Hindawi

Research Article

Poisson Incorporated Credibility Regression Modelling of Systematic Mortality Risk for Populations with Finite Data

Joab Odhiambo [ID],[1,2] **Patrick Weke** [ID],[1] **Philip Ngare** [ID],[1] **Raphael Naryongo** [ID],[3] **and Stanley Sewe** [ID][4]

[1]*Department of Mathematics, University of Nairobi, P. O. Box: 30197-00100, Nairobi, Kenya*
[2]*Department of Mathematics, Meru University of Science and Technology, Meru, Kenya*
[3]*Department of Statistics and Actuarial Science, Jomo Kenyatta University of Agriculture and Technology, Juja, Kenya*
[4]*Department of Mathematics and Actuarial Science, South Eastern Kenya University, Kitui, Kenya*

Correspondence should be addressed to Joab Odhiambo; omishjabali@gmail.com

This study considered the modeling of systematic mortality risk for populations with finite data using the Poisson incorporated Credibility regression model. For novelty, we have included the credibility regression approach to modelling mortality by assuming the number of annual deaths follow a Poisson distribution. Our model shows improvement in precision levels when estimating mortality risk compared to classical models used in European countries. We have illustrated that our model works optimally when using Kenyan mortality data, comparing male and female lives under the different strategies, thus making better predictions than the classical Lee–Carter (LC) and Cairns–Blake–Dowd (CBD) models. The mean absolute forecast error (MAFE), mean absolute percentage forecast error (MAPFE), root mean square error (RMSE), and root mean square forecast error (RMSFE) under the incorporated credibility regression model are much lower than the values obtained without incorporation of the Buhlmann credibility approach. The findings of this research will help insurance companies, pension firms, and government agencies in sub-Saharan countries model and forecast systematic mortality risks accurately. Finally, the results are essential in actuarial modelling and pricing, thus making life assurance products affordable for most people in low-income African countries.

1. Introduction

The concept of stochastic mortality risk modelling was introduced by Lee and Carter[1], making it possible to model both time and age as random variables. Moreover, mortality and longevity risk become random variables that change with age and time. This new approach led to more accurate mortality modelling, thus enabling researchers to get precise results that correspond to the actual demands of the market [2]. However, the primary source of modern risk faced by companies, pension funds, and life companies is longevity. An individual who has taken out life cover will tend to live longer than the expected age calculated by an actuary, thus posing a risk for those firms offering actuarial services.

The uncertainty posed by mortality and longevity risk has been further compounded by increased life expectancy in the 21st century. There has been an increase in life expectancy by over 10 years due to various factors including improved medical services, diet, and increased awareness of life choices. The increase in life expectancy has been experienced especially in developing countries, e.g., Kenya. Therefore, exposure to extreme mortality risk is a crucial issue. However, life companies, governments, and pension firms are yet to formulate new approaches of dealing with or managing these risks. In addition, life companies, governments, and pension funds are yet to innovate ways of coping with or managing systematic mortality risks especially in light of the prevailing global modern financial challenges [3].

The systematic mortality risk poses a massive challenge for the companies and governments that want to avoid insolvency from poor financial decisions that might occur

whilst they operate in the financial markets. However, many extensions of the model [1] have been made, which include [2, 4], among many others.

In many African developing countries such as Kenya, with limited data points, it may be difficult to trace, analyze, and make meaningful inferences when modelling mortality and longevity risk factors. However, the concept of credibility theory, primarily used in nonlife actuarial valuations, can help model mortality and longevity risks. Many researchers use deterministic models for instance in [5], which do not consider the population's stochastic nature of death rates. The model's limitations is the difficulty in parameter estimation due to insufficient data points.

The concept of credibility theory [6, 7] was introduced for mortality and longevity risk modelling by Hardy and Panjer[8]. This concept has been instrumental when dealing with insufficient data in a population such as those from many developing countries, especially in Africa. This study applies a novel approach to modelling mortality using a Poisson distribution but under an incorporated credibility regression, especially when dealing with data limitations [9–11].

The study assumes that the established and limited populations exhibit similar longevity trends and cohort mortality levels instead of the model proposed in [12]. In our applied model, the force of mortality has inverse proportional trends as the population with limited data increases, meaning that as the credibility factor increases, this reduces our over-reliance on an established population when estimating longevity and forecasting mortality from the conventional models.

