

Research Article

A Comparative Analysis of the Forecasted Mortality Rate under Normal Conditions and the COVID-19 Excess Mortality Rate in Malaysia

Robiaatul Adawiah Edrus ^{1,2}, Zailan Siri ¹, Mohd Azmi Haron,¹
Muhammad Aslam Mohd Safari ³ and Mohammed K. A. Kaabar ^{1,4}

¹Institute of Mathematical Sciences, Faculty of Science, University of Malaya, 50603 Kuala Lumpur, Malaysia

²School of Economics, Finance, and Banking, College of Business, Universiti Utara Malaysia, 06010 Sintok, Kedah, Malaysia

³Department of Mathematics and Statistics, Faculty of Science, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia

⁴Gofa Camp Near Gofa Industrial College and German Adebabay, Nifas Silk-Lafto 26649, Addis Ababa, Ethiopia

Correspondence should be addressed to Zailan Siri; zailansiri@um.edu.my and Mohammed K. A. Kaabar; mohammed.kaabar@wsu.edu

Received 4 February 2022; Accepted 14 April 2022; Published 6 May 2022

Academic Editor: Miaochoao Chen

Copyright © 2022 Robiaatul Adawiah Edrus et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

SARS-CoV-2, known as COVID-19, has affected the entire world, resulting in an unexpected death rate as compared to the death probability before the pandemic. Prior to the COVID-19 pandemic, death probability has been assessed in a normal context that is different from those anticipated during the pandemic, particularly for the older population cluster. However, there is no such evidence of excess mortality in Malaysia to date. Therefore, this study determines the excess mortality rate for specific age groups during the pandemic outbreak in Malaysia. Before determining the excess mortality rate, this study aims to establish the efficiency of various parametrized mortality models in reference to the data set before the pandemic. This study employs the hold-out, repeated hold-out, and leave-one-out cross-validation procedures to identify the optimal mortality law for fitting the mortality data. Based on the goodness-of-fit measures (mean absolute percentage error, mean absolute error, sum square error, and mean square error), the Heligman-Pollard model for men and Rogers Planck model for women are considered as the optimal models. In assessing the excess mortality, both models favour the hold-out technique. When the COVID-19 mortality data are incorporated to forecast the mortality rate for people aged 60 and above, there is an excess mortality rate. However, the men's mortality rate appears to be delayed and more prolonged than the women's mortality rate. Consequently, the government is recommended to amend the existing policy to reflect the post COVID-19 mortality forecast.

1. Introduction

The pandemic COVID-19, known as Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), is a highly contagious disease, affecting human's respiratory, hepatic, gastrointestinal, and neurological systems has made a significant negative impact on the entire world [1]. This pandemic has the potential to have an unparalleled influence around the globe, as the number of deaths continues to rise day by day. Malaysia has almost 17,000 COVID-19 daily cases on August 9, 2021, with 212 deaths, making it the Southeast Asian country with the third-highest cumulative cases after

Indonesia and the Philippines [2]. Despite the Malaysian government's implementation of a nationwide movement control order (MCO) on March 18, 2020, this pandemic substantially impacted humanity, particularly the number of deaths. Aside from the current overall population, predicting the estimated size of future populations requires understanding mortality rates. Although the death rate is an unexpected occurrence, it profoundly affects demographers, insurance companies, legislators, and the government. Therefore, all stakeholders should be aware of changes in mortality rates to continue providing the highest quality of service and deliver the best resources for future planning.

According to the Department of Statistics Malaysia, Malaysians aged 60 and more are expected to reach 3.5 million in 2020, accounting for 10.7% of the country's total population. However, the current COVID-19 pandemic is anticipated to lower the number of senior citizen populations, as COVID-19 mortality rates are higher in countries with a higher proportion of the population aged 65 and over [3,4]. Consequently, the biggest concern articulated prior to the pandemic outbreak concerning an ageing society should be reconsidered, as the mortality rate of elderly citizens has a significant influence during the pandemic. In terms of post-pandemic support for Malaysia's ageing population, Jamaluddin et al. [5] proposed the following steps: (i) reconsider whether the existing social security framework meets the needs of the population post-pandemic and in the light of an ageing population and (ii) reformation of the existing social security framework and a holistic execution plan to be contemplated in light of the projected economic outlook. Furthermore, Chung et al. [6] emphasised the importance of assisting older adults in Malaysia in adapting to their new surroundings and participating actively in social and economic activities.

COVID-19 is exerting a detrimental influence on humans all across the world, resulting in higher mortality and shorter life expectancy. Nevertheless, age-specific mortality rates have yet to be investigated in order to determine whether Malaysia has been experiencing an excess in mortality. Excess mortality rates occur when the total number of deaths during a crisis (i.e., global pandemic) exceeds what would be expected in normal circumstances [7]. If an excess mortality rate was identified, the government, demographers, insurance providers, and policymakers might revise their existing strategy for dealing with the current population numbers. As a result of this event, this study determines the excess mortality for a specific age group in Malaysia during the COVID-19 pandemic. Before investigating the excess mortality rate, this study investigates the optimum resampling method and parametrized mortality model to forecast the mortality rate under normal conditions. Parametrized functions, commonly known as mortality laws, are one-factor models that aim to represent the age pattern of death parsimoniously and they have advantages in the smoothness of predicted rates over time [8]. On the other hand, resampling methods are required as one of the quantitative tools to analyze the existing and anticipated conditions of mortality patterns [9].

As a consequence of the pandemic's domino effect, the number of deaths is increasing tremendously. Therefore, this study aims to forecast the mortality rate in normal conditions by employing various resampling methods to fit the Malaysian mortality rate in identifying the best model and resampling method. The main contribution of this research is twofold: (i) identifying the best resampling method and mortality model for forecasting the Malaysian mortality rate in normal conditions and (ii) determining the existence of an excess mortality rate during the pandemic. In the next section, the parametric mortality models employed in this study, as well as their development and application, are briefly described. Several resampling techniques for fitting

mortality are briefly discussed in the next section. The next section utilizes goodness-of-fit measures to identify the best model for forecasting the mortality rate under normal conditions in Malaysia. The next section also determines the existence of excess mortality during COVID-19 before concluding this research. Finally, the optimal mortality model and resampling technique to forecast normal and excess mortality rates for specific age groups during a pandemic are summarised in the final section.

2. Parametrized Mortality Models

This section describes the parametric mortality models involved to fit the mortality rates. The mortality models describe the process of individuals in a population dying off over the timeframe of a significant portion of their lives [10]. The mortality rate formulation is expressed in Equation 1.

$$m_{x,t} = \frac{D_{x,t}}{E_{x,t}}, \quad (1)$$

where

- (i) $m_{x,t}$: the mortality rate for a specific age group of x at a specific time t ,
- (ii) $D_{x,t}$: number of deaths for a specific age group of x at a specific time t , and
- (iii) $E_{x,t}$: exposures for a specific age group of x at a specific time t .

The number of populations at the beginning and end for a certain age group at a specific time is averaged to compute exposures. The number of deaths and number of populations are extracted for the Department of Statistics, Malaysia. The "MortalityLaws" package by Pascariu [10], which was developed in the R software, is utilized to fit the mortality rate. For both men and women, the mortality rate is provided in a five-year age group ranging from 0 to 84 years old. The data set covers the years 1995 to 2018 and fits the mortality rate in normal conditions. Table 1 shows a descriptive analysis for data characteristics used in this research study.

