IAJ
Indonesian Actuarial Journal
Persatuan Aktuaris Indonesia (PAI)

Vol. 01, No. 02, December (2025)
e-ISSN 3110-6463
https://10.65689/iajvol01no2pp185-202

Comparison of Estimation Methods for the Weibull Distribution in Indonesian Life Tables

Terra Dei Alibazah [1*], Agung Prabowo [2]

[1] *Mathematics Program of Study, Department of Mathematics,Universitas Jenderal Soedriman, Puwrokerto, Indonesia*
[2] *Statistics Program of Study, Department of Mathematics,Universitas Jenderal Soedriman, Puwrokerto, Indonesia*

*Corresponding email: terra.alibazah@gmail.com

Abstract
This study evaluates the performance of the Ordinary Least Squares, Nonlinear Least Squares, and Maximum Likelihood Estimation methods in fitting the two-parameter Weibull distribution to the 2019 Indonesian Life Table and the 2023 Indonesian Population Life Table, focusing on ages 59-111. The lower age bound is selected to isolate the adult and late-life mortality in which the Weibull model's assumptions are most applicable. Although both life tables exhibit monotonically increasing mortality with the same terminal age, they differ in the age at which mortality acceleration becomes pronounced. Goodness-of-fit was assessed through a comparison of root mean squared errors, root mean squared logarithmic errors, and residual plots. The results indicate that the Ordinary Least Squares method, while computationally stable, tends to overestimate survival beyond the terminal age. The Nonlinear Least Squares method better aligns with the empirical survival yet similarly extends the terminal age. The Maximum Likelihood Estimation method provides more realistic terminal ages but inflates survival at infancy and midlife stages. These findings highlight that estimation methods and data segment selection strongly influence the reliability of Weibull-derived life tables. Applications in actuarial and demographic practices require improved estimation strategies to better capture late-age mortality.

Article History:
Received: 14 September 2025
First Revised: 9 December 2025
Accepted: 30 December 2025
Published: 31 December 2025

Keywords:
Indonesian life tables, maximum likelihood estimation, nonlinear least squares, ordinary least squares, Weibull distribution

1. INTRODUCTION

Accurate actuarial calculations highly depend on reliable mortality data, particularly the probability of death and survival at various ages. These life tables serve multiple parties within the actuarial industry, primarily to estimate life insurance premiums and annuities [1], [2], although other industries such as public health have applied life table concepts to examine the duration of vaccine immunity and efficacy as in [3].

One prominent life table employed in fields of demography and actuarial sciences is the 2019 Indonesian Life Table (*Tabel Mortalita Indonesia 2019* or the TMI 2019). This life table was constructed using data of the insured Indonesian population from 52 life insurance companies [4]. The TMI 2019 had no significant differences from previous Indonesian Life Tables with regards to the population number of deaths [5], hence its widespread use in the insurance industry and demographic observations. On the other hand, the 2023 Indonesian Population Life Table (*Tabel Mortalitas Penduduk Indonesia* 2023 or the TMPI 2023) was collaboratively developed by the Indonesian Actuarial Association, the Mathematics and Natural Sciences faculty in Bandung Institute of Technology, and Indonesia's governing ministries. The TMPI 2023 used data provided by Indonesia's national health insurance (JKN), covering 92% of the general population, and captures mortality data from 2018 to 2022 [6], encompassing the COVID-19 pandemic.

These two life tables represent different segments of the Indonesian population. The TMI 2019 reflects a subset of insured individuals excluding the JKN national insurance program. In contrast, the TMPI 2023 captures a significantly broader demographic due to its inclusion of participants from the JKN program. Although both are drawn from insured populations, the two life tables differ in scope and underlying demographic characteristics, with the TMPI 2023 potentially sharing similarities with a population life table;

that is, it might represent a more heterogeneous population in terms of health status, socioeconomic background, lifestyle and occupation [1]. Nonetheless, the two different life tables still provide an opportunity to evaluate the behavior of parameter estimation methods across different population bases.

Parametric models such as De Moivre, Gompertz, Makeham, and Weibull distributions offer interpretable and practical approximations for continuous modeling of life table data. Several recent studies show that these survival functions could fit Indonesian life table data, such as Gompertz's law in [7], [8]. The Gamma-Gompertz-Makeham model also performed well when used to calculate life annuities based directly on mortality data, such as life expectancy in [9]. Meanwhile, the Makeham law was effective when used to construct mortality rates based on crude mortality data and fit using the Heligman-Pollard model before being evaluated alongside the TMPI 2023 [10]. The Weibull distribution is commonly used in survival analysis due to its flexibility and simplicity compared to other models, such as the exponential distribution [11]. However, as noted in [12], the distribution has several limitations. The Weibull shape parameter allows it to represent decreasing, constant, or increasing hazard rates depending on the population segment in question, i.e., decreasing within the early childhood age segment. While this makes the Weibull distribution a theoretically viable model, it also implies that applying the distribution to the entire lifespan in a life table may result in biologically implausible outcomes. The discrete nature of age intervals in life tables means that fitting the Weibull model requires careful selection of age samples. Furthermore, life table data provide aggregated counts rather than exact event times, so information at each age is inherently incomplete. This is analogous to Type-II or interval censored data in survival analysis, where the precise timing of events is unknown, and the selection of estimation method must account for the uncertainty, such as in [13], [14], [15].

A previous study [16] applied the Weibull distribution using the Ordinary Least Squares (OLS) method on segment data from the TMI 2019 ages 62-111, predicting full survival at infant ages and survival past the empirical terminal age. This aligns with the behavior of the Weibull hazard estimation in that it assumes full functionality (survival) in early ages and rather slow failure times (death) as the cumulative survival function approaches 0. Another estimation method, the Nonlinear Least Squares (NLS), is also commonly used to estimate Weibull parameters as performed in [17]. The study found that an optimized NLS method was more efficient in estimating a mixed Weibull distribution compared to another commonly used estimation method, the Maximum Likelihood Estimation (MLE). However, as found in [18], the MLE method performed better than the OLS in estimating Weibull parameters to find failure times and reliability through Monte Carlo simulated data, and was asymptotically efficient with large sample sizes. Another study [19] estimated Weibull parameters using the MLE and OLS methods on simulated and real data, finding that the OLS method performed better when handling small and large sample sizes, as well as estimating shape parameters, while the MLE method was better at handling large sample sizes and estimating scale parameters. Similarly, [20] proposed three new estimators of Weibull parameters, which included a modified MLE method with reduced bias and mean squared error in small samples. The modification outperformed the standard MLE method in estimating shape parameters when the sample size was not very small.

Each estimation method examined in this study carries inherent limitations when applied to life table data. The OLS method assumes equal variance in errors across all ages, which may not hold when modeling survival curves that are often heteroscedastic. Human mortality rates inherently increase with age, and their residual variance tends to increase at older ages. The NLS method accommodates the nonlinear form of the Weibull model, but is sensitive to initial, often guessed values. The MLE method offers desirable statistical properties in deriving incidence rates and corresponding confidence intervals from aggregated data [21]. However, fitting the MLE method to cumulative survival data can produce unstable or implausible estimates as they are not raw, event-based observations. This study therefore adopts an alternative input using the number of individuals who died before reaching the next age, which is more consistent with the method's underlying assumptions.

This study investigates and compares the performance of the OLS, NLS and MLE methods in estimating the two-parameter Weibull distribution using Indonesian mortality data from the TMI 2019 and the TMPI 2023, specifically covering ages 59 which is the Indonesian retirement age as of 2025 [22] to the terminal age

of 111, where mortality patterns are consistent with the distribution's underlying assumptions. Previous studies typically rely on a single estimation method for the Weibull model. This study compares the OLS, NLS, and MLE methods on two distinct population data sets and evaluates how each method influences terminal age estimation. Furthermore, the study also examines whether observed post-pandemic differences in survival patterns reflect genuine demographic change or arise from methodological sensitivities. By doing so, this study establishes a reference point from which subsequent methodological improvements and optimizations may extend.

