MATEMATIKA, MJIAM, Volume 43, Number 1, 309–320

Modelling Malaysian Mortality Improvement Using a Hybrid Logistic Spline

[1]Nur Idayu Ah Khaliludin*, [1]Norrlaili Shapiee, [2]Zarina Mohd Khalid, [2]Siti Rohani Mohd Nor, [3]Nurliyana Juhan, [4]Rose Irnawaty Ibrahim and [4]Yumn Suhaylah Yusoff

[1]Pusat Tamhidi, Universiti Sains Islam Malaysia,
71800 Nilai, Negeri Sembilan, Malaysia.

[2]Department of Mathematical Sciences, Universiti Teknologi Malaysia,
81310 UTM Johor Bahru, Johor, Malaysia.

[3]Preparatory Centre for Science and Technology,
Universiti Malaysia Sabah, Sabah, Malaysia.

[4]Fakulti Sains dan Teknologi, Universiti Sains Islam Malaysia,
71800 Nilai, Negeri Sembilan, Malaysia.

*Corresponding author: nuridayu@usim.edu.my

Article history
Received: 7 November 2024
Received in revised form: 11 August 2025
Accepted: 18 September 2025
Published on line: 15 November 2025

Abstract The unexpected rise in life expectancy and the declining mortality rates among older individuals have positioned Malaysia as an ageing nation. As this demographic shift becomes increasingly significant, the modeling of mortality rates is essential to address the specific needs of the elderly population effectively and promptly. However, the task of modeling these rates presents challenges, as mortality data frequently display irregular patterns attributed to various uncertainties. It cannot be formulated as a single mathematical equation as some age ranges have different shapes of mortality. The existing mortality models consisted of two separate mortality estimations between young and old age. This resulted in high correlation of parameters, rough mortality curve and incoherence estimation. To solve this problem, this research proposes a hybrid method that integrates a spline with three parametric logistic models. This research applies the proposed model to the Malaysian data from 2011 to 2021. The result is validated using Semi-Parametric Bootstrap. The bias and standard variation resulted from the bootstrap method declined as the number of replications increased, which indicates that the hybrid model able to estimate the underlying mortality risk. Thus, as this research provides a viable estimation of mortality improvement, this study will certainly bring a substantial benefit to the pension and health care provider and the insurance companies to calculate the mortality risk and build a sustainable reserve fund.

Keywords Mortality, Ageing, Bootstrap, Logistic, Spline.

Mathematics Subject Classification AMS 62P05.

---

1 Introduction

The implications of an aging population have heightened the focus on analyzing mortality rates. In response, various mortality models have been created to examine mortality trends globally. Yet, the inaccuracy of data proposes a major challenge to apply accurate mortality models in Malaysia on how to deal with the random variations due to low number of deaths and people surviving at old ages [1–3].

Pitacco provides a comprehensive review of the earliest mortality models such as the Gompertz model, Makeham model and Heligman-Pollard model [4]. Gompertz law of mortality states that the logarithm function of mortality rates is a linear function of age, and it is applicable to limited age range for instance, 30 to 90 years of age. The extension of this law is Makeham model and Heligman-Pollard model where more parameters and a constant are added to the Gompertz formula to capture the young age mortality. These models although improve the estimation of Gompertz model, they are difficult to fit in practice because of the high correlation in the estimated parameters. The high correlation also compromises the interpretability of the parameters [5].

However, studies by [5] and [6] of comparative studies on models for the oldest age. They proved that the constant rate of increment mentioned in the Gompertz laws above is not suitable for old age. [7] proposed the utilisation of logistic functions, akin to the Beard and Tatcher models, to illustrate mortality rates approaching an asymptotic limit. However, it is noteworthy that these models are exclusively relevant to adult populations. In contrast, during earlier stages of life, where mortality rates are comparatively lower, the behavior of the logistic curve aligns more closely with the Gompertz or Makeham laws.