The Poisson regression model is widely applied in the study of real data problems, such as mortality studies, where the aim is to investigate the number of deaths. In addition, in health insurance, where the target is to explain the number of claims made by individuals [13–16] illustrated how to use the Poisson regression model when dealing with the response variable consisting of count data. Furthermore, [14] demonstrated the almost unbiased ridge estimator in the zero-inflated Poisson regression model, which can be applied to the mortality study for an infinite number of deaths.

Most of the recent actuarial studies on modelling mortality under a framework of credibility theory were made in several studies, including [10, 17–19] in their published papers. The authors first applied Bühlmann's credibility theory to the mortality data of three modern developed countries: Japan, the UK, and the US. In contrast, in the second paper, they applied the concept of Bühlmann credibility to the three conventional models, namely, the model proposed in [1], the model proposed in [20], and the linear relational model proposed in [21] as a way of improving forecasting performances for the UK dataset.

We conduct our valuation using the Kenyan mortality data and compare the obtained values with the tabulated values from the same classical models for illustration purposes. Furthermore, we illustrate how our precision increases as we increase the data that affect the credibility

factor, thus reducing the reliance on the established data [12]. This study helps us estimate the correct number of deaths and lives living beyond the expected life expectancy, which is vital for planning for actuaries, life assurance companies, governments, and pension firms.

Nevertheless, it has been noted that age-specific mortality displayed a downward trend with time. Additionally, suppose a case with a finite mortality data experience for a given age, but an elaborate data experience for the whole age range is provided. In that case, one can apply Poisson incorporated credibility regression techniques to capture the mortality trends [22]. The research uses the Poisson credibility regression approach as a complementary approach to Bühlmann's credibility as [10] when forecasting the mortality rates [4] for developing African countries such as Kenya, with limited data experience.

Finally, Odhiambo et al. [23] proposed using Bühlmann's credibility approach to model systematic mortality risk, especially for sub-Saharan African populations with limited mortality data, such as Kenya. In addition, Odhiambo et al. [24] proposed incorporating the deep learning technique to the Cairns–Blake–Dowd (CBD) model while modelling systematic mortality risk to enhance the accuracy of the models used in cases with limited data for systematic mortality risk modelling and forecasting. These methods improve the accuracy and robustness of the models used in SMR forecasting, thus enhancing the valuation and pricing of actuarial products in Kenya.

The study is organized as follows: in Section 2, we briefly give reviews of the conventional models [1, 20] and the random coefficients regression models. In Section 3, we propose the Poisson-incorporated credibility regression approach as a complementary approach to Bühlmann's credibility when forecasting the mortality of a given population with finite data for novelty purposes instead of conventional Poisson modelling of deaths.

In Section 4, we present the extrapolation methods used to estimate future mortality rates under the Poisson incorporated credibility regression approach. We also conduct an empirical illustration using Kenyan data, both males and females. Furthermore, we evaluate the forecasting performances of the Poisson-incorporated credibility regression approach using various methods proposed in [1, 20], e.g., MAFE, MAPFE, and RMSFE measures. Finally, we use Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) to evaluate and select the best model.

In Section 5, we apply the Bühlmann credibility approach, which is essential in modelling and forecasting the mortality data, especially for countries with limited data in many developing countries globally, mainly in African countries such as Kenya.

Finally, we conclude in Section 6, noting that using the Bühlmann credibility approach application can help improve modelling precision whenever there is a problem with mortality or life table data availability. In the appendix, we have systematic mortality risk values comparing different ages.

2. Poisson Modelling of Systematic Risk Mortality

2.1. Poisson Modelling Preliminaries.

The credibility modelling of mortality as in [8] has been effective when modelling different mortality tables for different countries regarding their demographic characteristics. The ages of a population can be from time x to time t, where $t = 0, 1, 2, 3 \ldots$.

Definition 1. Let D_{x+t} denote the number of people who die with time from time x and $\triangle (x + t)$ that follow a conditional Poisson distribution of

$$D_{x+t} \sim \text{Poisson}\left(\mu (x, t) \times e(x, t)\right), \tag{1}$$

where the values of $x, t = 1, 2, 3 \ldots$ for all $x < t$.