2. METHODS

2.1 Mortality Data

Life tables often serve only to provide the ages and probability that an individual aged x will die before the next year, or q_x. On the other hand, the probability of an individual aged x surviving to the next year (that is, achieving age $x+1$) can be written as p_x, which is the complement of q_x. In other words,

$$p_x = 1 - q_x. \tag{1}$$

In the special case of a newborn, or life-age-0, the probability of survival to attain age x is

$$_xp_0 = s(x), \tag{2}$$

where *s(x)* is the survival function. Probability theory for life tables always requires that for $x = 0$ (that is, for the case of a newborn), the survival function is always 1, which can be denoted as

$$s(0) = 1, \tag{3}$$

and will be nonincreasing after age 0 [23]. In other words, the survival function will approach zero as an individual gets older. Another assumption for these probability functions is that survival at age x will have the same conditional distribution of survival, hypothesizing that a newborn has survived to age x. Therefore, the probability that an individual aged x surviving to age $x+1$ can be written as

$$_tp_x = \frac{_{x+t}p_x}{_xp_0} = \frac{s(x+t)}{s(x)}, \tag{4}$$

and the probability of death for an individual aged x before reaching age $x+t$ is simply the complement of Equation (4), which can be written as

$$_tq_x = 1 - \frac{s(x+t)}{s(x)}. \tag{5}$$

Additionally, the number of surviving individuals in the cohort at age x, denoted by l_x is obtained using

$$l_{x+1} = l_x \times p_x\,. \tag{6}$$

For consistency and illustrative purposes, the cohort size l_0 which denotes the number of individuals alive at age 0, is set to 100,000 for all life tables. This convention is in line with illustrative methodology as demonstrated in the U.S. model life tables [23]. From there, the number of deaths that occur for the cohort at age x before reaching the next year is calculated using

$$d_x = l_x - l_{x+1}. \tag{7}$$

Finally, the cumulative survival function for age x years is given by

$$s(x) = \frac{l_x}{l_0}, \tag{8}$$

which will result in the initial survival function as noted in (3) such that

$$s(0) = \frac{l_0}{l_0} = 1. \tag{9}$$

Table 1 presents the q_x values for three age groups (ages 0-2, 58-60 and 109-111) to illustrate the structure of the Indonesian life table data. All calculations and parameter estimation procedures are performed separately for each gender and each life table.

Table 1. Probabilities of Death from the TMI 2019 and the TMPI 2023

TMI 2019			**TMPI 2023**		
Age	Male q_x	Female q_x	Age	Male q_x	Female q_x
0	0.00524	0.00266	0	0.009791	0.007880
1	0.00053	0.00041	1	0.002526	0.002096
2	0.00042	0.00031	2	0.001079	0.000900
...	...	...	...	...	...
58	0.00939	0.00601	58	0.014309	0.010777
59	0.00971	0.00636	59	0.015405	0.011524
60	0.00999	0.00671	60	0.016604	0.012311
...	...	...	...	...	...
109	0.55733	0.54477	109	0.538649	0.518532
110	0.59244	0.58702	110	0.566271	0.559684
111	1	1	111	1	1

The lower bound of the estimation interval was set at age 59 based on exploratory inspection of the age-specific probability of death, which exhibits irregular and nonmonotonic behavior at younger ages. Such patterns violate the assumptions of the required shape hazard rate function shape in the two-parameter Weibull model, or what is called the force of mortality in actuarial contexts [23]. The implications of this restriction will subsequently be evaluated by observing the force of mortality function, with its formula given by

$$\mu(x) = \frac{-s'(x)}{s(x)}. \tag{10}$$

2.2 The Weibull Distribution

The Weibull survival function as found within engineering contexts [11] is written as

$$s(x) = \exp(-\lambda x^{\alpha}), \tag{11}$$

where x denotes time, α denotes the shape parameter, and λ denotes the scale parameter, with α and $\lambda > 0$. However, for applications in actuarial science and survival modeling over discrete age intervals, the parameters of the Weibull distribution can be expressed with $\lambda = \beta^{-\alpha}$, where α is interpreted as the shape parameter and $\beta > 0$ is the scale parameter. These parameters, adopted in [16], results in the survival function

$$s(x) = \exp\left(-\left(\frac{x}{\beta}\right)^{\alpha}\right). \tag{12}$$

With time, the hazard increases when $\alpha > 1$ and decreases when $\alpha < 1$. Furthermore, the scale parameter β corresponds to a characteristic life duration where

$$\begin{aligned} s(\beta) &= \exp\left(-\left(\frac{\beta}{\beta}\right)^{\alpha}\right) \\ &= \exp(-1) \\ &\approx 0.3679, \end{aligned} \tag{13}$$

meaning that by age β, the survival rate has decreased to 0.3679 or approximately 36.8%. This provides both a reference point at which a known survival proportion has occurred, and more interpretability for age-based models than using the raw scale parameter of λ in time-to-failure calculations.

The survival function as given by (12) can be rewritten in use for reconstructing a Weibull-derived life table. First, substituting (12) into (4) will yield

$$\begin{aligned} {}_tp_x &= \frac{s(x+t)}{s(x)} \\ &= \frac{\exp\left(-\left(\frac{x+t}{\beta}\right)^{\alpha}\right)}{\exp\left(-\left(\frac{x}{\beta}\right)^{\alpha}\right)} \\ &= \exp\left(-\left(\frac{x+t}{\beta}\right)^{\alpha} + \left(\frac{x}{\beta}\right)^{\alpha}\right) \\ &= \exp\left(-\frac{(x+t)^a - x^a}{\beta^a}\right). \end{aligned} \tag{14}$$

Similarly, rewriting (5) the same way yields

$${}_tq_x = 1 - \exp\left(-\frac{(x+t)^a - x^a}{\beta^a}\right). \tag{15}$$

2.3 Parameter Estimation

OLS estimates model parameters by minimizing the squared differences between observed and predicted values, assuming a linear relationship between them. As written in [11], applying a double-logarithmic transformation to the Weibull survival function yields a linear relationship between $\log(-\log(s(x)))$ and $\log(x)$. Applying such a transformation to (12) yields

$$\log\big(-\log(\mathrm{s(x)})\big) = \alpha\log(\mathrm{x}) - \alpha\log(\beta). \tag{16}$$

In this transformed model, the slope corresponds to the shape parameter α, while the intercept equals $\alpha\log(\beta)$. The scale parameter β can then be obtained through

$$\beta = \exp\left(-\frac{\text{intercept}}{\alpha}\right). \tag{17}$$

NLS estimates model parameters by minimizing squared residuals in a nonlinear model. However, its performance depends on the initial guess of the value that each parameter might have. These guesses are then applied to the model without requiring a log transformation. This means that, unlike OLS, NLS retains the original structure of the Weibull survival function in (12). Let the predicted survival be

$$\hat{s}(x_i) = \exp\left(-\left(\frac{x_i}{\beta}\right)^{\alpha}\right). \tag{18}$$

Then, the NLS finds the parameter values that minimize the sum of squared residuals through

$$\mathrm{SSR}(\alpha,\beta) = \sum_{i=1}^{n}[\mathrm{s_{obs}}(x_i) - \hat{s}(x_i)]^2, \tag{19}$$

where the $\mathrm{s_{obs}}(x_i)$ denotes the empirically observed survival probability at age x_i.

