affecting mortality rates

Country	Year	Macroeconomic variables
Australia	1950–2016	• GDP • Inflation rate
England	1950–2016	• GDP • Inflation rate
France	1950–2016	• GDP • Inflation rate
Japan	1950–2016	• GDP • Unemployment rate
USA	1950–2016	• GDP • Unemployment rate • Inflation rate

Each data set is divided into two periods: A training set and a testing set. To ensure a fair comparison, the data set is divided into two periods using the 80/20 train-test ratio (Alenezy et al. 2021).

THE LEE–CARTER MODEL

The LC model (1992) is as follows:

$$\ln(m(x,t)) = a(x) + b(x)k(t) + \varepsilon(x,t), \quad x = 0,\dots,x_m \quad t = 1,\dots,t_n \tag{1}$$

where $m(x,t)$ is the age-specific death rate for the x interval and the year t, $k(t)$ is the mortality index in the year t, $a(x)$ is the average age-specific mortality, $b(x)$ is a deviation in the mortality due to changes in the $k(t)$ index, and $\varepsilon(x,t)$ is the residual at age x and time t which is independent and identically distributed following a normal distribution $N(0, \sigma^2)$ with mean 0 and variance σ_ε^2. Lee and Carter (1992) estimated $a(x)$ as the average of $\ln(m(x,t))$ over time, and the $b(x)$ and $k(t)$ are estimated by singular value decomposition (SVD). Since the solution cannot be unique, the following constraints were imposed (Lee & Carter 1992):

$$\sum_{x_0}^{x_m} b(x) = 1 \text{ and } \sum_{t_1}^{t_n} k(t) = 0. \quad (2)$$

SVD is applied to the following matrix $Z(x,t)$ (Wang 2007):

$$Z(x,t) = ln(m(x,t)) - \hat{a}(x) \quad (3)$$

produced,

$$ULV' = SVD(Z(x,t)) = L(1)U(x,1)V(t,1) + \cdots + L(X)U(x,X)V(t,X) \quad (4)$$

where U represents the age component, L is the singular value and V represents the time component. The first step in forecasting mortality via LC model is estimating $a(x)$, $b(x)$ and $k(t)$ using historical age specific mortality rates. The estimates of $\hat{a}(x)$ can be obtained by finding the average over time of $ln(m(x,t))$. The estimates of $\hat{b}(x) = U(x,1)$ and $\hat{k}(t) = L(1)V(t,1)$ can be obtained by approximating the first term.

ARIMA MODEL FOR $k(t)$

A random walk with drift, ARIMA (0,1,0) model, which was first used by Lee and Carter (1992), is widely used to forecast the time-varying index, $k(t)$.

$$\hat{k}(t) = \hat{k}(t-1) + \theta + \varepsilon(t) \quad (5)$$

where θ is the drift parameter and $\varepsilon(t)$ are the normally distributed error terms with mean 0 and variance σ_k^2. The forecasted values of the adjusted $k(t)$ and the estimated $a(x)$ and $b(x)$ are substituted into equation (1) to get the forecasted values of $m(x,t)$.

$$\hat{m}(x,n+h) = \hat{m}(x,n)exp\left\{\hat{b}(x)\left(\hat{k}(n+h) - \hat{k}(n)\right)\right\},$$
$$h = 1,2,\ldots \quad x = 1,2,\ldots,n \quad (6)$$

where n is the last year from which data are available; h is the forecast horizon; and x represents the age group (Andreozzi, Blaconá & Arnesi 2011).

WAVELET TRANSFORM FORMULA

WT is built on the Fourier transform, which depicts any function as the sum of sine and cosine functions. WT is a function of time t that obeys a basic rule known as the admissibility condition (Mehra 2018):

$$C_\varphi = \int_{-\infty}^{\infty} \frac{|\varphi(f)|^2}{f} df < \infty, \quad (7)$$

where $\varphi(f)$ is the Fourier transform and a function of the frequency f of a father wavelet $\phi(t)$. The smooth and low-frequency components of a signal are generated by the father wavelet, while the detailed and high-frequency components are generated by the mother wavelet. The following equations represent the father and mother wavelets, respectively, where $j = 1, 2, 3, \ldots, J$ in a J-level wavelet decomposition:

$$\phi_{j,k} = 2^{\left(\frac{-j}{2}\right)}\phi\left(t - \frac{2^j k}{2^j}\right),$$
$$\varphi_{j,k} = 2^{\left(\frac{-j}{2}\right)}\varphi\left(t - \frac{2^j k}{2^j}\right), \quad (8)$$

where J denotes the maximum scale sustainable by the number of data points. The father and mother wavelets satisfy:

$$\int_{-\infty}^{\infty} \phi(t)dt = 1,$$
$$\int_{-\infty}^{\infty} \varphi(t)dt = 0. \quad (9)$$