Due to the difficulty in estimating the mortality models for young ages, non-parametric smoothing models which are based on the spline method are used. Examples of the spline method used in modelling the mortality model are variable knot cubic spline [1] and weighted least squares smoothing spline [2]. The differences between these splines are the location of the age knot that specifies the behaviour of the mortality curvature. A modification made in the weighted least square smoothing spline by [5] is that the user specifies a set of weights based on their judgment of the age knot. The weakness of this method is that this requires ad-hoc assumptions such as the mortality rate at age 100 is 1. The model in question posits an exponential increase in mortality rates, leading to the exclusion of certain data from the highest age groups to ensure a consistent upward trend in the smoothed mortality rates. In this context, developed a geometrically structured variable knot regression spline to create a mortality table for the Unit-ed Kingdom. This process unfolds in two stages: initially, a variable knots linear spline is fitted, followed by the estimation of the optimal control polygon for higher-order splines. Consequently, this approach allows for the determination of the most effective age knot sequences, along with the order and coefficients of the spline. [5] advocate for the application of spline models in analysing mortality rates, highlighting the adaptability of splines as a valuable method for estimating mortality at advanced ages.

This research seeks to develop a mortality model capable of accurately estimating Malaysian mortality rates across all age groups, with a particular focus on older ages where data may be sparse or unavailable. As noted in [8], mortality rates among the elderly in Malaysia are the highest and have shown an increasing trend despite an overall decline. Additionally, life expectancy in Malaysia has remained constant at 74.5 years since 2013 [9–11]. Therefore, we advocate for a logistic-based model, which employs a sigmoidal curve to depict slow growth at the beginning and end, with a phase of rapid growth in between.

Yet, the logistic models often fail to capture certain critical aspects of young age mortality, such as the spike in accidental deaths during late adolescence. Therefore, to estimate mortality rates across all age groups and ensure a consistent transition of mortality rates from one age to the next, this research propose a hybrid approach that combines non-parametric techniques for estimating crude mortality rates in younger age groups and parametric models for old age mortality rates. Non-parametric techniques, such as the spline technique, are used to create a discernible pattern by reducing or smoothing out random fluctuations or irregularities in the data that do not represent the underlying trend or signal. This method begins by producing smoothed mortality rates derived from the crude mortality rates for the available age range. For the oldest age groups and those ages lacking mortality data, parametric models are employed to estimate the mortality rates.

To validate the model consistency, bias and accuracy to estimate the mortality rates, this study employed semi-parametric bootstrap method. According to [12] and [13] semi-parameter bootstrapping allows researchers to leverage the strengths of both parametric assumptions and non-parametric flexibility. Since this research using hybrid approach between the non-parametric and parametric mortality model, this adaptability enables the modeling of complex data structures without overly strict assumptions. Furthermore, in a comparative study, [19] and [20] found that semi-parameter bootstrapping exhibits greater robustness to outliers compared to standard parametric methods. To conclude, although the complexities and computational demands of the semi-parameter bootstrapping pose challenges to the researchers, but its advantages of being flexible and robust made it a valuable tool for statistical analysis regarding the mortality rates [6, 14, 15].

2 Methodology

This study aims to model the Malaysian mortality rates especially older age mortality rate and estimate the single. In this research, the methodologies can be divided into four sections in which the first section present the data collection (Section 2.1). These data are smoothed using the methodology discussed in Section 2.2 and Section 2.3. Finally, this model is assessed using simulation study discussed in Section 2.4

2.1 Data Collection

Traditionally, crude mortality rates are determined by dividing the total number of deaths by the population at risk. However, detailed data on the number of deaths and the population at risk for each specific age group is not publicly available. Instead, publicly accessible data is provided in an abridged format, where the number of deaths and the population are grouped into five-year age intervals. For the purposes of this study, the Department of Statistics of Malaysia (DoSM) has provided detailed data, including single age and abridged mortality rates, population numbers, and death counts for a ten-year period from 2011 to 2021. The following data is disaggregated by gender and age group. However, it should be aware that no obtainable data for infants, as this research is focused on mortality rates beyond infancy, particularly in older age groups. Moreover, infant mortality has unique characteristics and thus warrants a separate, dedicated study.

2.2 Estimation of One-Year Age Mortality Rates Via Akima Spline

The death toll for single age particularly fluctuates on account of natural variability in the mortality process among the at-risk population. Thus, to maintain a consistent progression of mortality rates across different ages and to calculate the absent probabilities, it becomes necessary to employ the spline method. This technique is utilised on death probabilities to guarantee a seamless transition between mortality rates of younger and older age groups.