MLE estimates model parameters by maximizing the likelihood that the observed data were generated by the assumed model. For the two-parameter Weibull distribution, the probability density function (p.d.f) previously defined in [16] is given by

$$f(x;\alpha,\beta) = \frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1} \exp\left(-\left(\frac{x}{\beta}\right)^{\alpha}\right), \tag{20}$$

where α is the shape parameter and β is the scale parameter. The log-likelihood function derived from the p.d.f is maximized to obtain the parameter estimates that best fit the empirical data. In this study, MLE is performed by optimizing the log-likelihood function over the observed age-specific death probabilities, given by

$$\ell(\alpha,\beta) = \sum_{i=1}^{n}\left[\log\left(\frac{\alpha}{\beta}\right) + (\alpha-1)\log(x_i) - \left(\frac{x_i}{\beta}\right)^{\alpha}\right]. \tag{21}$$

Prior studies often apply global metaheuristic algorithms and other optimizations to the estimation method, such as to MLE for computational efficiency [24], and to avoid biased parameters as those found in other distributions in [25] with the Gompertz model. This study will apply optimization when applicable to allow meaningful comparison of each estimation method while maintaining their baseline behaviors. NLS is optimized using the Gauss-Newton optimization, known also as the Port algorithm. For MLE, the L-BFGS-B method is used for parameter constraints that avoid memory-inefficient computations.

Parameter estimation of the Weibull model is performed in RStudio, where each Microsoft Excel workbook containing age, q_x, p_x, l_x, d_x, and $s(x)$ values for both life table and sexes is imported as a data frame. The data are then isolated the 59-111 age range of both sexes. For OLS, the double log transformed version of the Weibull survival function is fitted using the $lm()$ function, regressing $\ln\left[-ln\left(s(x)\right)\right]$ on $\ln(x)$. The resulting shape and scale parameters are then used as initial values for the NLS method, applied to the cumulative survival function $s(x)$ using the $\backslash nls()$ function. For MLE, the estimation is based on the number of deaths d_x, with log-likelihood maximization implemented through RStudio's $\backslash optim()$ function on the Weibull p.d.f.

A Student t-test is applied to the shape parameter and intercept from the OLS method to assess whether they are nonzero. The NLS method also permits the use of the Student t-test to evaluate statistical significance. For the MLE method, hypothesis testing is conducted using the Wald test. This test examines the squared ratio of the estimate to its standard error to see if it differs significantly from zero. For all methods, a significance level of 0.01 is used to evaluate the resulting p-values, where p-values below that level are considered significant.

The Student t-test applied for OLS and NLS is given by

$$t = \frac{\hat{\theta}}{SE(\hat{\theta})} \tag{22}$$

while the Wald test applied to the MLE method is given by

$$z = \frac{\hat{\theta}}{SE(\hat{\theta})} \tag{23}$$

where $\hat{\theta}$ denotes the corresponding parameter being tested. For all three methods and their significance tests, the hypotheses are as follows:

$$H_0: \hat{\theta} = 0$$
$$H_1: \hat{\theta} \neq 0$$

In addition to hypothesis testing, the goodness-of-fit of the Weibull models estimated via OLS, NLS, and MLE is evaluated using error-based performance measures in RStudio. Specifically, the root mean square error (RMSE) is employed to quantify the average deviation between $\hat{y}_i$, which denotes the empirical survival probabilities, and y_i, which represents fitted survival probabilities on the original probability scale. The RMSE is a commonly used measure for assessing prediction accuracy in continuous outcomes [26], and is given by

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(\hat{y}_i - y_i)^2}{n}} \tag{24}$$

As a robustness check, the root mean square logarithmic error (RMSLE) is also reported. The RMSLE applies a logarithmic transformation to the empirical and fitted survival probabilities, reducing the influence of extreme deviations and stabilizing relative errors when values approach zero [26]. This property is particularly relevant for survival probabilities at advanced ages, where small absolute differences may correspond to large relative changes. The formula for the RMSLE is given by

$$RMSLE = \sqrt{\frac{1}{n}(\log(\hat{y}_i - 1) - \log(y_i + 1))^2} \tag{25}$$

New life tables are constructed in Microsoft Excel for each sex by substituting the shape parameter α and scale parameter β from each method to the Weibull survival function as given by (12). The survival function can then be used to derive the other life table data. The empirical and Weibull cumulative survival functions are then plotted together for visual comparison to assess and interpret how well each estimation method models mortality. Figure 1 illustrates the research workflow, from the calculation of additional mortality data based on the TMI 2019 and the TMPI 2023 to the parameter estimation, goodness of fit tests and hypothesis testing in RStudio, to the construction of the Weibull-derived life tables.

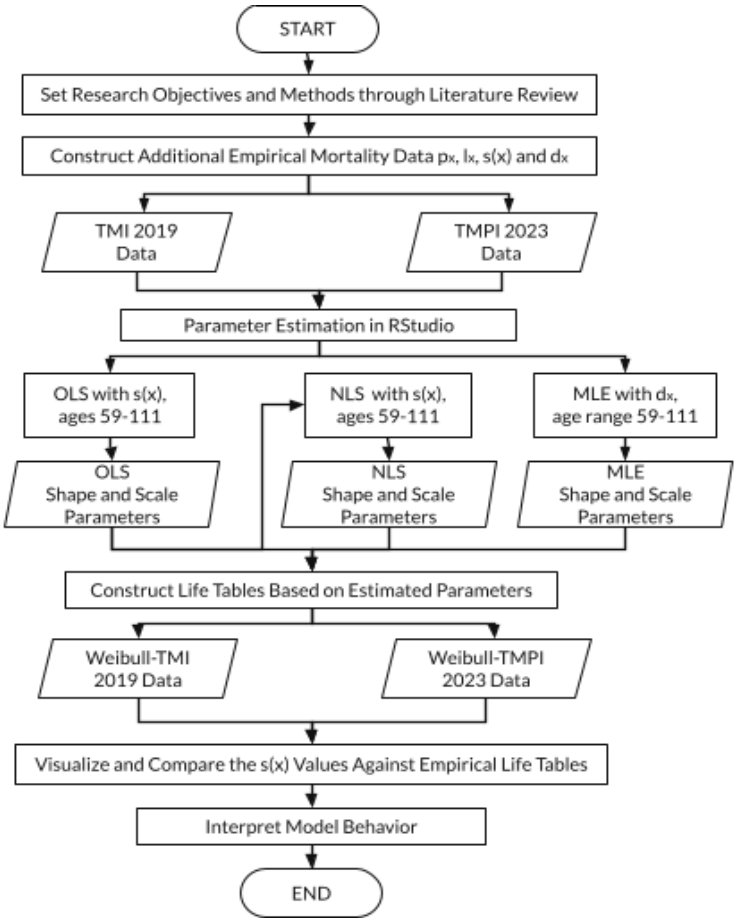

Figure 1. Research workflow

3. RESULT AND DISCUSSION

The p_x, l_x, d_x, and $s(x)$ values based on the TMI 2019 as well as the TMPI 2023 for male and female lives are presented in Table 2 and 3, respectively. They are then imported into RStudio for parameter estimation.