Table 2 describes the parametric mortality models to fit the Malaysian mortality rate. The HP model was developed by Heligman and Pollard [11], which consists of eight parameters with three terms: (i) the first term reflects the mortality rate during early childhood from 0 to 9 years old, (ii) the second term defines the adult mortality rate from 10 to 40 years old, and (iii) the third term illustrates the senescence mortality rate for ages over 40 years old. The Nigerian mortality rate was employed by Umar and Chukwudi [15] to investigate the performance of the HP and the Lee and Carter [16] models. Besides that, Silva et al. [17] applied the HP model to estimate life expectancy at birth in Mexico. Also, Kostaki [12] developed the KT model with nine parameters to eliminate a source of systematic error that affects the fit of the HP formula. The HP and KT models differ in the second term, which describes the spread of the accident hump to the left and right of its top [18]. The use of the KT model with mortality data [19–21] demonstrated that the cubic splines modification resulted

TABLE 1: Descriptive statistics.

Statistics	Age group	Male	Female	Age group	Male	Female
Mean	0-4	1.9189×10^{-03}	1.5988×10^{-03}	45-49	5.1574×10^{-03}	2.8259×10^{-03}
Median		1.7668×10^{-03}	1.4725×10^{-03}		5.0859×10^{-03}	2.7961×10^{-03}
Standard deviation		4.2349×10^{-04}	3.4035×10^{-04}		2.4775×10^{-04}	1.5721×10^{-04}
Variance		1.7935×10^{-07}	1.1584×10^{-07}		6.1382×10^{-08}	2.4714×10^{-08}
Mean	5-9	3.1772×10^{-04}	2.3448×10^{-04}	50-54	8.0693×10^{-03}	4.6852×10^{-03}
Median		2.8665×10^{-04}	2.1560×10^{-04}		7.9484×10^{-03}	4.5973×10^{-03}
Standard deviation		8.4148×10^{-05}	5.3021×10^{-05}		4.0884×10^{-04}	2.8896×10^{-04}
Variance		7.0810×10^{-09}	2.8112×10^{-09}		1.6715×10^{-07}	8.3497×10^{-08}
Mean	10-14	4.0113×10^{-04}	2.5182×10^{-04}	55-59	1.2566×10^{-02}	7.5739×10^{-03}
Median		3.8506×10^{-04}	2.3364×10^{-04}		1.2481×10^{-02}	7.3453×10^{-03}
Standard deviation		7.7224×10^{-05}	6.2299×10^{-05}		7.9368×10^{-04}	6.2704×10^{-04}
Variance		5.9636×10^{-09}	3.8812×10^{-09}		6.2992×10^{-07}	3.9317×10^{-07}
Mean	15-19	1.0895×10^{-03}	3.5803×10^{-04}	60-64	1.9787×10^{-02}	1.2390×10^{-02}
Median		1.0535×10^{-03}	3.4772×10^{-04}		1.8959×10^{-02}	1.1582×10^{-02}
Standard deviation		2.0656×10^{-04}	6.3557×10^{-05}		2.0697×10^{-03}	1.6288×10^{-03}
Variance		4.2666×10^{-08}	4.0395×10^{-09}		4.2837×10^{-06}	2.6529×10^{-06}
Mean	20-24	1.3148×10^{-03}	4.3422×10^{-04}	65-69	3.0495×10^{-02}	2.0521×10^{-02}
Median		1.2013×10^{-03}	4.1210×10^{-04}		3.0506×10^{-02}	2.0650×10^{-02}
Standard deviation		3.3529×10^{-04}	7.0974×10^{-05}		3.2912×10^{-03}	3.3641×10^{-03}
Variance		1.1242×10^{-07}	5.0374×10^{-09}		1.0832×10^{-05}	1.1317×10^{-05}
Mean	25-29	1.3885×10^{-03}	5.2769×10^{-04}	70-74	4.7377×10^{-02}	3.5058×10^{-02}
Median		1.2929×10^{-03}	5.0828×10^{-04}		4.6831×10^{-02}	3.5481×10^{-02}
Standard deviation		3.7518×10^{-04}	7.3722×10^{-05}		5.7506×10^{-03}	6.0824×10^{-03}
Variance		1.4076×10^{-07}	5.4349×10^{-09}		3.3070×10^{-05}	3.6995×10^{-05}
Mean	30-34	1.8474×10^{-03}	7.3459×10^{-04}	75-79	7.2970×10^{-02}	6.0594×10^{-02}
Median		1.9535×10^{-03}	7.2005×10^{-04}		6.9602×10^{-02}	5.9331×10^{-02}
Standard deviation		3.1984×10^{-04}	7.1331×10^{-05}		9.5451×10^{-03}	9.5346×10^{-03}
Variance		1.0230×10^{-07}	5.0881×10^{-09}		9.1109×10^{-05}	9.0909×10^{-05}
Mean	35-39	2.4806×10^{-03}	1.0957×10^{-03}	80-84	1.0441×10^{-01}	9.6537×10^{-02}
Median		2.5103×10^{-03}	1.0635×10^{-03}		9.7229×10^{-02}	9.6189×10^{-02}
Standard deviation		2.3528×10^{-04}	8.6181×10^{-05}		1.3292×10^{-02}	9.0265×10^{-03}
Variance		5.5355×10^{-08}	7.4271×10^{-09}		1.7668×10^{-04}	8.1478×10^{-05}
Mean	40-44	3.5092×10^{-03}	1.7307×10^{-03}			
Median		3.5147×10^{-03}	1.7259×10^{-03}			
Standard deviation		1.5476×10^{-04}	1.0189×10^{-04}			
Variance		2.3950×10^{-08}	1.0383×10^{-08}			

TABLE 2: List of parametrized mortality models, acronyms, and equations.

Abbreviations of Models	Models	References	Equations
HP	Heligman-Pollard	[11]	$m_x = A^{(x+B)^C} + D^* \exp[-E(\ln(x/F))^2] + (GH^x/1 + GH^x)$
KT	Kostaki	[12]	$m_x/1 - m_x = \begin{cases} A^{(x+B)^C} + D^* \exp[-E_1(\ln(x/F))^2] + GH^x, & \text{for } x \leq F \\ A^{(x+B)^C} + D^* \exp[-E_2(\ln(x/F))^2] + GH^x, & \text{for } x > F \end{cases}$
WT	Wittstein	[13]	$m_x = 1/BA^{-(Bx)^N} + A^{-(M-x)^N}$
RP	Rogers Plank	[14]	$m_x = A_0 + A_1^* \exp[-Ax] + A_2^* \exp[B(x-U) - \exp[-C(x-U)]] + A_3^* \exp[Dx]$

in the smoothing process to capture the mortality variations over time easily.

Furthermore, the WT model, which was introduced by Wittstein [13] and consisted of four parameters with two terms, was applied to investigate the human's mortality pattern [22]. The WT model is an alternative to the existing used logistic function in fitting observed probabilities at the oldest ages [23]. The WT model was postulated to overcome the Gompertz-law

anomaly on enhancing a model that works for many countries with minor flaws such as lack of model fit for a particular age group [24]. In addition, Rogers and Planck [14] developed the RP model, which is a multiexponential that compromises of four terms and nine parameters for modelling migration. The RP model consists of a constant, exponentially dropping child mortality, a double exponential accident hump, and Gompertzian senescent mortality [8].