Due to its dual practicality, the Akima spline technique is chosen to fulfill the task. It has the potential to estimate mortality rates for each individual age (1, 2, 3, 4,. . .,85), and also refine the mortality graph by reducing outliers and excessive fluctuations [16]. This results in a more precise and smooth depiction of mortality rates across various age groups.

Therefore, as to achieve the calculation of missing mortality rates for a person aged x in year t $(x = 1, 2, ...98; t = 2011, ..., 2021)$, $m_{x,t}$ this method employs four sets of four consecutive data points including $m_{x,t}$, i.e, $(m_{x-3,t}, m_{x-2,t}, m_{x-1,t}, m_{x,t})$, $(m_{x-2,t}, m_{x-1,t}, m_{x,t}, m_{x+1,t})$, $(m_{x-1,t}, m_{x,t}, m_{x+1,t}, m_{x+2,t})$, $(m_{x,t}, m_{x+1,t}, m_{x+2,t}, m_{x+3,t})$. The formula is for the missing rate is shown as equation (1)

$$m_{x,t} = m'_{x1,t}, w_{x1,t} + m'_{x2,t} w_{x2,t} + m'_{x3,t} w_{x3,t} + m'_{x4,t} w_{x4,t}, \tag{1}$$

where $m'_{xs,t}$ is the first derivative of the mortality rates in the s^{th} consecutive data points $(s = 1, 2, 3, 4)$ and it can be expounded in the equation (2). Whereas, $w_{xs,t}$ is the weight respective to the s^{th} consecutive data points as depicted in equation (4).

$$\begin{aligned} m'_{x1,t} &= F_t\left(x, x-3, x-2, x-1\right), \\ m'_{x2,t} &= F_t\left(x, x-2, x-1, x+1\right), \\ m'_{x3,t} &= F_t\left(x, x-1, x+1, x+2\right), \\ m'_{x4,t} &= F_t\left(x, x+1, x+2, x+3\right). \end{aligned} \tag{2}$$

The generalised function $F_t\left(i, j, k, l\right)$ is defined as equation (3)

$$\begin{aligned} F_t(i,j,k,l) = {} & \frac{(m_{j,t}-m_{i,t})(x_{k,t}-x_{i,t})^2(x_{l,t}-x_{i,t})^2(x_{l,t}-x_{i,t})}{(x_{j,t}-x_{i,t})(x_{k,t}-x_{i,t})(x_{l,t}-x_{i,t})(x_{k,t}-x_{j,t})(x_{l,t}-x_{j,t})} \\ & + \frac{(m_{k,t}-m_{i,t})(x_{l,t}-x_{i,t})^2(x_{j,t}-x_{i,t})^2(x_{j,t}-x_{l,t})}{(x_{j,t}-x_{i,t})(x_{k,t}-x_{i,t})(x_{l,t}-x_{i,t})(x_{k,t}-x_{j,t})(x_{l,t}-x_{k,t})} \\ & + \frac{(m_{l,t}-m_{i,t})(x_{j,t}-x_{i,t})^2(x_{k,t}-x_{i,t})^2(x_{k,t}-x_{j,t})}{(x_{j,t}-x_{i,t})(x_{k,t}-x_{i,t})(x_{l,t}-x_{i,t})(x_{l,t}-x_{k,t})(x_{l,t}-x_{j,t})} \end{aligned} \tag{3}$$

The weight $w_{xs,t}$ is the product of the sum of square of deviations from a straight line of the least-square fit, $V_t(i, j, k, l)$ and the distance factor, $D_t(i, j, k, l)$. Furthermore, i, j, k, and l are the data sets of four consecutive data points including $m_{x,t}$.

$$W_{xi,t} = \frac{D_t\left(x, x-i, x-i-1, x-i-2\right)}{V_t\left(x, x-i, x-i-1, x-i-2\right)}, \tag{4}$$

$$V_t\left(i,j,k,l\right) = \sum m^2 - b_0 \sum m + b_1 \sum mx. \tag{5}$$

In equations (5) to (7), symbol $\sum$ is the summation over four data points m_i, m_j, m_k and m_l. The distance factor $D_t\left(i,j,k,l\right)$ which is the sum of square of error can be represented as

$$D_t\left(i,j,k,l\right) = \left(x_{k,t} - x_{i,t}\right)^2 \left(x_{l,t} - x_{i,t}\right)^2 \left(x_{j,t} - x_{i,t}\right)^2. \tag{6}$$

It is important to note that the first index of the function in equations (2), (4) and (6) must be i, representing the missing mortality rate in question, while the remaining indices can be arranged in any order. This method is computationally intensive. Consequently, the implementation Akima method is employed by executing the algorithm from the Association for Computing Machinery Collected Algorithm (CALGO) in R Studio. Within R Studio, the ForeignBase package was employed by invoking the. Fortran code.