Table 2. TMI 2019 with Additional Mortality Data

TMI 2019 Male						**TMI 2019 Female**					
Age	q_x	p_x	l_x	d_x	$s(x)$	**Age**	q_x	p_x	l_x	d_x	$s(x)$
0	0.0052	0.99476	100000	524	1	**0**	0.00266	0.99734	100000	266	1
1	0.0005	0.99947	99476	53	0.99476	**1**	0.00041	0.99959	99734	41	0.99734
2	0.0004	0.99958	99423	42	0.99423	**2**	0.00031	0.99969	99693	31	0.99693

Table 2. TMI 2019 with Additional Mortality Data (continued)

TMI 2019 Male						**TMI 2019 Female**					
Age	q_x	p_x	l_x	d_x	$s(x)$	**Age**	q_x	p_x	l_x	d_x	$s(x)$
...	...	...	...	...	...	**...**	...	...	...	...	...
58	0.0094	0.99061	89296	838	0.89296	**58**	0.00601	0.99399	93097	560	0.93098
59	0.0097	0.99029	88457	859	0.88457	**59**	0.00636	0.99364	92538	589	0.92538
60	0.0099	0.99001	87598	875	0.875984	**60**	0.0067	0.99329	91949	617	0.91949
...	...	...	...	...	...	**...**	...	...	...	...	...
109	0.5573	0.44267	10	6	0.000103	**109**	0.54477	0.45523	81	44	0.00081
110	0.5924	0.40756	5	3	0.000046	**110**	0.58702	0.41298	37	22	0.00037
111	1	0	2	0	0.000019	**111**	1	0	15	0	0.00015

Table 3. TMPI 2023 with Additional Mortality Data

TMPI 2023 Male						**TMI 2023 Female**					
Age	q_x	p_x	l_x	d_x	$s(x)$	**Age**	q_x	p_x	l_x	d_x	$s(x)$
0	0.009791	0.99021	100000	979	1	**0**	0.00788	0.99212	100000	788	1
1	0.002526	0.99747	99021	250	0.990209	**1**	0.002096	0.9979	99212	208	0.99212
2	0.001079	0.99892	98771	107	0.987708	2	0.0009	0.9991	99004	89	0.99004
...	...	...	...	...	...	**...**	...	...	...	...	...
58	0.01431	0.9857	83915	1201	0.83915	**58**	0.01078	0.98922	93097	935	0.93098
59	0.01541	0.9846	82715	1274	0.82715	**59**	0.01152	0.98848	92538	989	0.92538
60	0.0166	0.9834	81440	1352	0.81441	**60**	0.01231	0.98769	91949	1045	0.91949
...	...	...	...	...	...	**...**	...	...	...	...	...
109	0.53865	0.46135	28	15	0.00028	**109**	0.51853	0.48147	187	97	0.00081
110	0.56627	0.43373	13	7	0.00013	**110**	0.55968	0.44032	90	50	0.00037
111	1	0	6	0	0.00006	**111**	1	0	40	0	0.00015

Figure 2 presents the force of mortality derived from the additional life table data based on the TMI 2019 and the TMPI 2023. The force of mortality functions for both life table and sexes exhibit a bathtub curve, charactertized by decreasing mortality rates at early ages and a predominantly increasing trend after the midlife. This further supports restricting the sample age range to 59-111 as compatible with the assumptions of the two-parameter Weibull distribution..

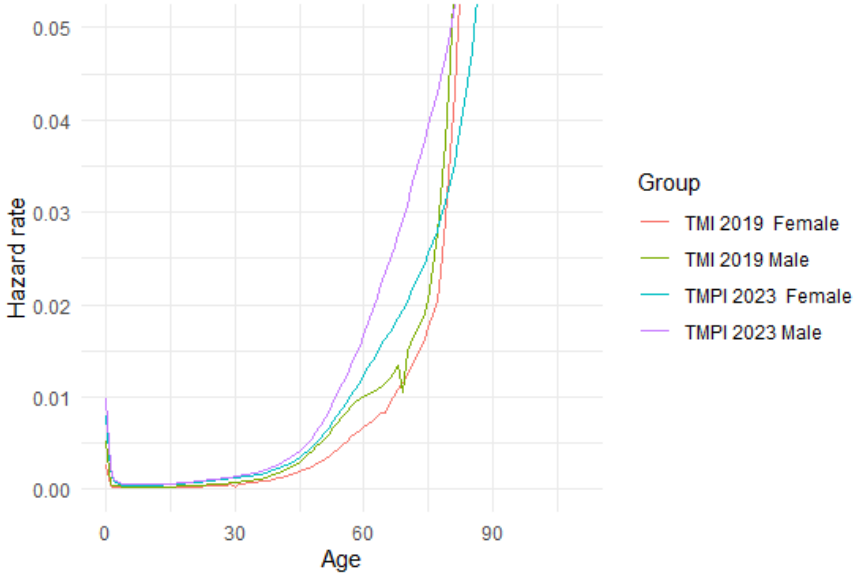

Figure 2. Comparison of the force of mortality functions for the TMI 2019 and the TMPI 2023

Table 4 presents the summary statistics obtained from the OLS, NLS and MLE methods applied to the TMI 2019. All three methods produced nonzero estimates with p-values below the significance level of 0.01. Therefore, the null hypotheses for each method were rejected, indicating that the estimated parameters are statistically significant. Similarly, summaries of the OLS, NLS and MLE results for the TMPI 2023 are presented in Table 5. All methods produced nonzero parameters with associated p-values well below the 0.01 significance level. The null hypotheses are rejected, indicating that these parameters are statistically significant.

Table 4. Summary of Weibull Parameter Estimation using the OLS, NLS and MLE methods on the TMI 2019

Method	Parameter	Estimates		SE		Statistic Test Value		p-value	
		Male	Female	Male	Female	Male	Female	Male	Female
OLS	α (slope)	7.45935	7.610619	0.191267	0.17371	38.99965	43.81163	<0.0001	<0.0001
	Intercept	-33.0359	-34.0597	0.847274	0.76951	-38.99086	-44.2615	<0.0001	<0.0001
	β	83.83053	87.81939	—	—	—	—	—	—
NLS	α	7.70783	8.20448	0.24835	0.18156	31.0364	45.18885	<0.0001	<0.0001
	β	85.98417	89.45235	0.27727	0.18986	310.10626	471.1457	<0.0001	<0.0001
MLE	α	10.54536	10.12106	0.02771	0.02610	380.62317	387.752	<0.0001	<0.0001
	β	86.95265	89.89246	0.02918	0.03075	2980.1104	2923.76	<0.0001	<0.0001

Table 5. Summary of Weibull Parameter Estimation using the OLS, NLS and MLE methods on the TMPI 2023

Method	Parameter	Estimates		SE		Statistic Test Value		p-value	
		Male	Female	Male	Female	Male	Female	Male	Female
OLS	α (slope)	5.992802	5.831957	0.125649	0.16055	47.69442	36.32476	<0.0001	<0.0001
	Intercept	-26.3245	-25.9404	0.556604	0.71121	-47.29483	-36.4739	<0.0001	<0.0001
	β	80.857	85.45388	—	—	—	—	—	—
NLS	α	5.3277	5.479194	0.075942	0.12778	70.15489	42.8789	<0.0001	<0.0001
	β	82.35629	87.55478	0.148503	0.26396	554.57525	331.6869	<0.0001	<0.0001
MLE	α	7.83376	8.128394	0.021326	0.12778	369.28795	365.8267	<0.0001	<0.0001
	β	86.95265	89.69542	0.039848	0.0397	2142.8686	2259.465	<0.0001	<0.0001

For the TMI 2019 male life table, the OLS method estimated $\alpha = 7.45935$ and $\beta = 83.83053$, the NLS method estimated $\alpha = 7.70783$ and $\beta = 85.98417$, and the MLE method estimated $\alpha = 10.54536$ and $\beta = 86.95265$. Similarly, for the TMI 2019 female life table, the parameter values estimated using the OLS method yielded $\alpha = 7.61062$ and $\beta = 87.81939$; the NLS method yielded $\alpha = 8.20448$ and $\beta = 89.45235$; and the MLE method yielded $\alpha = 10.12106$ and $\beta = 89.89246$. The relatively high shape parameter ($\alpha > 1$) for both sexes suggest a rapid mortality rate after the age of the respective β values (approximately after the ages of 83-89). According to the characteristic life as given by (13), around 36.8% of the male population is expected to be alive at about 83.83 years (OLS), 85.98 years (NLS), and 86.95 years (MLE); 36.8% of the female population is expected to be alive at approximately 87.82 years (OLS), 89.45 years (NLS), and 89.89 years (MLE).