3. Resampling Methods

This section briefly describes the resampling methods to select the parametric mortality model that best fits the Malaysian case for both men and women in normal circumstances. First, each model is fitted to acquire the parameters for their particular models for the entire study period by using the observed mortality rate expressed in Equation 1. Then, for each model and gender, a different loss function is applied to optimize the model (refer to Table 3). Refer to [10] for further information about the loss function. After acquiring all parameters for each model, hold-out, repeated hold-out, and leave-one-out cross-validation (LOOCV) methods are used to resample each parametric mortality model. Table 4 illustrates the goodness-of-fit measures in fitting the observed mortality data for each model, which are the mean absolute percentage error (MAPE), mean absolute error (MAE), sum square error (SSE), and mean square error (MSE).

Table 4 reveals that the HP model has the lowest values for all goodness-of-fit measures for the men's mortality rate, followed by the WT, RP, and KT models. The lower the goodness-of-fit measures, the better the parametric mortality model. The bold values indicate the best mortality models according to their respective goodness-of-fit measures. All of the goodness-of-fit measures display a consistent result for each model's performance when fitting the observed mortality rate. Based on Table 4, the HP model is the best-parametrized mortality model for fitting the observed mortality rate since the model has scored the best values for all measurements. Although the HP fits the men's mortality rate well, it cannot fit the women's mortality rate due to an overparameterization issue [25,26]. This issue is also applicable to our research study here. Therefore, Table 5 solely illustrates the goodness-of-fit measures for the KT, WT, and RP models in fitting the women's mortality rate.

Table 5 shows that the RP model has scored the best value for all goodness-of-fit measurements. Note that the lower the goodness-of-fit measure, the better the parametric mortality model. Therefore, the RP model is the best fit for the women's mortality rate. The RP model is the best based on MAPE values, followed by the WT and KT models. On the other hand, the MAPE value yields inconsistent results when compared to the other three measurements. In contrast, the RP model appears to be the best mortality model followed by the KT and WT models based on the MAE, SSE, and MSE values. It is important to note that the first procedure fits the observed mortality rate to acquire the parameters for each model and gender. While other studies applied MAE and MAPE to determine the level of prediction in forecasting crude oil prices [27,28], this study research study utilizes the same accuracy measures to determine the best parametric model in forecasting the mortality rate. Then, the next step is to apply hold-out, repeated hold-out, and LOOCV resampling methods to forecast the parameters for each model and gender.

3.1. Hold-Out Methods. Hold-out methods, also known as out-of-sample methods, require two sets of data: training

TABLE 3: Loss function in optimizing the parametric mortality model to obtain the fitted value for both genders.

Models	Men	Women
HP	Poisson	None
KT	LF6	LF6
WT	LF5	LF2
RP	LF3	LF2

and testing. The training set is a set of data designed to fit the model, whereas the testing set is used to evaluate the model's forecasting performance [29]. The training and testing sets are randomly divided depending on the number of samples. For instance, the training set contains 75% of the sample, while the testing set contains the remaining of 25%. However, this varies according to the situation, such as 80% and 20% or 2/3 and 1/3 [30,31]. The sample should be ordered chronologically because this study involves time-series data [32]. Refer Figure 1(a) in for an example of how to divide the data. For each mortality model, the methods for applying the hold-out method were as follows:

- (1) The data are separated into two sets, with the training set having a 2:1 ratio to the testing set. In addition, the training set spans the years 1995 through 2010, whereas the testing set spans the years 2011 through 2018.
- (2) The testing set parameters are forecasted using the parameters from the training set.
- (3) For all age groups, the parameters are fitted to acquire the forecasted mortality rate.
- (4) The goodness-of-fit measurements are used to assess each mortality model's forecasting performance for each age group.

3.2. Repeated Hold-Out Methods. The second resampling method is the repeated hold-out, which follows a similar procedure to the hold-out but involves several iterations [33]. An infinite new sample with a certain Bayesian distribution can be generated using this method [34]. Unlike the hold-out method, this method selects the training and testing sets at random for each iteration. The repeated hold-out, as proposed by Bergmeir et al. [35], applies the standard procedure for time series data without any adjustments. Refer Figure 1(b) in for an example of how the repeated hold-out method is applied in the analysis. The repeated hold-out method follows a similar procedure but takes into accounts the sample's peculiarities as follows:

- (1) The sample is divided into two sets, with two-third of the sample serving as the training set and one-third serving as the testing set. The procedure is repeated for a number of iterations. For the first iteration, the training set corresponds to the years 1996, 1997, 1999, 2001, 2002, 2003, 2007, 2008, 2009, 2012, 2013, 2014, 2015, 2016, 2017, and 2018. Whereas the testing set corresponds to the years 1995, 1998, 2000, 2004, 2005, 2006, 2010, and 2011 for the first iteration.

TABLE 4: Goodness-of-fit measures in fitting the men’s mortality rate.

Models	MAPE (%)	MAE	SSE	MSE
HP	4.5572	7.0954×10^{-04}	4.4854×10^{-05}	2.6385×10^{-06}
KT	33.5167	1.3949×10^{-03}	1.7418×10^{-04}	1.0246×10^{-05}
WT	13.4823	1.0489×10^{-03}	5.9971×10^{-05}	3.5277×10^{-06}
RP	22.7458	1.0584×10^{-03}	6.8785×10^{-05}	4.0462×10^{-06}

Note: bold values indicate the lowest values from their respective columns.

TABLE 5: Goodness-of-fit measures in fitting of women.

Models	MAPE (%)	MAE	SSE	MSE
KT	71.1260	2.0610×10^{-03}	2.2707×10^{-04}	1.3357×10^{-05}
WT	13.6325	2.3856×10^{-03}	6.0150×10^{-04}	3.5382×10^{-05}
RP	5.2998	1.0288×10^{-03}	1.5532×10^{-04}	9.1365×10^{-06}

Note: bold values indicate the lowest values from their respective columns.

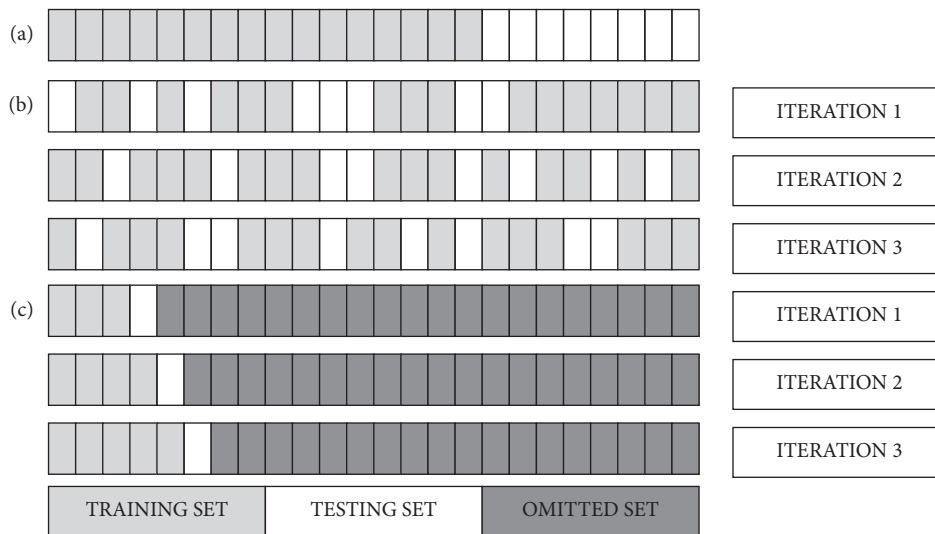


FIGURE 1: Schematic diagrams for (a) hold-out, (b) repeated hold-out, and (c) LOOCV resampling method.