2.3 Estimation of Older Age

In Section 1, the logistic-based models discussed, particularly the trio of Beard, Kannisto, and Wilmoth models, represent iterative modifications of one another. Both the Kannisto and Beard models incorporate the term $\alpha e^{\beta x}$ into the numerator of the traditional logistic model as a limiting rate, while the Wilmoth model adopts a distinct methodology. Additionally, the Kannisto and Wilmoth models are structured to maintain an asymptote of 1, thereby preventing the output from exceeding this value. However, an examination of the mortality curve in Malaysia revealed that the asymptote of 1 in these models is overly restrictive for the Malaysian demographic context. The Malaysian mortality curve adheres to the logistic pattern, albeit with certain deviations in its S-shape. As a result, a modification of the Wilmoth model was implemented. In contrast to the Kannisto, Beard, and Wilmoth models, the revised model allows the mortality data to dictate its maximum or limiting rate, thus establishing the numerator as a specific value, α. In this proposed logistic framework, the parameters β and ς regulate the rate of mortality fluctuations, where β indicates the age at which the mortality curvature shifts direction, and ς characterizes the degree of 'wriggliness' or the sigmoidal form of the curve.

The following equation is the proposed logistic equation (7) to depict the older age mortality.

$$\hat{m}_{x,t} = \frac{\alpha}{1 + \exp\left(-\frac{x-\beta}{\varsigma}\right)} \tag{7}$$

The parameter α represents the asymptote, which is the stable or peak mortality rate that varies between 0 and 1. When constructing a life table, it is invariably assumed that each individual will reach a definitive maximum age, denoted as ω. It is crucial to recognize that the maximum rate of mortality, denoted as α, may not always coincide with the highest attainable age, represented by ω. In other words, the age at which the highest mortality rate occurs may not be the same as the maximum possible lifespan.

Moreover, the parameter ς defines the degree of 'wriggliness' or the sigmoidal form of the curve, with its values extending from 0 to infinity. Within the standard logistic function, ς represents the rate at which growth decreases as size increases. In relation to the proposed model, a ς value of 0 produces a flat, horizontal line, whereas a higher ς value results in a line that consistently rises.

The proposed model delineates the inflection point at which the curve changes its shape, the apex of the mortality rate, and the curvature of the mortality graph. This, in turn, increases the adaptability of our model.

However, to estimate the entire age spectrum, we continue to utilise the Akima model for ages 0 up to a certain older age, denoted as x_0, and supplement it with our proposed model. x_0 represents the threshold age at which the transition from the Akima model to the proposed model occurs. Although the age of 60 is officially recognized as the threshold for old age in Malaysia, it is more advantageous to allow mortality data to dictate the appropriate age classification. The integration of spline techniques with parametric models for old age is widely endorsed in various countries, including the United Kingdom [7–9]. Therefore, our comprehensive model is represented in equation (6).

In this context, $\hat{m}_{x,t}$ denotes the estimated mortality rate for an individual of age x, $s\left(x; m; t\right)$ refers to the Akima spline method, and $\dfrac{\alpha}{1+\exp\left(-\dfrac{x-\beta}{\varsigma}\right)}$ constitutes the proposed model.

2.4 Simulation Study via Semi-Parametric Bootstrap

This simulation is intended to assess various beneficial properties of the estimator $\hat{\theta}$ from a frequentist perspective: 1) $\hat{\theta}$ should be unbiased for θ in finite sample: $E\left(\hat{\theta}\right) = \theta$ and consistent such that as $n \to \infty$, $\hat{\theta} \to \theta$ where n is the number of observations. 2) The sample estimate of variance, denoted as $Var\left(\hat{\theta}\right)$ should provide consistent estimation of the sample variance of $\hat{\theta}$. 3) the constructed confidence intervals must possess the property that at least $100(1-\alpha)\%$ of intervals contain θ. 4) It is essential that the variance $Var\left(\hat{\theta}\right)$ be minimized to ensure $\hat{\theta}$ serves as an efficient estimator of θ.