A function that is an input represented by wavelet transform can be built in any time-series data as a sequence of projections onto father and mother wavelets indexed by $\{k\} = 2j$ where $k = \{0, 1, 2, \ldots\}$, and $\{S\} = 2j$ where $\{j = 1, 2, 3, \ldots J\}$. The analysis of real discretely sampled data necessitates the construction of a lattice to perform calculations. Mathematically, it is convenient to use a dyadic expansion, as shown in equation (9). The expansion coefficients are given by the projections:

$$S_{j,k} = \int_{-\infty}^{\infty} \phi_{j,k} f(t)dt,$$
$$d_{j,k} = \int_{-\infty}^{\infty} \varphi_{j,k} f(t)dt. \quad (10)$$

The wavelet approximation coefficients to $f(t)$ that lead to $k(t)$ in the wavelet framework of Lee and Carter (1992) is defined by:

$$k(t) = S_j(t) + D_j(t)$$
$$= \sum_{-\infty}^{\infty} S_{j,k}\phi_{j,k}(t) + \sum_{-\infty}^{\infty} d_{j,k}\phi_{j,k}(t) \quad (11)$$

WT is used to calculate the wavelet approximation coefficient in (Mehra 2018) for a discrete signal, where $S_j(t)$ and D_j (t) introduce the smooth and detailed coefficients, respectively. The smooth coefficients emphasise the most critical features of the data, while the detailed coefficients detect the main features in the dataset (Mehra 2018; Percival & Walden 2000).

ANFIS MODEL

The ANFIS utilises both fuzzy logic and ANN (Harandizadeh & Armaghani 2021). For training, the ANFIS employs the ANN learning algorithm. ANN's operations consist of forward and backward steps from the ANFIS learning algorithm. The forward step consists of five layers.

Figure 1 depicts the ANFIS architecture, which has two inputs and one output. Based on Figure 1, the fuzzy inference system under consideration has two inputs (y_1, y_2) and one output (z) for explanation simplification. Note that the input y_1 and y_2 represent the macroeconomic data, and the output z represents the log mortality rates. A standard rule base of fuzzy if-then rules for the first order of Sugeno fuzzy model can

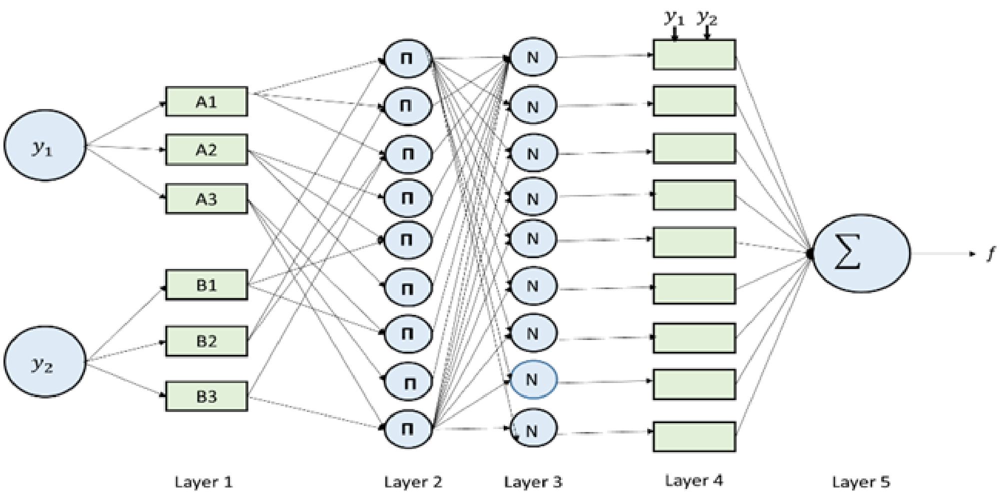

FIGURE 1. ANFIS architecture of two inputs and one output with four rules

be expressed as follows: If y_1 is A_1 and y_2 is B_1, then $f_1 = p_1 y_1 + q_1 y_2 + r_1$, where p,r and q are denoted as linear output parameters.