- (2) Since the data are in a random order, the parameters for the testing set are obtained using the package “imputeTS” [36].
- (3) The parameters of the testing set are fitted to acquire the forecasted mortality rate for all age groups in the testing set.
- (4) The goodness-of-fit measures for each age group are computed to evaluate the forecasting ability for each mortality model.

Using the looping function in the R software, different training and testing sets are generated randomly for each iteration and model. The number of iterations is repeated 100 times as applied by Atance et al. [9].

3.3. *Leave-One-Out Cross-Validation.* The third method is leave-one-out cross-validation (LOOCV), nearly identical to the first hold-out method but differs in the proportion of the training and testing sets. For a more detailed discussion of the analysis involved in this analysis, see [37]. This method has

several advantages over other methods: (a) it minimizes sample bias because the training set is made up of $n-1$ observations, which covers almost the entire sample and (b) it selects the training and testing sets without involving randomness because almost all data are used for fitting and testing purposes [38]. The training set has a window with a defined period. For the following iteration, a new datum of time series is added chronologically, also known as an “assessment on a rolling forecasting origin one-step-ahead” [39]. For a visual representation of how this method works, see Figure 1(c) in . The following are the approaches for modifying this method:

- (1) The first three years (1995, 1996, and 1997) are utilized as a training set in the first iteration because three is the least number of samples to fit the mortality model [40].
- (2) The ARIMA model is used to forecast the parameter for the year 1998 based on the first three time-series data.
- (3) The number of training data is raised by one to acquire the next forecasted parameter for the

following iteration. Finally, the forecasted parameters are fitted to provide the forecasted mortality rate for each testing set's age group.

- (4) Goodness-of-fit measures are calculated for each age group to evaluate the forecasting ability of each mortality model.

4. Goodness-of-Fit Measures

This section presents the goodness-of-fit measurements for each model to assess its forecasting ability: (i) MAPE, (ii) MAE, (iii) SSE, and (iv) MSE.

Table 6 summarizes previous studies on goodness-of-fit measures for selecting the best model for fitting and forecasting the mortality rate in normal conditions. Equation 2 displays the MAPE formula, Equation 3 represents the MAE formula, Equation 4 illustrates the SSE formula, and Equation 5 shows the MSE for the hold-out method specifically.

$$\text{MAPE} = \frac{1}{n_d} \sum_t \sum_x \left| \frac{m_{x,t} - \hat{m}_{x,t}}{m_{x,t}} \right|, \quad (2)$$

$$\text{MAE} = \frac{1}{n_d} \sum_t \sum_x |m_{x,t} - \hat{m}_{x,t}|, \quad (3)$$

$$\text{SSE} = \sum_t \sum_x (m_{x,t} - \hat{m}_{x,t})^2, \quad (4)$$

$$\text{MSE} = \frac{1}{n_d} \sum_t \sum_x (m_{x,t} - \hat{m}_{x,t})^2. \quad (5)$$

Meanwhile, Equation (6) indicates the goodness-of-fit measure for repeated hold-outs, while Equation (7) exhibits the goodness-of-fit measure for the LOOCVs.

$$\text{Repeated hold - out} = \frac{1}{k} \sum_{j=1}^k \text{Goodness - of - fit measure}_j, \quad (6)$$

$$\text{LOOCV} = \frac{1}{n} \sum_{i=1}^n \text{Goodness - of - fit measure}_i. \quad (7)$$

The number of iterations involved in a repeated hold-out is represented by k . Since the iteration is repeated 100 times, k is equivalent to 100. On the other hand, n represents the number of observations in this study, which is 24. The forecasting performance of each mortality model is assessed using goodness-of-fit measures. Table 7 shows the goodness-of-fit measures for the men, while Table 8 shows the goodness-of-fit measures for the women.

Table 7 indicates the goodness-of-fit measure for each parametrized mortality model of resampling methods for the men's mortality data. In terms of resampling methodologies, the repeated hold-out method favours the KT model for men's mortality, whereas the hold-out method favours the HP, WT, and RP models. Note that the HP model is the best

TABLE 6: Summary of goodness-of-fit measures.

Goodness-of-fit	Articles
MAPE	[41–46]
MAE	[17, 47–49]
SSE	[17, 50–52]
MSE	[17, 41, 48, 53]

mortality model for fitting the men's mortality data, and the hold-out method is the best resampling method for men's data. Although the RP model has scored the lowest values for all goodness-of-fit measurements of the hold-out resampling method, the HP model has scored the lowest values for three goodness-of-fit measurements. The fact that the HP model has the lowest values for MAE, SSE, and MSE for the hold-out resampling method resulting the HP model as the optimal model for the men's mortality data.

Table 8 illustrates the goodness-of-fit measures in fitting the women's mortality rate. The hold-out resampling method is the best in fitting the women data based on the KT and RP models. However, when it comes to fitting the WT model, the repeated hold-out resampling method is the best. Note that the RP model fits the women's mortality data well, and the hold-out is the best resampling method for women data. Overall, the HP model for the men's mortality rate and the RP model for women's mortality data are the best models for fitting the mortality rate. Despite the fact that the repeated hold-out is the best resampling method for KT men and WT women, the hold-out has been demonstrated to be the best resampling method in other situations.

5. Forecasting Mortality under Normal Conditions and COVID-19

The HP model is applied for the men's data, and the RP model is utilized for the women's data using the hold-out resampling method to forecast the mortality rate under normal conditions and account for the COVID-19 mortality data. The mortality rate under normal conditions is forecasted from 2019 to 2030 without considering the COVID-19 mortality data. Then, the COVID-19 mortality data for 2019 and 2020 are included to predict the mortality rate until 2030. The excess mortality rate is determined by comparing the mortality forecast under normal conditions to the mortality forecast using the COVID-19 data. Figure 2 depicts the mortality rate for men aged 60 years and over. The straight line represents the observed mortality rate from 1995 until 2020, while blue dotted line indicates the fitted mortality rate under normal conditions from 1995 until 2030. The red dotted line indicates the forecasted mortality rate by considering the COVID-19 mortality data. When the mortality rate for people aged 60 and over is forecasted, the mortality rate shows a decreasing pattern after 2020. However, when the COVID-19 data for 2019 and 2020 are incorporated in the forecasting, the mortality rate begins to rise in 2023 and continues to rise after that. When using the COVID-19 data in forecasting, it demonstrates that there is

TABLE 7: Results of the goodness-of-fit measures for the men's mortality.