The non-parametric bootstrap method is employed for the simulation analysis. The process of bootstrapping can be seen in these steps: 1) A sample of size n is selected with replacement from the sample population, denoted as S. The selected sample is referred Referring the selected sample as the bootstrap sample $S^* = \{m_{11}, ..., m_{1n}\}$. 2) Repeating Procedure 1 R times which resulting in R distinct bootstrap samples. The b^{th} bootstrap sample is designated as $Sb^* - \{m_{b1}, ..., m_{bm}\}$. 3) Marked as L, the proposed logistic model is applied to each bootstrap sample to derive the corresponding bootstrap statistics, L^*. 4) The bias B^* and variance $V\left(L^*\right)$ of the bootstrap statistics are figured. 5) 95% confidence interval of the form $\theta = (L - B^*) \pm 1.96\, se(L^*)$ is constructed. Here, $se(L^*) = \sqrt{V(L^*)}$ is the standard error for the bootstrap estimate.

The resampling procedure is executed R times to achieve reliable parameter estimation. In this study, R is designated as 100, 1,000, 10,000, and 100,000. Should the bootstrap estimates, represented as θ, display variability in relation to the model estimates θ, it indicates that the proposed model may inadequately capture the true distribution of the observed Malaysian mortality data. Therefore, with an increase in the value of R, the bootstrap estimates of bias, also denoted as θ, are expected to diminish. This trend suggests that the bootstrap estimates are likely converging toward the actual parameter values.

3 Results and Discussion

In this section, non-parametric bootstrap technique is exerted to evaluate the bias and consistency of the model parameters. The simulation process is initiated with the generation of multiple samples derived from the Malaysian mortality dataset. These simulated samples are then applied to the proposed mortality model, resulting in the acquisition of new statistical outcomes, which encompass revised parameters and mortality rates. From these outcomes, metrics such as bias, variance, and confidence levels are computed. This entire procedure is replicated across 100, 1,000, 10,000, and 100,000 scenarios to ensure a comprehensive evaluation.

The mean, bias and standard deviation are tabulated in Table 1. The parameter estimates for α, β and ς are 0.142, 75.3, and 7.69 respectively. Bias is the difference between the mean and the parameter estimate. From Table 1, the bias for every parameter declined as the number of replications increased. This suggests that the bootstrap estimates are converging toward the actual parameter values. Furthermore, as the displayed variation decreased, this also indicates that the proposed model adequately captures the true distribution of the observed Malaysian mortality data.

Table 1: Mean, bias and standard deviation (SD) of α, β and ζ for every replication (R).

Parameter	**Replication**	**Mean**	**Bias**	**Standard Deviation**
α	100	0.145	0.003	0.011
	1000	0.144	0.002	0.011
	10,000	0.143	0.001	0.011
	100,000	0.143	0.001	0.011
β	100	75.69	0.42	1.47
	1000	75.47	0.20	1.43
	10,000	75.44	0.18	1.46
	100,000	75.44	0.18	1.46
ς	100	7.84	0.15	0.64
	1000	7.75	0.07	0.64
	10,000	7.74	0.06	0.64
	100,000	7.74	0.05	0.64

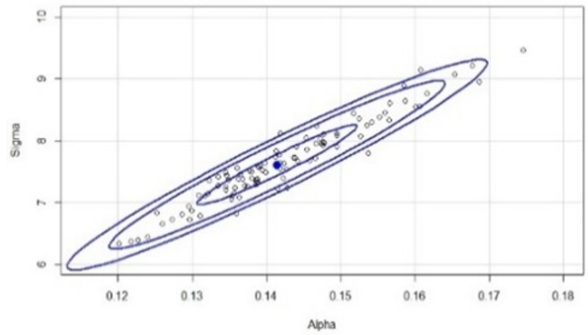
The confidence interval (CI) at a 5% critical value for both males and females, derived using bootstrap simulation with $R = 100$, is presented in Table 2 and graphically represented in Figure 1. It is observed that the interval gap for males is wider, indicating a higher degree of uncertainty in the mortality rates for males in the oldest age bracket. This can be primarily attributed to the lesser availability of data for males compared to females within these age groups. As stated by the Department of Statistics, Malaysia (DOSM), the total number of deaths for individuals aged 80 and above is 88,715 for males and 123,020 for females. Concurrently, the living population within these ages comprises 859,500 males and 1,039,500 females. However, it is anticipated that the overall uncertainties at these ages would be elevated due to the scarcity of data.