The value $e(x, t)$ denotes the population surviving past a given age from x to $x + t$ exposed to death and $\mu (x, t)$ is defined as the force of mortality of the established population from age x to $x + t$.

The value will be estimated from the established models [20] using poorly established data from a country such as Kenya. It is essential to assume that the rates will change from one discrete age to another depending on an individual's age, as in [25].

We use two mortality models to model the value of $\mu (x, t)$, which will be multiplied by the number of people exposed to calculate the total number of deaths. The results are what will have been experienced within a given population with limited data sets, such as in the case of the Kenyan demographic setup.

The Poisson distribution assumes that the annual deaths are uniformly distributed throughout the study year, called the UDD assumption. There is a constant force of mortality known as the CFM assumption between two integer ages when modelling and forecasting using the mortality models.

2.2. Mathematical Preliminaries of Stochastic Mortality Models.

This section briefly reviews the two classical mortality models, namely, the models proposed in [1, 20], which are applied within the research study in the subsequent subsections.

2.3. The Lee–Carter Mortality Model.

In the original model [1], the logarithm links of the year recorded mortality rates denoted as $M(x, t) = \ln (m(x, t))$ for all ages from $x = 0, 1, 2, \ldots, n - 1$, and a predictor model is given by the following equation:

$$M(x, t) = \ln (m(x, t)) = \alpha_x + \beta_x k_t + \varepsilon (t, x), \tag{2}$$

where α_x is defined as the age parameter reflecting the mean mortality at an exact age of x, k_t is the period parameter indicating the overall mortality trends in year t and β_x is the age parameter indicating the differences from the mean mortality at age x, with changes in the overall level of mortality index.

The error term is defined as $\varepsilon (t, x)$ and is assumed to be normally distributed with mean zero, and a constant variance, σ^2. In addition, it reflects a specific period as well as age effects that have not been captured by the mortality model. Therefore, after assuming that the randomness of the model is independent, identical, and homoscedastic with a mean of zero, as in [1], a way of approximation is suggested using the method of SVD (singular value decomposition), which is subjected to the constraints $\sum_{x=0}^{K-1} \beta_x = 1$ and $\sum_{x=0}^{K-1} k_t = 0$ when obtaining the model estimates of the parameters. The estimates are as follows:

$$\alpha_x = \frac{1}{t_{n-1} - t_0 + 1} \sum_{x=0}^{K-1} \ln (m(x, t)),$$

$$k_t = \sum_{x=0}^{K-1} \left[\ln (m(x, t)) - \alpha_x\right], \tag{3}$$

$$\beta_x = \frac{1}{\sum_{x=0}^{K-1} k_t} \sum_{x=0}^{K-1} \left[\ln (m(x, t)) - \alpha_x\right] k_t.$$

Upon selection of the SVD method of estimation [1], we often apply the use of the time series models when projecting mortality using the parameters. The values of k_t are obtained using the following formula:

$$\widehat{M}(x, t + h) = \widehat{\alpha_x} + \widehat{\beta_x} \times \widehat{k}_{t+h} \quad \text{for } h = 1, 2, 3 \ldots H. \tag{4}$$

2.4. The CBD Mortality Model.

The original model [20] links the logit transformation of the single-year probabilities of mortality or deaths denoted by $M(x, t) = \text{logit } q(t, x)$ with model predictor as follows:

$$M(x, t) = \log itq(t, x) = k_t^{(1)} + k_t^{(2)} (x - \overline{x}) + \varepsilon (t, x). \tag{5}$$

where $k_t^{(1)}$ is defined as the period parameter indicating the overall level of mortality within year t and $k_t^{(2)}$ is the period parameter showing how mortality index affects every age, whilst $\overline{x}$ is defined as the mean age of the selected fitting age interval. $\varepsilon (t, x)$ is the specific random effects that have not been captured on the model [20], and is expected to follow a Gaussian distribution with a mean of zero and a constant standard deviation.