Models	Measures	Hold-out	Repeated hold-out	LOOCV
HP	MAPE (%)	5.6770	5.1238	7.8842
	MAE	7.8611 $\times 10^{-04}$	8.1409×10^{-04}	1.2178×10^{-03}
	SSE	5.1695 $\times 10^{-05}$	7.8236×10^{-05}	1.697110^{-04}
	MSE	3.0409 $\times 10^{-06}$	4.6021×10^{-06}	9.9829×10^{-06}
KT	MAPE (%)	34.9447	33.1597	35.7525
	MAE	1.4566 $\times 10^{-03}$	1.4572×10^{-03}	2.1629×10^{-03}
	SSE	1.8423×10^{-04}	1.8385 $\times 10^{-04}$	4.3728×10^{-04}
	MSE	1.0837×10^{-05}	1.0815 $\times 10^{-05}$	2.5723×10^{-05}
WT	MAPE (%)	14.7517	18.2086	27.8585
	MAE	1.4752×10^{-01}	1.7130 $\times 10^{-03}$	3.3979×10^{-03}
	SSE	6.1264 $\times 10^{-05}$	4.0727×10^{-04}	1.1290×10^{-03}
	MSE	3.6038 $\times 10^{-06}$	2.3957×10^{-05}	6.6410×10^{-05}
RP	MAPE (%)	19.1942	22.7014	39.1911
	MAE	1.0162 $\times 10^{-03}$	1.1772×10^{-03}	2.6291×10^{-03}
	SSE	6.6918 $\times 10^{-05}$	1.0276×10^{-04}	5.2375×10^{-04}
	MSE	3.9363 $\times 10^{-06}$	6.0449×10^{-06}	3.0809×10^{-05}

Note: bold values indicate the lowest measurements from their respective rows.

TABLE 8: Results of the goodness-of-fit measures for the women's mortality.

Models	Measures	Hold-out	Repeated hold-out	LOOCV
KT	MAPE (%)	74.1485	71.3582	66.8255
	MAE	2.1718 $\times 10^{-03}$	2.2875×10^{-03}	2.5198×10^{-03}
	SSE	2.2437 $\times 10^{-04}$	2.8220×10^{-04}	2.6727×10^{-04}
	MSE	1.3198 $\times 10^{-05}$	1.6635×10^{-05}	1.5722×10^{-05}
WT	MAPE (%)	14.342	13.8261	16.6043
	MAE	2.5703×10^{-03}	2.4004 $\times 10^{-03}$	2.5159×10^{-03}
	SSE	7.2283×10^{-04}	6.1679 $\times 10^{-04}$	7.5042×10^{-04}
	MSE	4.2519×10^{-05}	3.6282 $\times 10^{-05}$	4.4142×10^{-05}
RP	MAPE (%)	5.6581	5.9980	10.3610
	MAE	1.0511 $\times 10^{-03}$	1.0693×10^{-03}	1.4759×10^{-03}
	SSE	1.5562 $\times 10^{-04}$	1.6674×10^{-04}	2.8049×10^{-04}
	MSE	9.1539 $\times 10^{-06}$	9.8081×10^{-06}	1.6500×10^{-05}

Note: bold values indicate the lowest measurements from their respective rows.

an aftereffect of mortality rate for men population aged 60 and over.

Figure 3 displays the mortality rate of women for the age group of 60 years and over. Based on the observed mortality rate, all age groups showed a decreasing pattern. However, when the mortality rate is forecasted, considering the COVID-19 mortality data, the mortality rate indicates a slight increase in 2023 and then a reduction after that. The women's mortality rate increases with age, but not as much as the men's mortality rate. Compared to the women's mortality rates, the men's mortality rates have a delayed effect but last longer. This is most likely related to the fact that women have a higher chance of outliving men.

6. Discussion

As the mortality rates for the elderly population of 60 years and over improve, the ageing population becomes a major concern [54, 55]. The mortality rate of an ageing population has been steadily increasing from 1950 to 2015 in Malaysia, consistent from 1950 to 2015 [56]. Furthermore, Malaysia's number of

people aged 60 and over is predicted to increase, resulting in an ageing population. However, when the mortality rate is forecasted using the COVID-19 mortality data for 2019 and 2020, the men's mortality rate displays a delayed effect. Still, the trend begins to indicate an increasing pattern starting in 2023. The mortality rate has increased significantly after 2023 across all age groups for men aged 60 and over.

Furthermore, when the COVID-19 mortality data are considered, the women's mortality rate also shows an impact, although it does not last as long as the men's mortality rate. Due to the COVID-19, it seems that the population of men has a greater impact than the population of women. The women's mortality rate has only impacted a short-term until 2023, after which it will return to a stable trend as forecasted under normal conditions. The COVID-19 mortality rate exhibits a bigger impact when the age group increases. The result is consistent with the studies in [7, 57], which demonstrated an excess mortality rate, particularly in countries with a large population of people aged 65 years and above during the COVID-19 pandemic. The excess mortality increased with age above 70 years and correlated in time