Layer-1 In this layer, each node is a square node with a node function.

$$O_{1,i} = \mu_{Ai}(y_1), \text{ for } i = 1,2,3 \text{ and } O_{1,i} = \mu_{Bi\text{-}3}(y_2) \quad \text{for } i = 4,5,6 \tag{12}$$

where y_1 and y_2 are denoted as inputs to node i, and A_i and B_i are defined as linguistic labels for inputs. In other words, O(1,i) is defined as the membership function of A_i and B_i. Typically, $\mu_{Ai}(y_1)$ and $\mu_{Bi}(y_2)$ are chosen to be bell-shaped with a maximum value of 1 and minimum value of 0, such as $\mu_{Ai}(y_1)$ and $\mu_{Bi-3}(y_2) = exp\,(\frac{-(y_i - c_i)}{a_i})^2)$, where the set of parameters a_i, c_i. These parameters are referred to as premise parameters in this layer. The fuzzification process converts crisp values into linguistic values by using the Gaussian function as the shape of the membership function.

Layer-2 Each node in this layer is a labelled Π that multiplies the incoming signals and outputs the product. For instance,

$$O_{2,i} = w_i = \mu_{Ai}(y_{-1}). \quad \mu_{Bi\text{-}3}(y_2), i = 1,2,3,\ldots,9 \tag{13}$$

Each node output describes the firing strength of a rule. In this layer, the t-norm operator (the AND operator) is used by the inference stage.

Layer-3 Each node in this layer is a circle node called N. The i^{th} node measures the ratio of the i^{th} rule firing strength to the sum of all rules firing strengths:

$$O_{3,i} = \overline{w}_i = \frac{w_i}{(w_1 + w_2 + \cdots + w_9)}, \qquad i = 1,2,3,\ldots,9 \quad (14)$$

In short, the ratio of the strengths of the rules is calculated in this layer.

Layer-4 In this layer, each node i is a square node with the following node function:

$$O_{4,i} = \overline{w}_i . f_i = \frac{w_i . (p_i y_1 + q_i y_2 + r_i)}{\sum_i w_i}, \quad i = 1,2,3,\ldots,9 \quad (15)$$

where $\overline{w}_i$ is the output of layer 3 and $\{p_i, q_i, r_i\}$ is the parameter set. Parameters in this layer are referred to as consequent parameters. This layer measures the parameters for the subsequent parts.

Layer-5 A circle node named Σ calculates the overall output of this layer as the summation of all incoming signals.

$$O_{5,i} = \text{overall output} = \sum_i \overline{w}_i, f_i = \frac{\sum_i w_i, f_i}{\sum_i w_i}. \qquad (16)$$

The backward step is a database estimation technique composed of two parts: the antecedent part contains the membership function parameters and the subsequent part contains the linear equation coefficients. Since the Gaussian function is used as the membership function in this process, two parameters of this function, namely the mean and variance, are optimised. In this step, parameter learning is carried out using the least square technique (Figure 2).