Table 2: Bootstrap parameter estimates at 95% Confidence Intervals.

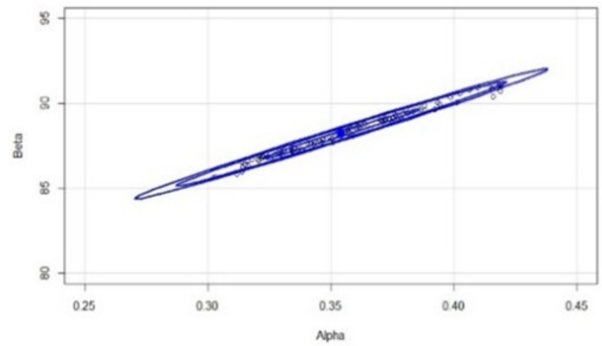
Simulation	**Male**			**Female**		
	α	β	ζ	α	β	ζ
95% lower CI	0.1166	71.963	6.290	0.2854	85.1946	7.2562
95% upper CI	0.1606	77.714	8.781	0.4074	90.7473	8.6961
Gap	0.0440	5.7517	2.4912	0.1220	5.5527	1.4399

The study presents scatterplots of bootstrap replications for all parameters derived from the logistic function used to model male and female mortality data in Malaysia from 2011 to 2021 (Figure 1). These scatterplots provide a visual representation of the uncertainty and variability associated with the estimated parameters. To further illustrate the relationships between the parameters, concentration ellipses are drawn at three different confidence levels (50%, 90%, and 99%) using the estimated covariance matrix of the parameters. The innermost circle is for the confidence interval 99%, the middle circle is for 90% and the outermost circle is for 50% confidence interval. These ellipses help identify the regions where the true parameter values are most likely to fall, given the observed data and the assumptions of the model. It is evident from this figure that the innermost circle, which is the most concentrated circle, signifies that 99% of the parameter estimates are proximate to the actual value, while the remaining estimates fall within the 90% and 50% confidence intervals. The diameter and size of the circle disclose the range between the lower and upper confidence levels. It means that if the circle is small or thin, the death variability is small. It is observed that the circle representing females is smaller than that of males, indicating a lower mortality risk for females compared to males. In general, most of the parameter estimates reside within the ellipses, leading us to conclude that the parameters for the proposed model are unbiased and consistent.

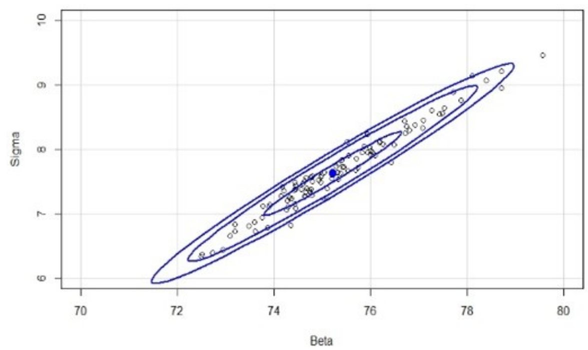
In addition to the consistency and bias test, the mortality estimates are also measured in term of its accuracy using the Root Mean Squared Errors (RMSE) and are compared with six established mortalit models namely Gompertz, Makeham, Beard, Kannisto, Heligman Pollard and Wilmoth. These models are ranked by RMSE, where rank 1 is the model that best fits the data. Table 3 signifies that the proposed model presents the most reliable estimates. It is then followed by the Heligman Pollard (HP) model which has the second higher accuracy and predictive ability rates. HP model achieved this order as it has the greatest number of parameters to capture complex relationship between age and mortality risk. Gompertz are top three as it is well-known as to model the old age mortality rates. Kannisto, Makeham and Beard models results are close together and this is actually expected as the model is an improvement of each other.