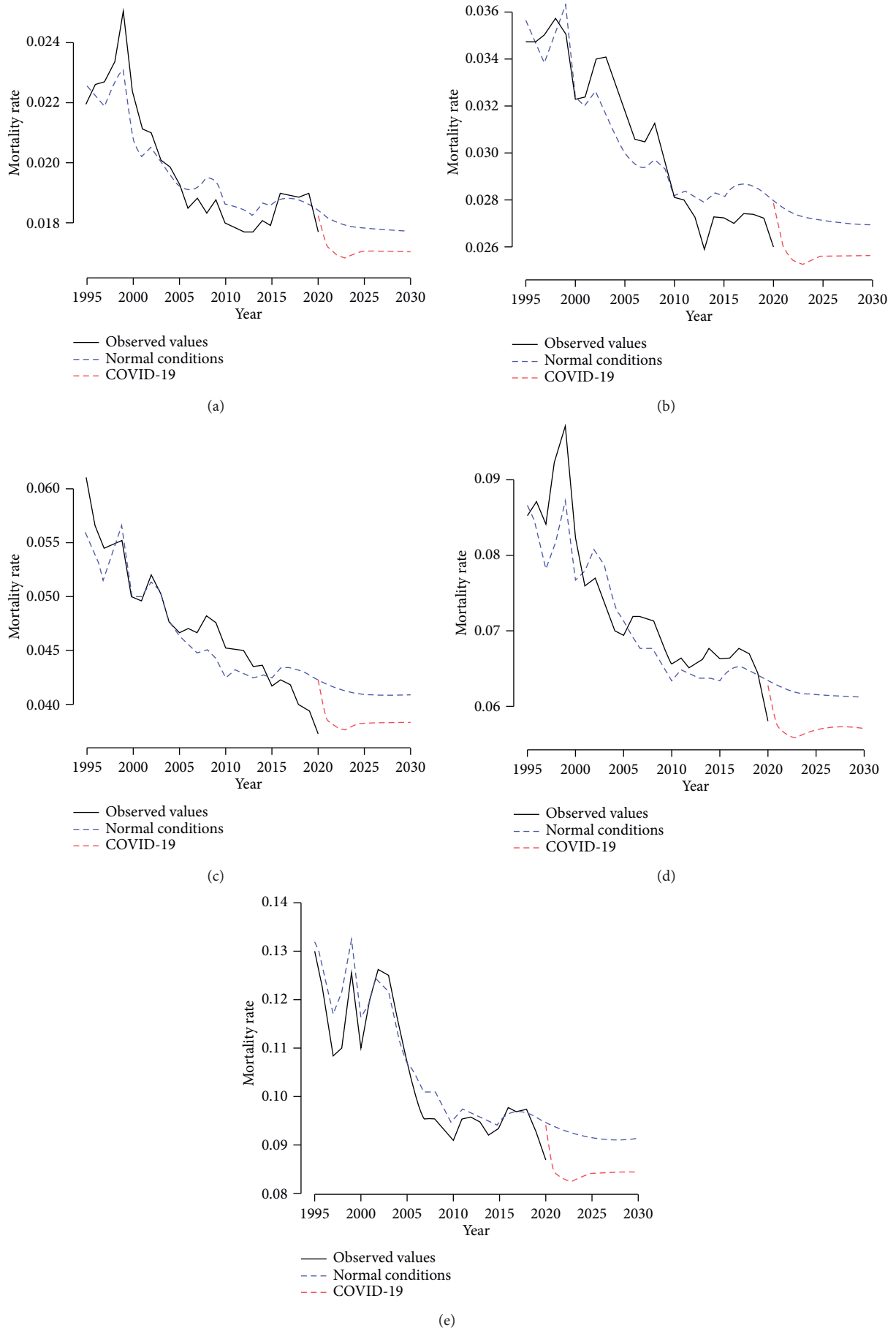


FIGURE 2: Men's mortality rate for the age groups of (a) 60-64, (b) 65-69, (c) 70-74, (d) 75-79, and (e) 80-84.

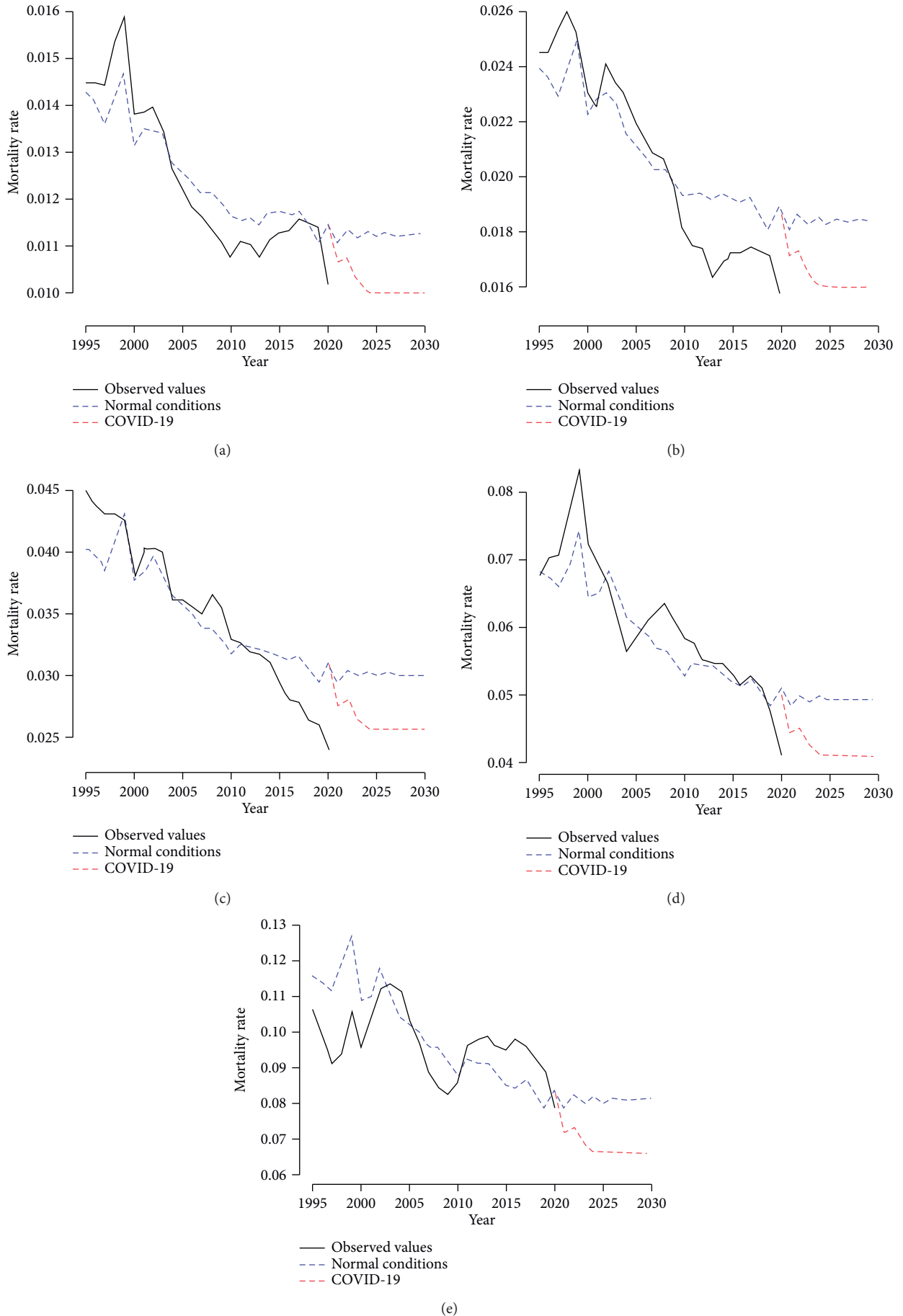


FIGURE 3: Women’s mortality rate for the age groups of (a) 60–64, (b) 65–69, (c) 70–74, (d) 75–79, and (e) 80–84.

with the COVID-19 reported mortality rate time series in Italy [58].

Based on the forecasted mortality rates for men and women, the excess mortality rate only takes place after 2023. This is due to the fact that the mortality rate shows a decreasing pattern and increases after 2023. It can be concluded that the Malaysia's population experience the ageing population to the point that the mortality rate due to confirmed COVID-19 mortality has not impacted immediately as soon as the COVID-19 outbreak in 2020. This is most likely due to the fact that an increasing of the ageing population resulting from the tremendous achievements of public health policies and social and economic development in Malaysia [59].

Moreover, the mortality rates were higher for ages above 50 due to the pandemic in Spain but no evidence in the Czech Republic [7]. The empirical result suggests that the COVID-19 mortality has a delayed and longer effect on the men's mortality rate. On the other hand, COVID-19 has a quick and short effect on the mortality rate of women. Although the forecasting only considers the mortality data for two years due to data availability, the results could be used as a benchmark to predict the post-COVID-19 mortality rate for future planning.

7. Conclusion

This study determines the excess mortality rate for particular age groups during the COVID-19 in Malaysia. In order to determine the excess mortality rate, this study employed various parametrized mortality models to forecast the mortality rate under normal conditions, including the Heligman-Pollard, Kostaki, Wittstein, and Rogersplanck models. The data set involved in this study ranges from the year 1995 to 2018. Furthermore, this study identifies the optimal mortality law for fitting the mortality data utilising multiple resampling methods, such as hold-out, repeated hold-out, and leave-one-out cross-validation. The optimal model for men and women's mortality rates is determined using a variety of goodness-of-fit measures (mean absolute percentage error, mean absolute error, sum square error, and mean square error). The Heligman-Pollard model for men and the Rogers Planck model for women are the optimal models based on goodness-of-fit measures. Both models favour the hold-out technique as the best resampling method. While our study has been conducted in a yearly basis, there is no outlier detected by observing the boxplot. The post-COVID-19 mortality rate is projected for ten years up to 2030 using the COVID-19 mortality data for 2019 and 2020. The empirical results revealed that the COVID-19 mortality in Malaysia has an excess mortality rate, particularly among those aged 60 and over. The men's mortality rate appears to have a delayed and longer effect than the women's mortality rate based on the forecasted mortality rate. This is most likely due to the risk of a woman living longer than a man. In conclusion, this study recommends amending the existing policy to reflect the post-COVID-19 mortality forecast. Furthermore, the mortality rate reveals that excess mortality rate is caused not just by the illness

itself but also by other psychological matters such as suicide and health treatment delay.