Similarly, as shown in equation (1), we briefly present the model parameter estimates that can be obtained easily by regressing the logit $q(t, x)$ on the value of $(x - \overline{x})$ for every t value:

$$\widehat{k}_t^{(1)} = \frac{1}{t_{n-1} - t_0 + 1} \sum_{x=0}^{K-1} \log itq(x, t), \tag{6}$$

and the values of $\widehat{k}_t^{(2)}$ are given as follows:

$$\widehat{k}_t^{(2)} = \frac{\sum_{x=0}^{K-1} \log itq(x, t)(x - \overline{x})}{\sum_{x=0}^{K-1} (x - \overline{x})^2}. \tag{7}$$

According to [25], we made an assumption that number of deaths follows a Poisson distribution with a mean of $m(t, x) \times E(t, x)$, which means that $m(t, x)$ is the central

deaths rate and $E(t, x)$ is the number of people exposed. However, we estimate the model parameters through the maximization of the log-likelihood function of equation (5).

With an assumption of period estimates being identical and independent for each other, the use of extrapolation using a random walk having a drift parameter (ϑ_i, for $i = 1, 2$) and then mortality forecasts for the period of $h = 1, 2, 3, \ldots, H$ are obtained by the following equation:

$$\widehat{M}(x, t + h) = \left(k_t^{\widehat{(1)}} + \vartheta_1 h\right) + \left(k_t^{\widehat{(2)}} + \widehat{k}_{t+h}\right)(x - \overline{x}). \quad (8)$$

3. The Credibility Regression of Mortality Risk Models

This section deals with a mortality modelling approach within a credibility regression framework with different coefficients. Also, it describes the procedure of parameter estimation while providing a particular case with fixed coefficients within it.

3.1. The Credibility Regression Structure under Random Coefficient (RC)

Definition 2. Let $D(t, x)$ be the number of observed deaths at a given exact age x within t years' time and $E(t, x)$ as the mean population exposure to risk at aged x within year t.

Furthermore, we define the ratio of $D(t, x)/E(t, x)$ as $m(t, x)$, which denotes the age-specific mortality rates as illustrated in [26]. Besides, let $p(x, t) = 1 - q(x, t) = 1 - e^{-m(x,t)}$ denote the one-year probabilities of death, which is the assumption of the constant force of mortality (CFM) within the different individual age groups.

With the response variable assumption defined by $M(x, t)$ as the transformation of natural logarithm and logit of measure of mortality, $m(t, x)$ and $q(x, t)$ respectively for ages $x = 0, 1, 2, 3 \ldots k - 1$, for integer ages $t = 1, 2, 3, \ldots n - 1$, and for calendar years are provided in [27]. We denote B_x as a random age-related risk parameter, which is defined by $M(x) = (M_{(t_0+x)}, M_{(t_1+x)}, M_{(t_2+x)}, \ldots M_{(t_n+x)})\prime$ as mortality index and W_x is the design matrix of the applied explanatory variables under [28]. Thus, the pair describing the evolution of mortality in age x is given by (B_x, Mx), under these assumptions:

(1) The pairs of $(B_{x0}, Mx_0), (B_{x1}, Mx_1), \ldots, (B_{xk-1}, Mx_{k-1})$ are independent while the pairs of $(B_{x1}, B_{x2}, \ldots, B_{xk-1})$ are independent and identically distributed.

(2) The $E(Mx/B_x) = W_x \phi(B_x)$, where W_x is the fixed nxq design matrix having a full rank of $q(< n)$ and $\phi(B_x)$ is an undefined regression vector having a length of p.

(3) The $\text{Cov}(Mx_0/B_x) = \text{diag}(d_{t0t0}(M_x), d_{t1t1}(M_x), \ldots, d_{tn-1tn-1}(M_x))$ where $d_{tt}(Mx) = \sigma_1^2 Mx + \sum_{k=2}^q \sigma_k^2(Mx)W_{kt,x}^2$ with $\sigma_1^2(Mx) = \sigma_{01}^2(Mx) + \sigma_{11}^2(Mx)$ or when written in the matrix formulation:

$$\text{Cov}\left(\frac{Mx}{B_x}\right) = \begin{bmatrix} \sigma_1^2 Mx + \sum_{k=2}^q \sigma_k^2(Mx)W_{t0,x}^2 & \cdots & 0 \\ 0 & \cdots & 0 \\ 0 & \cdots & \sigma_1^2 Mx + \sum_{k=2}^q \sigma_k^2(Mx)W_{tn-1,x}^2 \end{bmatrix}. \quad (9)$$