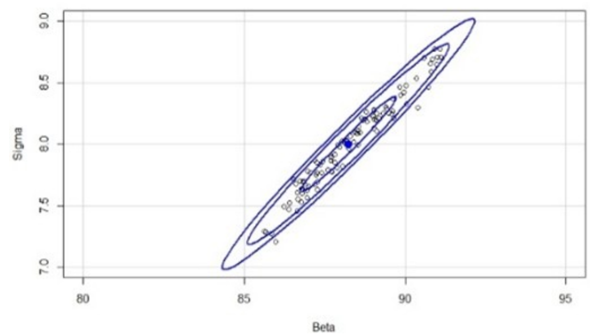
(a) Male α and male β



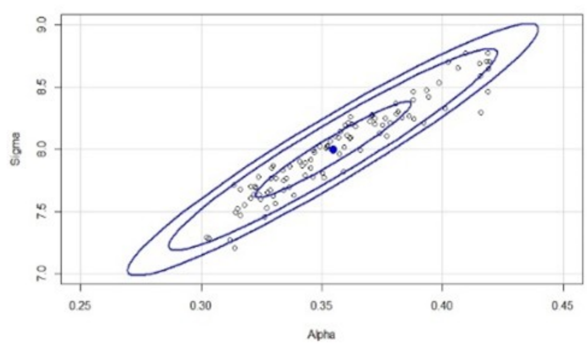
(b) Female α and female ς



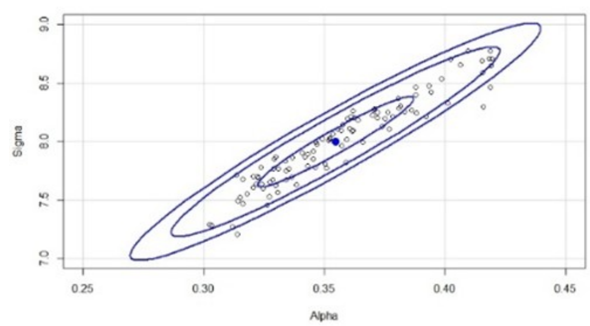
(c) Male β and male ς



(d) Female β and female ς



(e) Male β and male α



(f) Female β and female α

Figure 1: Concentration ellipses at 50%, 90% and 99% confidence level between different parameters for male and female.

Table 3: RMSE and RMSE Rank for Different Mortality Models

Model	**RMSE**	**RMSE Rank**
Hybrid Logistic Spline	0.0182	1
Gompertz	0.0356	3
Makeham	0.0412	6
Beard	0.0536	7
Kannisto	0.0392	5
Heligman Pollard (HP)	0.0180	2
Wilmoth	0.0365	4

4 Conclusion

The decline in late-life mortality and the rapid aging of the population are crucial due to the increased costs associated with retirement and insurance payouts and the increased cost of social care. However, estimating mortality in older age groups is challenging, due to the poor quality of data and small numbers of individuals alive at older ages resulting in low observed death counts.

Ultimately, this research aims to address the identified challenges in mortality estimation. It presents a comprehensive methodology for calculating mortality rates across all age groups, while accounting for uncertainties related to age, gender, and parameter estimations. The approach employs the Akima interpolation technique to derive precise mortality rate estimates for younger populations and introduces an innovative logistic mortality model tailored for older age groups, which often have limited data. This model facilitates the estimation of mortality rates for the oldest cohorts. The validation process for the mortality model utilises a semi-parametric bootstrap method, which assesses various characteristics of the model, including parameter bias, consistency, and robustness. Findings from the bootstrap analysis demonstrate that the proposed model maintains consistency, is unbiased, and exhibits robustness in its estimations. The results obtained from this research are coherent and agrees with [1], [2] and [18]. The research shows that the hybrid method has greatly enhanced mortality estimates, allowing government, demographers, actuaries, pension, and insurance providers to make more accurate predictions for social care expenses, retirement, and insurance payments. As the world are progressing towards AI, future research may incorporate this model into AI model such as in the random forest and applies the model to the country that shares similar mortality trends and life expectancy to Malaysia.