Data Availability

Data are available on request. The data are collected from publicly available sources except for the ones that have been mentioned otherwise. The raw data supporting the conclusions of this manuscript will be made available by the authors without undue reservation.

Conflicts of Interest

There are no conflicts of interest related to this work.

Authors' Contributions

Robiaatul Adawiah Edrus contributed to literature review, data collection, data analysis, research framework, methodology, validation, and initial draft. Zailan Siri performed conceptualization, validation, methodology, formal analysis, and supervision. Mohd Azmi Haron performed conceptualization, validation, methodology, formal analysis, review, editing, and supervision and provided resources. Muhammad Aslam Mohd Safari contributed to conceptualization, validation, methodology, formal analysis, initial draft, and supervision and provided the software. Mohammed K. A. Kaabar contributed to validation, methodology, supervision, initial draft, and editing and provided resources. All authors read and approved the final version.

References

- [1] S. M. Shaharudin, S. Ismail, N. A. Hassan, M. L. Tan, and N. A. F. Sulaiman, "Short-term forecasting of daily confirmed COVID-19 cases in Malaysia using RF-SSA model," *Frontiers in Public Health*, vol. 9, p. 625, 2021.
- [2] Ministry of Health, "Malaysia," 2021, <https://covidnow.moh.gov.my>.
- [3] X. Q. Wang, G. Song, Z. Yang et al., "Association between ageing population, median age, life expectancy and mortality in coronavirus disease (COVID-19)," *Ageing*, vol. 12, no. 24, pp. 24570–24578, 2020.
- [4] M. A. Torres Acosta and B. D. Singer, "Pathogenesis of COVID-19-induced ARDS: implications for an ageing population," *European Respiratory Journal*, vol. 56, no. 3, Article ID 2002049, 2020.
- [5] S. Z. Jamaluddin, F. Yuen Wah, and M. Abu Taher, "Covid-19: a preliminary assessment on the social security framework for an aged Malaysia," *Commonwealth Law Bulletin*, vol. 47, no. 1, pp. 55–71, 2021.
- [6] C. F. Chung, S. Tan, J. B. Singau, K. H. Pazim, and K. Mansur, "The repercussions of COVID-19 Pandemic on the wellbeing of older people in Malaysia: a literature review," *International Journal for Studies on Children, Women, The Elderly Disabilities*, vol. 11, pp. 17–23, 2020.
- [7] O. Šimpach and M. Pechrová, "Implications of the SARS-cov-2 pandemic for mortality forecasting: case study for the Czech republic and Spain," *Engineering Proceedings*, vol. 5, p. 58, 2021.
- [8] H. Booth and L. Tickle, "Mortality modelling and forecasting: a review of methods," *Annals of Actuarial Science*, vol. 3, no. 1, pp. 3–43, 2008.