The structure of the parameters is defined in the following way:

$$b = E(\phi(B_x)),$$
$$\Phi = \text{cov}(\phi(B_x)),$$
$$s^2 = E[\sigma^2(B_x)] = E[\sigma_1^2(B_x), \sigma_2^2(B_x), \ldots \sigma_p^2(B_x)], \quad (10)$$
$$\triangle_x = E\left(\frac{Mx_0}{B_x}\right).$$

In the above regression setting, we can estimate $\triangle x$. Consequently, we estimate the values of regression coefficients using the method of generalized least squares method or GLS. This gives an individual estimator of $\phi(B_x)$ to be

$$\widehat{\phi}_x = \left(W_x^1 \triangle_x^{-1} W_x\right)^{-1} W_x' \triangle_x^{-1} Mx,$$
$$\text{Cov}\left(\frac{\widehat{\phi}_x}{B_x}\right) = \left(W_x' \triangle_x^{-1} W_x\right)^{-1}. \quad (11)$$

3.2. Credibility Regression Using Fixed Coefficients (FC).
When we fix the coefficients of regression and an assumption of some weights [29], it may appear on the regression line. For example, taking the number of people who are exposed as $E(t, x)$, for values of $t = 0, 1, 2, 3 \ldots n - 1$ that is then applied. We apply assumptions (1) and (2) outlined in the previous subsection as they are but with a modification of the assumption (3), where $\text{Cov}(Mx0/B_x) = \sigma^2(Mx)P_x$ and P_x is a square $m \times m$ positive diagonal matrix having the weights, $P_x = \text{diag} [E(t_0, x), E(t_1, x), E(t_2, x), \ldots E(t_{n-1}, x)]$ [27]. Its structural parameters are then defined as follows:

$$a = E(\phi(B_x)),$$
$$\Theta = \text{cov}(\phi(B_x)), \quad (12)$$
$$s^2 = E[\sigma^2(B_x)],$$

whereas the coefficients vector of ordinary least squares estimator of the $\phi(B_x)$ is defined as follows:

$$\widehat{\phi}_x = \left(W_x^1 \triangle_x^{-1} W_x\right)^{-1} W_x' \triangle_x^{-1} Mx. \quad (13)$$

From the above assumptions, it easy to note that the credibility estimator of the model for the fixed coefficients (FC) model is provided as follows:
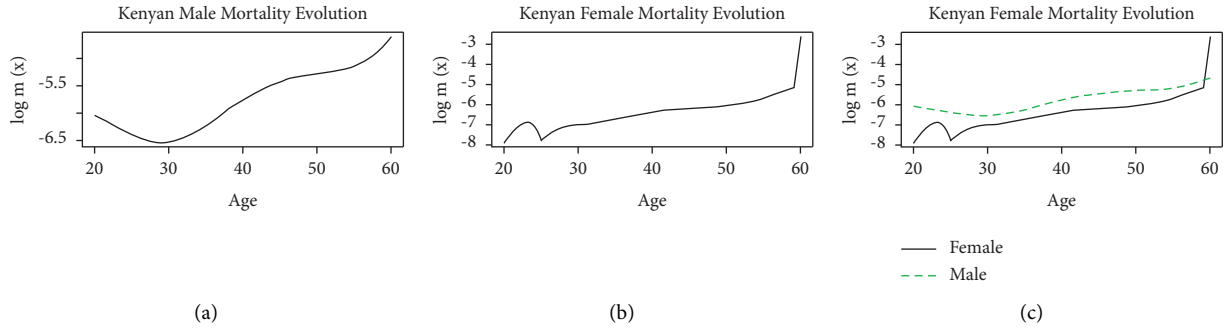
(a)

(b)

(c)