On the other hand, for the TMPI 2023 male life table, estimates for the Weibull parameter using the OLS method yielded $\alpha = 5.99280$ and β =80.85701, while the NLS method estimated $\alpha = 5.32770$ and $\beta = 82.35629$, and finally the MLE method produced $\alpha = 7.87556$ and $\beta = 85.38967$. For the female life table, the OLS method estimated $\alpha = 5.83196$ and $\beta = 85.45388$, the NLS estimated $\alpha = 5.479194$ and $\beta = 87.55478$, and the MLE estimated $\alpha = 8.128394$ and $\beta = 89.69542$. As with the Weibull parameters obtained based on the TMI 2019, the relatively high α estimates with both sexes returning values of $\alpha > 1$ imply after the ages listed by the β parameter values, the mortality rate increases rapidly. Furthermore, (13) also implies that around 36.8% of the male population is expected to be alive at about 80.86 years (OLS), 82.36 years (NLS), and 85.39 years (MLE), while 36.8% of the female population is expected to be alive at approximately 85.45 years (OLS), 87.55 years (NLS), and 89.69 years (MLE).

Figure 3 illustrates the standard error residuals produced by each estimation method for male and female lives in the TMI 2019 and the TMPI 2023. Residuals in the OLS method exhibit larger deviations from zero, with noticeable curvature and an increasing spread at older ages. The NLS residuals are generally closer to zero than those of the OLS method, but still display similar departures across the age range. Finally, the MLE residuals tend to cluster most tightly around zero near the terminal age range, though the residuals tend to remain below zero in earlier age ranges.

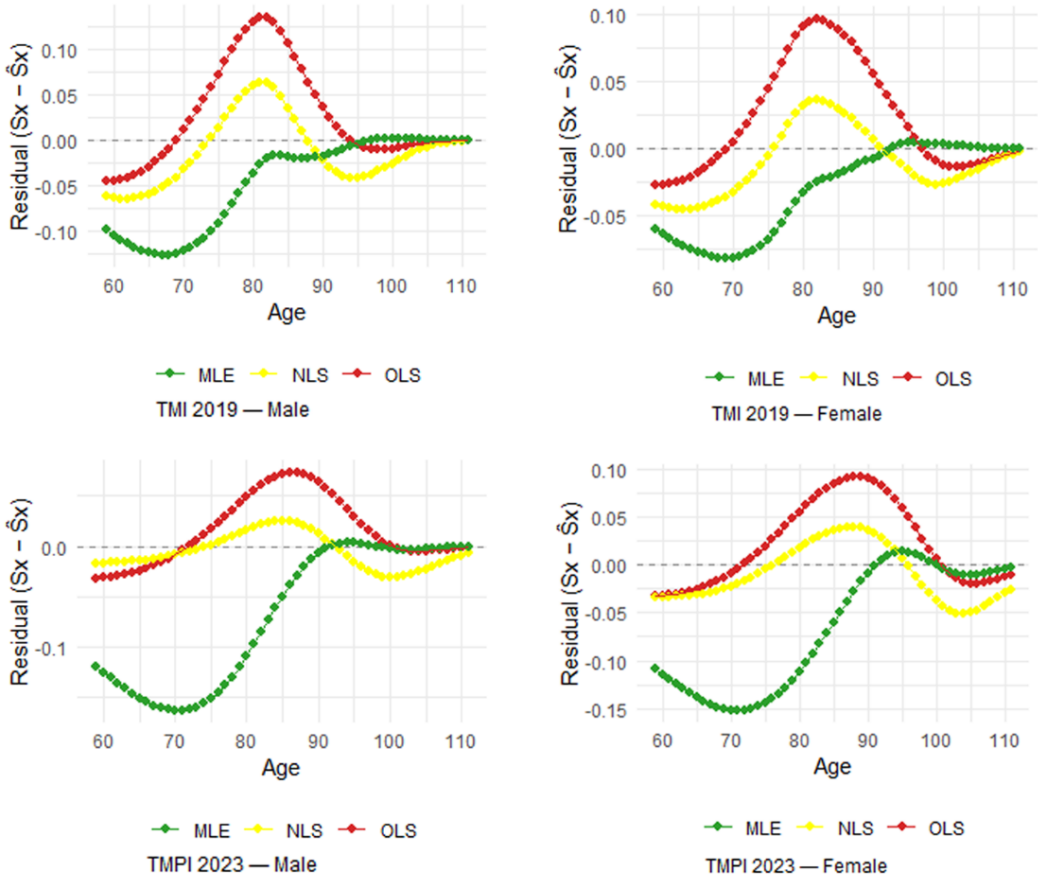

Figure 3. Comparison of standard error residuals of the OLS, NLS and MLE methods

The goodness-of-fit results based on empirical survival probabilities from the TMI 2019 and the corresponding estimates obtained using the OLS, NLS and MLE methods are reported in Table 6. Furthermore, Table 7 presents the goodness-of-fit results for the TMPI 2023 data, which exhibit similar patterns to those observed for the TMI 2019. The NLS method consistently yields the lowest error values across both datasets, while the MLE method produces higher errors compared to the other estimation methods. This suggests that survival probabilities estimated via the MLE method tend to deviate more substantially from the empirical data, resulting in a comparatively poorer fit of the Weibull model within the examined age range.

Table 6. Comparison of RMSE and RMSLE results based on the TMI 2019

	OLS		**NLS**		**MLE**	
	Male	**Female**	**Male**	**Female**	**Male**	**Female**
RMSE	0.0597039	0.0463935	0.0389529	0.0269628	0.0683003	0.0457085
RMSLE	0.0396241	0.0302316	0.0255082	0.01696003	0.0371489	0.0244641

Table 7. Comparison of RMSE and RMSLE results based on the TMPI 2023

	OLS		**NLS**		**MLE**	
	Male	**Female**	**Male**	**Female**	**Male**	**Female**
RMSE	0.0367107	0.0483109	0.0177192	0.0308751	0.0970955	0.0923145
RMSLE	0.028415	0.0356707	0.015073	0.0246954	0.0575267	0.0526656

Table 8 illustrates a preview of the q_x and $s(x)$ values for select ages of male lives, based on the TMI 2019 using the OLS, NLS and MLE methods. The terminal age obtained using the OLS method is 116 years. Meanwhile, the terminal age when using the NLS method is 118 years. Finally, the MLE method yielded a lower terminal age than the empirical terminal age, which is 109 years.

Table 8. Weibull-derived Life Table based on the TMI 2019 for Male Lives

TMI 2019 with OLS			TMI 2019 with NLS			TMI 2019 with MLE		
Age	q_x	$s(x)$	**Age**	q_x	$s(x)$	**Age**	q_x	$s(x)$
0	0.000000	1	**0**	0.000000	1	**0**	0.000000	1
1	0.000000	1	**1**	0.000000	1	**1**	0.000000	1
2	0.000000	1	**2**	0.000000	1	**2**	0.000000	1
...	...	...	**...**	...	...	**...**	...	...
58	0.008677	0.937933520	**58**	0.006750	0.95305018	**58**	0.002758	0.986116
59	0.009675	0.929795537	**59**	0.007559	0.9466175	**59**	0.003241	0.983396
60	0.010769	0.920799415	**60**	0.008449	0.93946191	**60**	0.003799	0.980209
...	...	...	**...**	...	...	**...**	...	...
114	0.486626	0.00004994	**116**	0.497252	0.0000430	**107**	0.601051	0.0001343
115	0.506026	0.00002564	**117**	0.517243	0.0000216	**108**	0.633527	0.0000536
116	1	0.00001266	**118**	1	0.0000104	**109**	1	0.0000196

The Weibull-derived survival curves based on the parameters from the OLS and NLS methods, as illustrated in Figure 4, begin to diverge from the empirical TMI 2019 curve starting in the retirement age range, with both models extending the survival trend beyond 111 years. The OLS and NLS methods yielded terminal ages of approximately 5 and 7 years higher, respectively, than in the TMI 2019. In contrast, the curve based on the MLE method parameters deviates from the empirical curve earlier but converges with it in the elderly age range, yet resulting in a lower estimated terminal age.