- [9] D. Atance, A. Debón, and E. Navarro, “A comparison of forecasting mortality models using resampling methods,” *Mathematics*, vol. 8, no. 9, pp. 1550–1570, 2020.
- [10] M. D. Pascariu, “Package “Mortalitylaws”” . *Mortalitylaws: Parametric Mortality Models, Life Tables And HMD*, R Package Version, 2020.
- [11] L. M. A. Heligman and J. H. Pollard, “The age pattern of mortality,” *Journal of the Institute of Actuaries*, vol. 107, no. 1, pp. 49–80, 1980.
- [12] A. Kostaki, “A nine-parameter version of the Heligman-Pollard formula,” *Mathematical Population Studies*, vol. 3, no. 4, pp. 277–288, 1992.
- [13] T. Wittstein, “The mathematical law of mortality,” *Journal of the Institute of Actuaries and Assurance Magazine*, vol. 24, no. 3, pp. 153–173, 1883.
- [14] A. Rogers and F. Planck, *MODEL: A General Program for Estimating Parameterized Model Schedules of Fertility, Mortality, Migration, and Marital and Labor Force Status Transitions*, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1983.
- [15] Y. Haruna Umar and U. Joshua Chukwudi, “Modeling mortality rates using Heligman-Pollard and Lee-Carter in Nigeria,” *American Journal of Theoretical and Applied Statistics*, vol. 8, no. 6, pp. 221–239, 2019.
- [16] R. D. Lee and L. R. Carter, “Modeling and forecasting U.S. mortality,” *Journal of the American Statistical Association*, vol. 87, no. 419, pp. 659–671, 1992.
- [17] E. Silva, A. Ramos, and K. G. Olvera, “Inferences on mortality using the Heligman-Pollard model: the Mexican case,” *Communications in Statistics—Simulation and Computation*, vol. 49, no. 12, pp. 3227–3239, 2020.
- [18] K. N. Zafeiris and A. Kostaki, “Recent mortality trends in Greece,” *Communications in Statistics—Theory and Methods*, vol. 48, no. 1, pp. 112–126, 2017.
- [19] L. Zanotto, V. Canudas-Romo, and S. Mazzuco, “A mixture-function mortality model: illustration of the evolution of premature mortality,” *European Journal of Population*, vol. 37, no. 1, pp. 1–27, 2020.
- [20] C. H. Skiadas and K. N. Zafeiris, “Methods of modeling and estimation in mortality,” *Data Analysis and Applications 1: Clustering and Regression, Modeling-estimating, Forecasting and Data Mining*, Wiley, Hoboken, NJ, USA, 2019.
- [21] A. Kostaki and K. N. Zafeiris, “Dealing with limitations of empirical mortality data in small populations,” *Communications in Statistics: Case Studies, Data Analysis and Applications*, vol. 5, no. 1, pp. 39–45, 2019.
- [22] O. G. Michael, “On empirical analysis of gompertzian mortality dynamics,” *Journal of Science*, vol. 2, pp. 1–18, 2021.
- [23] P. Pflaumer, “Projecting age-specific death probabilities at advanced ages using the mortality laws of Gompertz and Wittstein,” *JSM Proceedings, Social Statistics Section*, vol. 103, 2018.
- [24] F. A. Lawal and J. B. Ndikubwimana, “Modelling Nigeria male mortality using functional data time series analysis approach,” *SSRN Electronic Journal*, pp. 1–14, 2021.
- [25] A. Alexopoulos, P. Dellaportas, and J. J. Forster, “Bayesian forecasting of mortality rates by using latent Gaussian models,” *Journal of the Royal Statistical Society: Series A*, vol. 182, no. 2, pp. 689–711, 2019.
- [26] S. Mazzuco, B. Scarpa, and L. Zanotto, “A mortality model based on a mixture distribution function,” *Population Studies*, vol. 72, no. 2, pp. 191–200, 2018.
- [27] W. Gao, M. Aamir, A. B. Shabri, R. Dewan, and A. Aslam, “Forecasting crude oil price using kalman filter based on the reconstruction of modes of decomposition ensemble model,” *IEEE Access*, vol. 7, pp. 149908–149925, 2019.
- [28] P. Xu, M. Aamir, A. Shabri, M. Ishaq, A. Aslam, and L. Li, “A new approach for reconstruction of imfs of decomposition and ensemble model for forecasting crude oil prices,” *Mathematical Problems in Engineering*, vol. 2020, Article ID 1325071, 23 pages, 2020.
- [29] L. J. Tashman, “Out-of-sample tests of forecasting accuracy: an analysis and review,” *International Journal of Forecasting*, vol. 16, no. 4, pp. 437–450, 2000.
- [30] A. Debón, F. Montes, J. Mateu, E. Porcu, and M. Bevilacqua, “Modelling residuals dependence in dynamic life tables: a geostatistical approach,” *Computational Statistics & Data Analysis*, vol. 52, no. 6, pp. 3128–3147, 2008.
- [31] L. Torgo, *Data Mining with R: Learning with Case Studies*, Chapman and Hall/CRC, Boca Raton, FL, USA, 2011.
- [32] C. Bergmeir and J. M. Benítez, “On the use of cross-validation for time series predictor evaluation,” *Information Sciences*, vol. 191, pp. 192–213, 2012.
- [33] A. Forsythe and J. A. Hartigan, “Efficiency of confidence intervals generated by repeated subsample calculations,” *Biometrika*, vol. 57, no. 3, pp. 629–639, 1970.
- [34] D. Barry and J. Hartigan, “A conversation with john hartigan,” *Statistical Science*, vol. 20, no. 4, pp. 418–430, 2005.
- [35] C. Bergmeir, R. J. Hyndman, and B. Koo, “A note on the validity of cross-validation for evaluating autoregressive time series prediction,” *Computational Statistics & Data Analysis*, vol. 120, pp. 70–83, 2018.
- [36] S. Moritz and T. Bartz-Beielstein, “ImputeTS: time series missing value imputation,” *R. R Journal*, vol. 9, no. 1, p. 207, 2017.
- [37] P. Burman, “A comparative study of ordinary cross-validation, v-fold cross-validation and the repeated learning-testing methods,” *Biometrika*, vol. 76, no. 3, pp. 503–514, 1989.
- [38] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning: With Applications in R*, Springer Science & Business Media, Berlin, Germany, 2013.
- [39] R. J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*, OTexts, Melbourne, Victoria, Australia, 2021.
- [40] R. J. Hyndman and Y. Khandakar, “Advances in automatic time series forecasting,” *Journal of Statistical Software*, vol. 27, 2008.
- [41] A. Raziq, N. Sheikh, M. Rafique, G. Murtaza, and A. Muhammad, “Forecasting fertility and child mortality rates of Pakistan using time series model,” *Linguistica Antverpiensia*, vol. 2021, no. 2, pp. 1194–1213, 2021.
- [42] C. C. L. Tsai and E. S. Cheng, “Incorporating statistical clustering methods into mortality models to improve forecasting performances,” *Insurance: Mathematics and Economics*, vol. 99, pp. 42–62, 2021.
- [43] C. C. L. Tsai and A. D. Wu, “Bühlmann credibility-based approaches to modeling mortality rates for multiple populations,” *North American Actuarial Journal*, vol. 24, no. 2, pp. 290–315, 2020.
- [44] C. C. L. Tsai and T. Lin, “A bühlmann credibility approach to modeling mortality rates,” *North American Actuarial Journal*, vol. 21, no. 2, pp. 204–227, 2017.
- [45] W. Van Wel, “Mortality modelling and forecasting using cross-validation techniques,” *MaRBL*, vol. 1, 2015.
- [46] C. Wan and L. Bertschi, “Swiss coherent mortality model as a basis for developing longevity de-risking solutions for Swiss pension funds: a practical approach,” *Insurance: Mathematics and Economics*, vol. 63, pp. 66–75, 2015.

- [47] S. Satpathy, M. Mangla, N. Sharma, H. Deshmukh, and S. Mohanty, "Predicting mortality rate and associated risks in COVID-19 patients," *Spatial Information Research*, vol. 29, no. 4, pp. 455–464, 2021.
- [48] S. R. Kessy, M. Sherris, A. M. Villegas, and J. Ziveyi, "Mortality forecasting using stacked regression ensembles," *Scandinavian Actuarial Journal*, pp. 1–36, 2021.
- [49] H. Nazari Kangavari, A. Shojaei, and S. S. Hashemi Nazari, "Suicide mortality trends in four provinces of Iran with the highest mortality, from 2006-2016," *Journal of Research in Health Sciences*, vol. 17, no. 2, Article ID e00382, 2017.
- [50] M. Haghbayan, S. S. Khatami, and F. N. Heravi, "The estimation of newly infected cases of Covid-19 with consideration of governmental action and behavior of people in Iran," *SRPH Journal of Medical Sciences and Healthcare Management*, vol. 3, pp. 1–7, 2021.
- [51] H. E. Abraha, A. B. Belachew, M. M. Ebrahim, M. H. Tequare, M. T. Adhana, and N. E. Assefa, "Magnitude, trend, and causes of under-five mortality from Kilite-Awlaelo health demographic surveillance database, northern Ethiopia, 2009–2017," *BMC Public Health*, vol. 20, pp. 1465–1468, 2020.
- [52] A. J. G. Cairns, D. Blake, K. Dowd et al., "A quantitative comparison of stochastic mortality models using data from England and Wales and the United States," *North American Actuarial Journal*, vol. 13, pp. 1–35, 2009.
- [53] N. Blanco, K. A. Stafford, M. C. Lavoie, A. Brandenburg, M. W. Górna, and M. Merski, "A simple model for the total number of SARS-CoV-2 infections on a national level," *Epidemiology and Infection*, vol. 149, pp. e80–10, 2021.
- [54] N. A. Siripatsopon, "Mortality model and its implications for the Thai elderly," *J. Assoc. Res.* vol. 25, pp. 441–457, 2020.
- [55] R. I. Ibrahim and N. Mohd Nordin, "Analysis of mortality improvement on the pension cost due to aging population," *Matematika*, vol. 36, no. 3, pp. 209–216, 2020.
- [56] N. S. Mohamad Ibrahim, S. Niza Shair, and A. Yuzi Yusof, "Mortality rates and life expectancy improvements among Malaysian elderlies," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 19, no. 1, pp. 134–139, 2020.
- [57] P. Vanella, U. Basellini, and B. Lange, "Assessing excess mortality in times of pandemics based on principal component analysis of weekly mortality data the case of COVID-19," *Genus*, vol. 77, no. 1, 2020.
- [58] C. Modi, V. Boehm, S. Ferraro, G. Stein, and U. Seljak, "total covid-19 mortality in italy: excess mortality and age dependence through time-series analysis," *Nature Communications*, vol. 12, 2020.
- [59] N. H. Ahmad Bahuri, H. Rizal, H. Abdul Majid, M. A. Said, and T. T. Su, "Development of the active ageing awareness questionnaire in Malaysia," *Healthcare*, vol. 9, no. No. 5, p. 499, 2021.