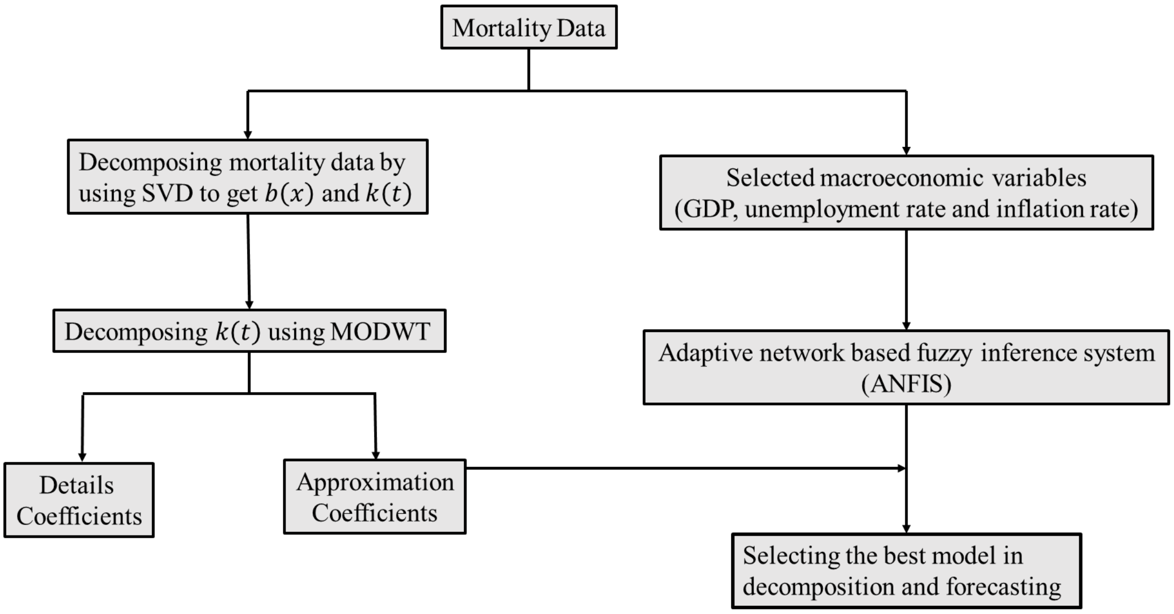

FIGURE 2. The flowchart of the MODWT with ANFIS

PERFORMANCE MEASURES

Several accuracy criteria, including MAPE, mean absolute error (MAE) and ME, were used as follows:

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right| \qquad (17)$$

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|A_t - F_t| \qquad (18)$$

$$ME = \frac{1}{n}\sum_{t=1}^{n}(A_t - F_t) \qquad (19)$$

The MAPE criterion, also known as the mean absolute percentage deviation (MAPD), is a statistical criterion for forecasting accuracy (Alenezy et al. 2021). It is always expressed in percentage and calculated using equation (17), where A_t is the actual value, F_t denotes the forecasted value and n denotes the sample size. The absolute value of each forecasted point in time is summed and divided by the number of fitted points in this equation. Also, equation (18) defines the MAE, while equation (19) defines the ME. As depicted on Figure 2, the following steps are taken based on the research framework from the study: First, estimate $a(x)$ as the average $ln(m(x,t))$ of overtime and decompose the mortality data using SVD to determine $b(x)$ and $k(t)$. Second, decompose the parameter of $k(t)$ using MODWT's functions to obtain approximation and detailed coefficients. The approximation coefficients are used in the forecasting process since they have the main components of the data. Third, the ANFIS mechanism requires the dependent variable (the approximation coefficients $k(t)$) and independent variable (GDP, UR and IF). Fourth, the best model is selected based on the ME and MAPE.

Results and Discussion

CORRELATION AND REGRESSIONS RESULTS

The input variables must be carefully chosen to avoid variable multicollinearity. Table 2 illustrates the correlation between the macroeconomic data (input variables) and mortality rate (output variable) $k(t)$. There are weak correlation (< 50%) between input variables for all countries. The absence of perfect multicollinearity is referred to as 'no multicollinearity', which is defined as an exact (non-stochastic) linear relationship between two or more input macroeconomic variables. The LC model is appropriate for testing all macroeconomic factors. Table 3 summarise the multiple regression between macroeconomic variables affecting mortality in each country studied and mortality index. According to Table 3, the macroeconomics variables explain more than 90% of the $k(t)$ variance that analysed for all populations except for the Japanese population, which accounted for 75.9% of the variance. In an attempt to optimise forecasting of future mortality rates, the LC model is then modified to include GDP, inflation and UR.

TABLE 2. The correlations between the input and the output variables

Countries		$k(t)$	GDP	IF	UR
	$k(t)$	1	-0.9850	0.1390	-0.3780
USA	GDP		1	-0.0870	0.2480
	IF			1	0.1820
	UR				1
Australia	$k(t)$	1	-0.9682	0.3855	-
	GDP		1	-0.2909	-
	IF			1	-
England	$k(t)$	1	-0.9680	0.3550	-
	GDP		1	-0.2392	-
	IF			1	-
France	$k(t)$	1	-0.9412	0.4779	-
	GDP		1	-0.0279	-
	IF			1	-
Japan	$k(t)$	1	-0.9770	-	-0.1266
	GDP		1	-	0.1775
	UR			-	1

FORECASTING RESULTS

Table 4 shows that the wavelet filters with Haar, d4, la8, bl14, and c6 wavelet coefficients were used to evaluate the accuracy of two ARIMA models representing $k(t)$. The number of autoregressive terms and the number of non-seasonal differences required for stationarity were varied to find an ARIMA (p,d,q) model that best represents $k(t)$. A good fit for the countries under consideration was found with $p = 5$ and $d = 1$. The bl14 filter has the smallest MAE and MAPE for all populations. This finding is consistent with Yaacob et al. (2021), who report that the bl14 filter best fits the log of mortality rates for males and females, and the total population of Australia, British, French, Japan and the USA.