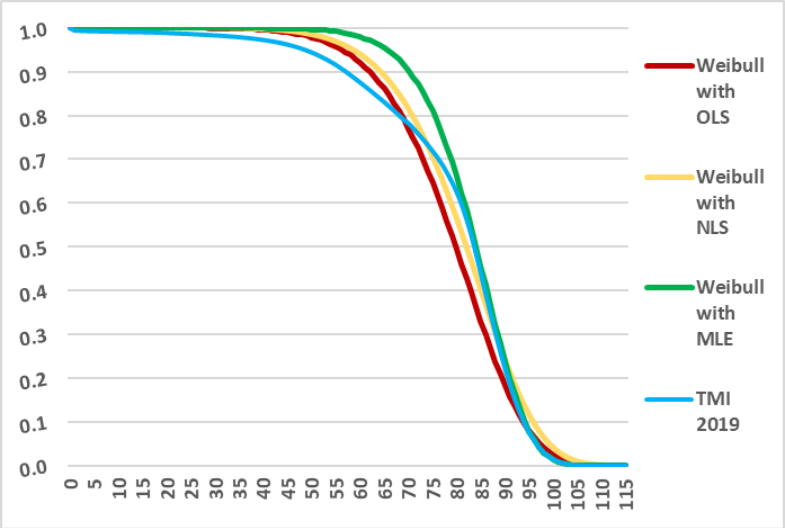

Figure 4. Comparison of cumulative male survival based on the TMI 2019 and its Weibull-derived tables using the OLS, NLS and MLE methods

As with the male cohort, Table 9 presents the q_x and $s(x)$ values from the Weibull-derived female life tables based on the TMI 2019. The terminal age estimated using the OLS method for female lives reached 121 years, the NLS method 120 years, and the MLE method 114 years.

Table 9. Weibull-derived Life Table based on the TMI 2019 for Female Lives

TMI 2019 with OLS			TMI 2019 with NLS			TMI 2019 with MLE		
Age	q_x	$s(x)$	**Age**	q_x	$s(x)$	**Age**	q_x	$s(x)$
0	0.000000	1	**0**	0.000000	1	**0**	0.000000	1
1	0.000000	1	**1**	0.000000	1	**1**	0.000000	1
2	0.000000	1	2	0.000000	1	2	0.000000	1
...	...	...	**...**	...	...	**...**	...	...
58	0.005894	0.95834697	**58**	0.004295	0.97181481	**58**	0.002237	0.98821205
59	0.006590	0.95269859	**59**	0.004852	0.96764055	**59**	0.002611	0.98600116
60	0.007355	0.94642004	**60**	0.005469	0.96294577	**60**	0.003039	0.98342698
...	...	...	**...**	...	...	**...**	...	...
119	0.485260	0.0000411	**118**	0.501238	0.0000611	**112**	0.581675	0.0000954
120	0.504259	0.0000212	**119**	0.522427	0.0000305	**113**	0.611226	0.0000399
121	1	0.0000105	**120**	1	0.0000145	**114**	1	0.0000155

Figure 5 provides a visual comparison of the cumulative survival functions for female lives, illustrating the point at which each estimation method diverges from the TMI 2019 survival data. The Weibull-derived cumulative survival function for female lives, based on all three estimation methods, begin to diverge from the empirical TMI 2019 curve at differing age ranges, and extends the terminal age beyond 111 years. The OLS and NLS methods' estimated survival curves diverge around the midlife and yielded much higher terminal ages. The MLE method's estimated survival probabilities diverge past the retirement age but aligns more closely with the empirical trend afterwards, ultimately producing an estimated terminal age closer to what was reported in TMI 2019.

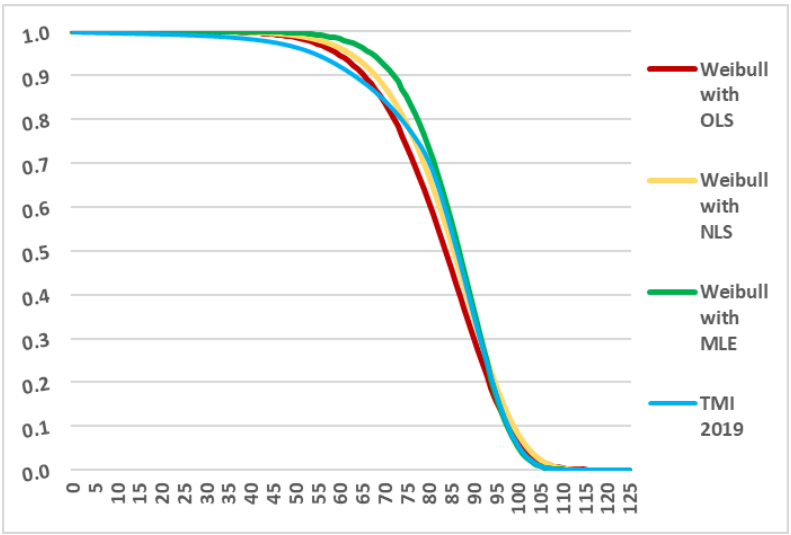

Figure 5. Comparison of cumulative female survival based on TMI 2019 and its Weibull-derived tables using OLS, NLS and MLE

Continuing the analysis with data from TMPI 2023, the estimated q_x and $s(x)$ values for male lives, derived using the same three estimation methods, are reported in Table 10. The estimated terminal age extends to 131 years under the OLS method, while the NLS estimation yields a terminal age of 130 years. The MLE-based life table reaches a considerably lower terminal age of 118 years.

Table 10. Weibull-derived Life Table based on TMPI 2023 for Male Lives

TMPI 2023 with OLS			TMPI 2023 with NLS			TMPI 2023 with MLE		
Age	q_x	$s(x)$	**Age**	q_x	$s(x)$	**Age**	q_x	$s(x)$
0	0.000000	1	**0**	0.000000	1	**0**	0.000000	1
1	0.000000	1	**1**	0.000000	1	**1**	0.000000	1

Table 10. Weibull-derived Life Table based on TMPI 2023 for Male Lives (continued)

2	0.000000	1	**2**	0.000000	0.999999998	**2**	0.000000	1
...	...	...	**...**	...	...	**...**	...	...
58	0.014623	0.87236029	**58**	0.014618	0.856895382	**58**	0.006828	0.953569682
59	0.015904	0.85960409	**59**	0.015722	0.844369365	**59**	0.007669	0.947058675
60	0.017271	0.84593337	**60**	0.016887	0.831094536	**60**	0.008596	0.93979605
...	...	...	...	...	...	...	...	...
120	0.419169	0.0000236	**129**	0.367920	0.0000180	**116**	0.541997	0.0000142
121	0.432312	0.0000137	**130**	0.377660	0.0000114	**117**	0.563140	0.0000065
122	1	0.0000078	**131**	1	0.0000071	**118**	1	0.0000028

Figure 6 presents a visual comparison of the survival curves produced by each method, and illustrates the divergence relative to the empirical life table data. The Weibull-derived survival functions estimated using the OLS and NLS methods diverge from the empirical survival in the early ages and towards the terminal age. However, in the 60-75 age range, the Weibull-derived curve aligns with the TMPI 2023 curve more, before diverging and extending the terminal age. For the MLE estimation method, the survival curve diverges from the infancy to the elderly age ranges, only aligning closer with the empirical curve around the characteristic life age range, and yielding a lower terminal age compared to the OLS and NLS methods.