Acknowledgements

This research was funded by Universiti Sains Islam Malaysia under the USIM RACER Research Grant with reference code PPPI/RACER/TAMHIDI/USIM 110223.

References

[1] Dodd, E., Forster, J. J., Bijak, J. and Smith, P. W. Smoothing mortality data: the English Life Tables, 2010 - 2012. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*. 2018. 181(3): 717-735.

[2] Khaliludin, N. I. A., Khalid, Z. M. and Rahman, H. A. Modified Logistic Model for Mortality Rates in Malaysia. Universiti Teknologi Malaysia. 2021.

[3] Khaliludin, N. I. A., Khalid, Z. M. and Rahman, H. A. On estimate of Malaysian mortality rates using interpolation methods. *Matematika*. 2019: 177-186.

[4] Pitacco, E. Heterogeneity in mortality: a survey with an actuarial focus. *European Actuarial Journal*. 2019. 9: 3-30.

[5] Hilton, J., Dodd, E., Forster, J. J. and Smith, P. W. Modelling frontier mortality using Bayesian generalised additive models. *Journal of Official Statistics*. 2021. 37(3): 569-589.

[6] Anderson, T. and Li, Y. Computational challenges in advanced bootstrapping techniques. *Journal of Statistical Computation and Simulation*. 2020.

[7] Richards, S. J. A Hermite-spline model of post-retirement mortality. *Scandinavian Actuarial Journal*. 2020. 2020(2): 110-127.

[8] Tang, K. H., Dodd, E. and Forster, J. J. Joint modelling of male and female mortality rates using adaptive p-splines. *Annals of Actuarial Science*. 2022. 16(1): 119-135.

[9] Department of Statistics Malaysia. Current Population Estimates, Malaysia, 2020. Department of Statistics. 2020.

[10] Department of Statistics Malaysia. Abridged Life Table 2019 - 2020. Department of Statistics Malaysia. 2021.

[11] Department of Statistics Malaysia. Key Findings Population and Housing Census of Malaysia, 2020 Version 2.0. Department of Statistics. 2022.

[12] Lee, J. and Wang, S. Misspecification and its effects on bootstrap confidence intervals. *Journal of Econometrics*. 2019.

[13] Smith, K., Thompson, P. and Patel, R. The dual benefits of semi-parameter bootstrapping: A methodological review. *Journal of Statistical Theory and Practice*. 2021.

[14] Brown, A., Johnson, R. and Smith, L. Parameter sensitivity in semi-parameter bootstrapping: A cautionary note. *Statistics in Medicine*. 2021.

[15] Johnson, M. Navigating the complexities of semi-parameter bootstrapping: A guide for practitioners. *Statistical Methods in Research*. 2023.

[16] Alexandru, M. B. and D. C. The Akima's fitting method for quartic splines. *Journal of Numerical Analysis and Approximation Theory*. 2022. 51(2).

[17] Dimitrova, D. S., Kaishev, V. K., Lattuada, A. and Verrall, R. J. Geometrically designed variable knot splines in generalized (non-) linear models. *Applied Mathematics and Computation*. 2023. 436: 127493.

[18] Dodd, E., Forster, J. J., Bijak, J. and Smith, P. W. Stochastic modelling and projection of mortality improvements using a hybrid parametric/semi-parametric age-period-cohort model. *Scandinavian Actuarial Journal*. 2021. 2021(2): 134-155.

[19] Idais, H., Yasin, M., Pasadas, M. and González, P. Optimal knots allocation in the cubic and bicubic spline interpolation problems. *Mathematics and Computers in Simulation*. 2019. 164: 131-145.

[20] Ministry of Health Malaysia. Malaysian Health at a Glance 2018. Putrajaya, Malaysia: Malaysian Healthcare Performance Unit. 2020.

[21] Roberts, C. and Cheng, L. The limits of semi-parameter bootstrapping in diverse datasets. *International Journal of Data Science*. 2022.

[22] United Nations. United Nations Decade of Healthy Ageing 2021-2030. New York, United States: Department of Economic and Social Affairs, Population Division. 2020.

[23] Zhang, Y., Chen, H. and Liu, F. High-dimensional data analysis with semi-parameter bootstrapping. *Computational Statistics & Data Analysis*. 2022.