TABLE 3. The multiple *r* squared between macroeconomic factors and mortality index

Countries	Variables	Estimate	Std.Error	t-statistic	Sign.	Performance test	Estimate
						R-squared	0.9804
	Intercept	141.5541	2.7478	51.5160	<2e-16***	R-square-Adjusted	0.9795
USA	GDP	-17.6340	0.3484	-50.6090	<2e-16***	F-statistics	1050.0
	IF	87.2424	15.5233	5.6200	4.65E-07**	Sign.	<2.2e-16
	UR	-66.5314	28.9685	-2.2970	0.025*		
Australia	Intercept	0.0000	1.0340	0.0000	1	R-squared	0.9493
	GDP	-34.6000	1.0890	-31.7760	<2e-16***	R-square-Adjusted	0.9477
	IF	4.1970	1.0890	3.8540	0.000272***	F-statistics	598.60
						Sign.	<2.2e-16
England	Intercept	0.0000	0.7988	0.0000	1	R-squared	0.9527
	GDP	-27.7200	0.8289	-33.4400	<2e-16***	R-square-Adjusted	0.9512
	IF	3.8790	0.8289	4.6800	<1.53e-05***	F-statistics	644.50
						Sign.	<2.2e-16
France	Intercept	0.0000	1.0980	0.0000	1	R-squared	0.9361
	GDP	-30.6700	1.1520	-26.6210	<2e-16***	R-square-Adjusted	0.9341
	IF	8.1780	1.1520	7.0980	1.25E-09***	F-statistics	468.70
						Sign.	<2.2e-16
Japan	Intercept	23.6503	8.8061	2.6860	0.00921**	R-squared	0.7599
	GDP	-23.8246	9.4538	-2.5200	0.01424*	R-square-Adjusted	0.7524
	UR	-0.7397	0.0883	-8.3730	7.09E-12***	F-statistcs	101.30
						Sign.	<2.2e-16

Note: '***' 0.001,'**'0.01, and '** 0.05

TABLE 4. Accuracy measures based on MAE and MAPE for total population by country (smallest values are bolded)

Country	MAE						MAPE					
	Haar	d4	la8	bl14	c6	ARIMA	Haar	d4	la8	bl14	c6	ARIMA
Australia	2.801	2.759	1.984	**1.782**	2.485	2.347	25.816	29.280	8.931	**8.998**	12.423	19.459
England	3.193	2.336	1.570	**1.488**	2.221	1.953	28.693	19.231	22.129	**14.878**	15.495	21.473
France	2.857	2.621	1.897	**1.833**	2.582	2.206	19.090	17.507	46.790	**13.623**	19.316	17.793
Japan	5.560	4.032	2.855	**2.619**	3.769	3.280	27.027	19.814	21.307	**7.239**	14.005	18.671
USA	1.838	1.606	1.203	**1.143**	1.652	1.468	43.721	9.985	7.807	**6.026**	12.935	18.099

The bl14 filter was chosen for the LC-WT-ANFIS model after an evaluation of the results in Table 4. As shown in Table 5, the LC-WT-ANFIS consistently outperforms its contenders in four countries (England, France, Japan and the USA). However, the LC model appeared to be the most effective for the Australian population. In general, the LC-WT-ANFIS model outperforms the original LC model in terms of phase properties. As an example, Figure 3 depicts the observed and forecasted log death rates for the USA population in 2011. In comparison to the original LC, Figure 3 shows that the estimation of log death rate values using wavelet and ANFIS provides a good fit for the USA in 2011. The log death rates for the LC-WT-ANFIS model closely follow the observed values across all age groups.