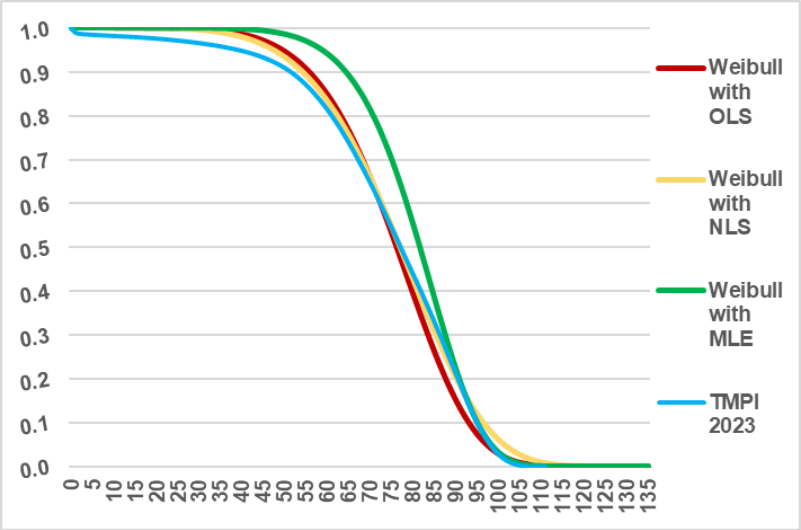

Figure 6. Comparison of cumulative male survival based on the TMPI 2023 and its Weibull-derived tables using the OLS, NLS and MLE methods

Finally, the Weibull-derived life tables based on the TMPI 2023 containing the estimated q_x and $s(x)$ values for female lives are illustrated in Table 9. The terminal ages produced using the OLS and NLS methods are 131 years and 138 years, respectively. The MLE method estimated the terminal age to be 121 years.

Table 9. Weibull-derived Life Table based on the TMPI 2023 for Female Lives

TMPI 2023 with OLS			TMPI 2023 with NLS			TMPI 2023 with MLE		
Age	q_x	$s(x)$	**Age**	q_x	$s(x)$	**Age**	q_x	$s(x)$
0	0.000000	1	**0**	0.000000	1	**0**	0.000000	1
1	0.000000	1	**1**	0.000000	1	**1**	0.000000	1
2	0.000000	1	**2**	0.000000	0.999999	**2**	0.000000	1
...	...	...	**...**	...	...	**...**	...	...
58	0.010879	0.90091705	**58**	0.010230	0.90057528	**58**	0.004299	0.9715102
59	0.011802	0.8911161	**59**	0.011032	0.89136244	**59**	0.004850	0.9673333
60	0.012785	0.88059937	**60**	0.011882	0.88152863	**60**	0.005460	0.9626418
...	...	...	**...**	...	...	**...**	...	...

Table 9. Weibull-derived Life Table based on the TMPI 2023 for Female Lives (continued)

129	**0.398718**	**0.0000160**	**136**	**0.367060**	**0.0000141**	**119**	**0.503713**	**0.000048**
130	0.410190	0.0000096	**137**	0.376609	0.0000089	**120**	0.524543	0.000024
131	1	0.0000057	**138**	1	0.0000056	**121**	1	0.000011

Figure 7 compares the cumulative survival function derived from each estimation method, highlighting how the Weibull-based survival functions diverge from the empirical life table curve. The curves based on the OLS and NLS methods' estimates show noticeable divergence from the empirical TMPI 2023 survival from infancy to the retirement age. However, curves exhibit a relatively closer fit after the characteristic life age range, before diverging again and extending to higher terminal ages. In contrast, the survival curve estimated using the MLE method deviates more substantially from infancy to old age before converging more closely to the empirical curve near the terminal age, although it produces a higher terminal age than was reported in the TMPI 2023.

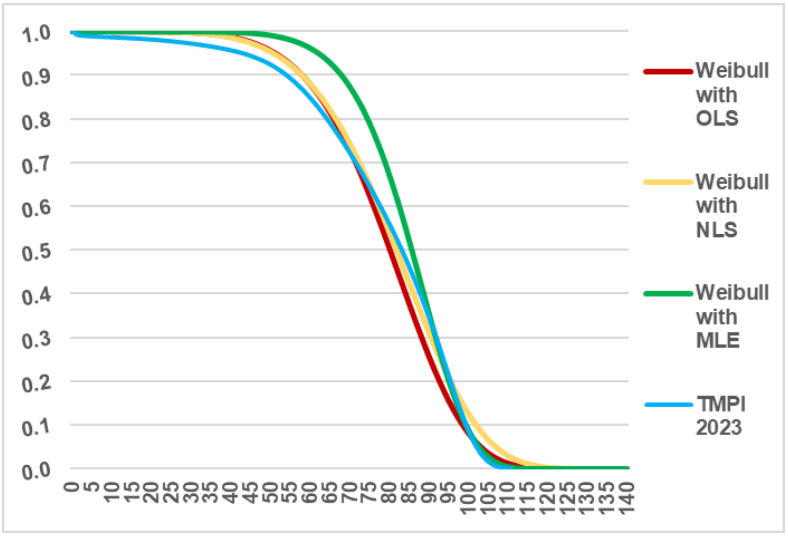

Figure 7. Comparison of cumulative female survival based on the TMPI 2023 and its Weibull-derived tables using the OLS, NLS and MLE methods

The results above show how each estimation method yields different survival patterns, especially at early and terminal ages. Although all three models fit reasonably well with the TMI 2019 and TMPI 2023 in the midlife age intervals, poor fit of certain methods becomes more pronounced at the extremes. This raises important considerations for how each model should be applied in practice.

The survival curves derived from the OLS and NLS methods align reasonably well with the empirical TMPI 2023 data in the midlife range. However, the two methods result in terminal ages often above 120 years. This

likely reflects how sensitive the OLS method is to data size and variance [27]. When the resulting parameters are used again in the NLS method, the model behaves in a similar manner. In the present study, the sample of $n = 53$ spanning ages 59–111 may have introduced survival bias in the middle range. The MLE method, in comparison, overestimates survival well into an individual's midlife. However, the cumulative survival decreases rapidly near the terminal age, producing lower terminal ages compared to the OLS and NLS methods. In other words, the fitted Weibull hazard accelerates nearing the terminal age. The Weibull model is known to overestimate mortality at very old ages [28]. Because the MLE method in this study estimated parameters from d_x at ages 59-111, it is especially sensitive to this limitation.

As a robustness check, the Weibull estimation was repeated using a lower age cutoff of 65 years. The resulting parameter estimates, omitted for brevity, exhibited only minor numerical differences and did not alter the relative goodness-of-fit rankings across the Ordinary Least Squares, Nonlinear Least Squares, and Maximum Likelihood Estimation methods. The consistency of these rankings when using an alternative lower age bound indicates that the main conclusions are robust to reasonable changes in the sample interval, provided the interval lies within an age range where the force of mortality follows a consistent monotonic trend.

4. CONCLUSIONS

This study highlights the strengths and limitations of the OLS, NLS, and MLE methods when fitting a two-parameter Weibull distribution to life table data. The OLS method provided stable estimation but inflated terminal ages, reflecting its sensitivity to sample size. The NLS method accommodated nonlinearity and yielded the smallest goodness-of-fit errors among the three methods, yet similarly extended the terminal ages and remained highly dependent on initial values. The MLE method, applied to d_x counts rather than cumulative survival, offered a more consistent likelihood basis, but still overestimated survival at early to midlife ages.

Across all three methods, the results reflect a known limitation of the Weibull distribution, in which its increasing hazard rate does not always capture the complexity of human mortality over the entire lifespan. This can introduce bias without careful parameterization and age-range selection. A key limitation of the current study is the sensitivity of all methods to tail behavior and sample selection. The lack of optimization in the OLS and NLS methods in particular constrains their reliability when applied across a larger sample of ages where the force of mortality is not monotonic.

Future research could extend this work in several directions. First, improved optimization strategies could be explored to enhance parameter estimation when applying survival models to life table data. Second, alternative mortality models, such as the Gompertz, logistic, or modified Weibull distributions, could be evaluated for comparative performance in estimating parameters. Finally, hybrid or multi-model approaches may offer a more flexible framework for capturing age-specific mortality across broader populations. Such approaches might help overcome the limitations identified in this study and improve the reliability of actuarial and demographic applications.