TABLE 5. Accuracy measures based on ME and MAPE for LC-WT-ANFIS vs. LC and LC-ANFIS by country

Country	ME			MAPE		
	LC	LC-ANFIS	LC-WT-ANFIS	LC	LC-ANFIS	LC-WT-ANFIS
Australia	**0.073**	0.849	0.096	**1.931**	18.294	2.247
England	0.094	0.156	**0.046**	2.424	3.584	**1.667**
France	0.133	0.301	**0.014**	2.606	6.622	**2.763**
Japan	0.139	0.127	**0.093**	2.644	2.459	**1.893**
USA	0.087	0.093	**0.006**	2.105	2.199	**1.263**
Average	0.105	0.305	**0.051**	2.342	6.631	**1.966**

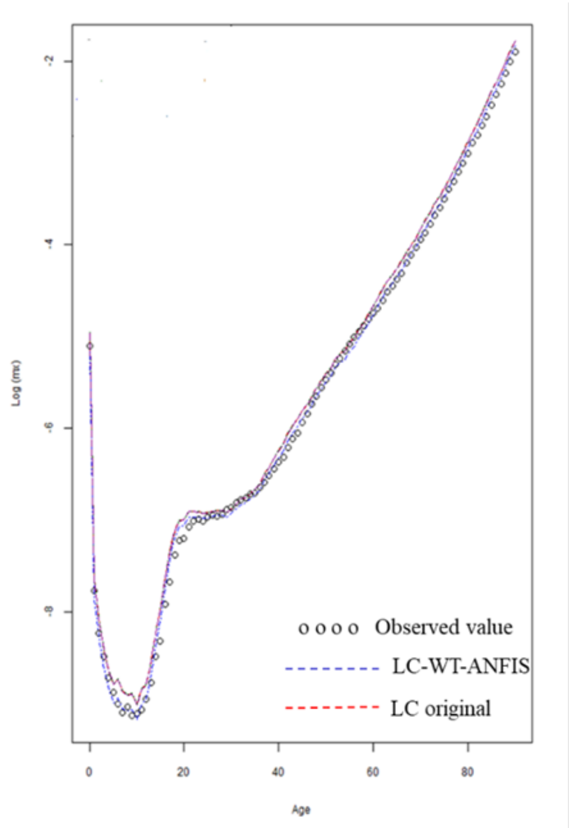

FIGURE 3. Observed and predicted log death rates for the USA total population (the year 2011) using the wavelet and ANFIS integration with the LC model

Conclusion

In this study, we have proposed a new model (LC-WT-ANFIS) successfully to enhance the forecasting accuracy of the LC model. The LC model is extrapolative in nature, and it makes no attempt to account for medical, behavioural or societal factors that affect mortality change (Lee & Carter 1992). Prior research has shown that accounting for real-world variations in variables such as income (GDP), health expenditure and lifestyle helps explain mortality decline and improves mortality rate forecasting (French & O'hare 2014). In this paper, the LC model was fitted to the mortality data for five countries, namely Australia, England, France, Japan and the USA. Macroeconomic variables were incorporated into the forecasting process using a wavelet and an ANFIS model $k(t)$. The forecast performance of LC-WT-ANFIS model was then compared to that of a conventional LC model using ME and MAPE. Although these variables reflected general trends and were not age-specific, this model produced the best forecast results in four of the five countries studied. In addition, when the values of ME and MAPE of all countries are considered, it is shown that incorporating wavelet and ANFIS into the LC model outperforms the conventional LC model in forecasting future index mortality. Wavelet and fuzzy methods can be used to enhance the LC model to improve forecast accuracy as these approaches can handle a large number of design parameters and a long training period. The applicability of this strategy is still constrained by data availability and the stability of covariate projections.

Acknowledgements

The authors thank the anonymous referees for their useful suggestions.

References

Alenezy, A.H., Ismail, M.T., Wadi, S.A., Tahir, M., Hamadneh, N.N., Jaber, J.J. & Khan, W. A. 2021. Forecasting stock market volatility using hybrid of adaptive network of fuzzy inference system and wavelet functions. *Journal of Mathematics (Special Issue - Transport and Mixing Fluid Flow in Dynamical Systems)* 2021: Article ID. 9954341.

Andreozzi, L., Blaconá, M.T. & Arnesi, N. 2011. The Lee Carter method for estimating and forecasting mortality: An application for Argentina. In *International Symposium on Forecasting-2011, Prague Proceedings* 2050(5): 1–17.

Atsalakis, G., Nezis, D., Matalliotakis, G., Ucenic, C.I. & Skiadas, C. 2007. Forecasting mortality rate using a neural network with fuzzy inference system. *International Workshop on Applied Probability. Compi'egne, FR.*

Boonen, T.J. & Li, H. 2017. Modeling and forecasting mortality with economic growth: A multipopulation approach. *Demography* 54(5): 1921-1946.

Booth, H., Maindonald, J. & Smith, L. 2002. Applying Lee-Carter under conditions of variable mortality decline. *Population Studies* 56(3): 325-336.

Booth, H., Hyndman, R.J., Tickle, L. & De Jong, P. 2006. Lee-Carter mortality forecasting: A multi-country comparison of variants and extensions. *Demographic Research* 15: 289-310.

Brouhns, N., Denuit, M. & Vermunt, J.K. 2002. Measuring the longevity risk in mortality projections. *Bulletin of the Swiss Association of Actuaries* 2(1): 105-130.

Cornish, C.R., Bretherton, C.S. & Percival, D.B. 2006. Maximal overlap wavelet statistical analysis with application to atmospheric turbulence. *Boundary-Layer Meteorology* 119(2): 339-374.

Datastream 2020. Thomson Reuters Datastream. https://www.library.ucsb.edu/node/7813. Accessed on November 2020.

Daubechies, I. 1992. *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics. https://doi.org/10.1137/1.9781611970104

French, D. 2014. International mortality modelling - An economic perspective. *Economics Letters* 122(2): 182-186.

French, D. & O'Hare, C. 2014. Forecasting death rates using exogenous determinants. *Journal of Forecasting* 33(8): 640-650.

Girosi, F. & King, G. 2007. *Understanding the Lee-Carter Mortality Forecasting Method*. Gking. Harvard. Edu.

Hainaut, D. & Denuit, M. 2020. Wavelet-based feature extraction for mortality projection. *ASTIN Bulletin: The Journal of the IAA* 50(3): 675-707.

Hanewald, K. 2011. Explaining mortality dynamics: The role of macroeconomic fluctuations and cause of death trends. *North American Actuarial Journal* 15(2): 290-314.

Hanewald, K., Post, T. & Gründl, H. 2011. Stochastic mortality, macroeconomic risks and life insurer solvency. *The Geneva Papers on Risk and Insurance-Issues and Practice* 36(3): 458-475.

Harandizadeh, H. & Armaghani, D.J. 2021. Prediction of air-overpressure induced by blasting using an ANFIS-PNN model optimized by GA. *Applied Soft Computing* 99: 106904.

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

Human Mortality Database. 2020. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). https://www.mortality.org/

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

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

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

Mallat, S.G. 1989. A theory for multiresolution signal decomposition: The wavelet representation. In *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11(7): 674-693. doi: 10.1109/34.192463

Mehra, M. 2018. *Wavelets Theory and Its Applications*. Singapore: Springer.

Nigri, A., Levantesi, S., Marino, M., Scognamiglio, S. & Perla, F. 2019. A deep learning integrated Lee–Carter model. *Risks* 7(1): 33.

Nor, S.R.M., Yusof, F. & Norrulashikin, S.M. 2021. Coherent mortality model in a state-space approach. *Sains Malaysiana* 50(4): 1101-1111.

Percival, D.B. & Walden, A.T. 2000. *Wavelet Methods for Time Series Analysis (Vol. 4)*. Cambridge: Cambridge University Press.

Renshaw, A.E. & Haberman, S. 2006. A cohort-based extension to the Lee–Carter model for mortality reduction factors. *Insurance: Mathematics and Economics* 38(3): 556-570.

Wiśniowski, A., Smith, P.W., Bijak, J., Raymer, J. & Forster, J.J. 2015. Bayesian population forecasting: Extending the Lee-Carter method. *Demography* 52(3): 1035-1059.

Wang, J.Z. 2007. *Fitting and Forecasting Mortality for Sweden: Applying the Lee-Carter Model*. Mathematical Statistics, Stockholm University.

World Bank. 2021. *World Development Indicators.* https://data.worldbank.org/ Accessed on January 2021.

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

Yan, W. 2012. Toward automatic time-series forecasting using neural networks. *IEEE Transactions on Neural Networks and Learning Systems* 23(7): 1028-1039.

Zhang, G., Patuwo, B.E. & Hu, M.Y. 1998. Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting* 14(1): 35-62.

[*]Corresponding author; email: aityqah@uitm.edu.my