5. REFERENCES

[1] M. Aalaei and K. Ebrahimnejad, “Modifying premiums for life insurance products using specific mortality tables,” Journal of Mathematics and Modeling in Finance, vol. 5, pp. 137–153, May 2025, doi: 10.22054/JMMF.2025.83627.1165.

[2] G. Siswono, U. Azmi, and W. Syaifudin, “Mortality Projection on Indonesia’s Abridged Life Table to Determine the EPV of Term Annuity,” Jurnal Varian, vol. 4, pp. 159–168, Apr. 2021, doi: 10.30812/varian.v4i2.1094.

[3] M. R. Sternberg et al., “Application of a life table approach to assess duration of BNT162b2 vaccine-derived immunity by age using COVID-19 case surveillance data during the Omicron variant period,” PLOS ONE, vol. 18, no. 9, Sept. 2023, doi: 10.1371/journal.pone.0291678.

[4] Asosiasi Asuransi Jiwa Indonesia, Tabel Mortalitas Indonesia IV. Jakarta: Asosiasi Asuransi Jiwa Indonesia, 2019.

[5] H. Utomo, "Perbandingan Tabel Mortalita Indonesia dan Tabel Mortalita CSO Menggunakan Uji Mann-Whitney dan Uji Kruskal-Wallis," Syntax Literate Jurnal Ilmiah Indonesia, vol. 6, pp. 1210–1215, Mar. 2021, doi: 10.36418/syntax-literate.v6i3.2364.

[6] M. Ruby et al., Tabel Mortalitas Penduduk Indonesia, vol. 1. in Tabel Mortalitas dan Morbiditas Penduduk Indonesia, no. 2, vol. 1. Jakarta: BPJS Kesehatan, 2023.

[7] N. Janna, T. A. Mulyani, Kusumaningsih, M. Dwi, N. A. Fitriyan, A. Prabowo, and Y. A. Ulyana, "Construction of the Mortality Table with Gompertz's Law Using the 2019 TMI Reference," International Journal of Mathematics, Statistics, and Computing, vol. 1, pp. 69–75, Dec. 2023, doi: 10.46336/ijmsc.v1i4.58.

[8] M. R. A. P. Putra, N. Nurjannah, and M. Kurniawaty, "Estimation of Gompertz Mortality Parameter Models on Indonesian Population Mortality Table 2023," CAUCHY: Jurnal Matematika Murni dan Aplikasi, vol. 10, pp. 545–555, June 2025, doi: 10.18860/cauchy.v10i2.33319.

[9] F. Castellares, S. Patricio, and A. Lemonte, "On Gamma-Gompertz-Makeham Assurances and Life Annuities," Austrian Journal of Statistics, vol. 53, pp. 1–24, Sept. 2024, doi: 10.17713/ajs.v53i4.1785.

[10] K. Sari, N. Arfan, A. Dzakiya, and J. Rusli, "Construction of Mortality Rate Using Heligman Pollard Model," ITM Web of Conferences, vol. 75, Feb. 2025, doi: 10.1051/itmconf/20257504003.

[11] D. G. Kleinbaum and M. Klein, Survival Analysis: A Self-Learning Text, 3rd ed. in Statistics for Biology and Health. Springer, 2012.

[12] J. F. Carriere, "Parametric Models for Life Tables," Transactions of Society of Actuaries, vol. 44, pp. 77–100, 1992.

[13] M. Liu, H. Zheng, and J. Yang, "Reliability evaluation for Weibull distribution with heavily Type II censored data," Quality and Reliability Engineering International, vol. 40, pp. 3620–3642, May 2024, doi: 10.1002/qre.3570.

[14] T. Tevetoglu and B. Bertsche, "Bias Corrected Weibull Parameter Estimation and Impact on Confidence Bounds," presented at the ESREL2020, Venice, Italy, Nov. 2020. doi: 10.3850/978-981-14-8593-0_3925-cd.

[15] Y. Yu, J. Gong, and K. Zhu, "Analytical Solution for Parameter Estimation of Weibull Distributions with Interval-Censored Data," Journal of Data Science and Intelligent Systems, May 2025, doi: 10.47852/bonviewJDSIS52024661.

[16] T. D. Alibazah, A. Prabowo, and B. H. Guswanto, "Weibull Mortality Tables based on Indonesian Mortality Table 2019," International Journal of Mathematics, Statistics, and Computing, vol. 1, pp. 76–83, Dec. 2023, doi: 10.46336/ijmsc.v1i4.38.

[17] Z. Lu, L. Dong, and J. Zhou, "Nonlinear Least Squares Estimation for Parameters of Mixed Weibull Distributions by Using Particle Swarm Optimization," IEEE Access, vol. 7, pp. 60545–60554, 2019, doi: 10.1109/ACCESS.2019.2915279.

[18] E. A. Nketiah, L. Chenlong, J. Yingchuan, and B. Dwumah, "Parameter estimation of the Weibull Distribution; Comparison of the Least-Squares Method and the Maximum Likelihood estimation," International Journal of Advanced Engineering Research and Science, vol. 8, no. 9, Oct. 2021, Available: https://ijaers.com/detail/parameter-estimation-of-the-weibull-distribution-comparison-of-the-least-squares-method-and-the-maximum-likelihood-estimation/

[19] N. A. M. Ikbal, S. A. Halim, and N. Ali, "Estimating Weibull Parameters Using Maximum Likelihood Estimation and Ordinary Least Squares: Simulation Study and Application on Meteorological Data," Mathematics and Statistics, vol. 10, no. 2, pp. 269–292, Mar. 2022, doi: 10.13189/ms.2022.100201.

[20] A. Jokiel-Rokita and S. Piątek, "Estimation of parameters and quantiles of the Weibull distribution," Statistical Papers, vol. 65, pp. 1–18, Dec. 2022, doi: 10.1007/s00362-022-01379-9.

[21] S. Voß, A. Hoyer, and R. Brinks, "Maximum likelihood estimation of age-specific incidence rate from prevalence," PLOS One, vol. 20, no. 5, May 2025, doi: 10.1371/journal.pone.0321924.

[22] P. Ardianto and T. A. Prasetyo, "Indonesia Raises Retirement Age to 59 Starting January 2025." Accessed: July 01, 2025. [Online]. Available: https://jakartaglobe.id/business/indonesia-raises-retirement-age-to-59-starting-january-2025

[23] N. L. Bowers, H. U. Gerber, J. C. Hickman, D. A. Jones, and C. J. Nesbitt, Actuarial Mathematics, 2nd ed. The Society of Actuaries, 1997.

[24] M. Teimouri and A. K. Gupta, "On the Three-Parameter Weibull Distribution Shape Parameter Estimation," Journal of Data Science, vol. 11, no. 3, pp. 403–414, Aug. 2022, doi: 10.6339/JDS.2013.11(3).1110.

[25] A. Al-Shomrani, "An Improvement in Maximum Likelihood Estimation of the Gompertz Distribution Parameters," Journal of Statistical Theory and Applications, vol. 22, Apr. 2023, doi: 10.1007/s44199-023-00057-5.

[26] C. Wang, G. Pang, J. Gama, L. Cao, M. Spiliopoulou, V. Kumar, X. Zhou, and X. Wu, Data Science: Foundations and Applications, Proc. 29th Pacific-Asia Conf. on Knowledge Discovery and Data Mining (PAKDD 2025), Sydney, Australia, June 10–13, 2025, Part VI. Singapore: Springer Nature, 2025.

[27] A. L. Burton, "OLS (Linear) Regression," The encyclopedia of research methods in criminology and Criminal Justice, vol. 2, pp. 509–514, 2021.

[28] A. R. Thatcher, V. Kannisto, and J. W. Vaupel, The force of mortality ages 80 to 120. in Odense Monographs on Population Aging, no. 5. Odense University Press, 1998.