FIGURE 1: Observed $\log m(t, x)$ of the period of 2010 in Kenya, for males (a) and females (b) at the ages 20, 30, 40, and 60. (c) Mean male and female $\log m(t, x)$ values over ages).
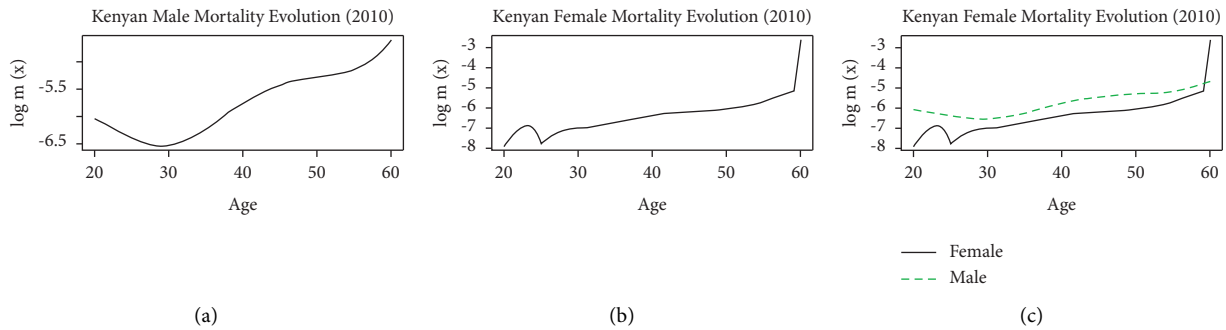

(a)

(b)

(c)

FIGURE 2: Observed $\text{logit} q(t, x)$ of the period 2010 in Kenya, for males (a) and females (b) at the ages 40, 30, 40, and 60. (c) Mean male and female $\text{logit} q(t, x)$ values over ages.

$$\hat{\phi}_x^{FC} = \widehat{O_x}\hat{\phi}_x + (1 - \widehat{O_x})\hat{a}, \tag{14}$$

where $\widehat{O_x} = \hat{v}[\hat{s}^2(W_x^1 \triangle_x^{-1} W_x)^{-1}) + \hat{v}]^{-1}$, known as the estimated credibility factor [29]. Also, we apply the following estimators:

$$\hat{s}^2 = \frac{1}{t_{n-1} - t_0 + 1} \sum_{x=0}^{K-1} (M(x) - \hat{\phi}_x W_x)' \widehat{\triangle_x^{-1}} (M(x) - \hat{\phi}_x W_x),$$

$$\hat{a} = \left(\sum_{x=0}^{K-1} \widehat{O_x}\right)^{-1} * \left(\sum_{x=0}^{K-1} \widehat{O_x}\hat{W}_x\right), \tag{15}$$

$$\hat{v} = \frac{1}{x_{n-1} - x_0} \sum_{x=0}^{K-1} [\widehat{O_x}(\hat{W}_x - \hat{a})(\hat{W}_x - \hat{a})'].$$

Besides, the values of the estimated parameters of $\hat{a}$ and $\hat{v}$ have been determined through the process of iteration, while imposing a constraint of $\hat{v} + \hat{v}'/2 = 0$ after every process of iteration.

Remark 1. We fitted for the random coefficient (RC) as well as the fixed coefficient (FC) of the credibility regression models when modelling mortality rates for a given set of age $x = 0, 1, 2, \ldots, k - 1$ for the year $t = 0, 1, 2, \ldots, n - 1$.

For both models [1, 20], the fitted mortality rates go up to year $n - 1$ and can be wholly written as an estimate of $\widehat{Mx} = \hat{\phi}_x W_x$.

3.3. Extrapolation Methods for Estimating Future Mortality Rates. When fitting current data of the response variable for the mortality data provided, $M(x) = (M_{(t_0+x)}, M_{(t_1+x)}, M_{(t_2+x)}, \ldots M_{(t_n+x)})'$, which is the mortality rate for a given year ahead and can be estimated as follows:

$$\widehat{M}_{t_{n-1}, x} = \widehat{W_{1,x}^c} + \widehat{W_{2,x}^c}(t_{n-1} - t_0 + 2), \tag{16}$$

where c = RC or FC. According to proposals of estimation of a single-year ahead estimates [18], $\widehat{M}_{t_{n-1}+1, x}$ is always fixed to the current fitting span, without having to remove $\widehat{M}_{t_0, x}$; hence, the fitting year span can be extended yearly systematically to $[t_0, t_{n-1} + 1], [t_0, t_{n-1} + 2], \ldots, [t_0, t_{n-1} + m - 1], [t_0, t_{n-1} + m]$. Also, within each of the estimation steps, we fitted the credibility regression models directly on a continuously extended response variable to yield $[\widehat{M}_{t_{n-1}+2, x}, \widehat{M}_{t_{n-1}+3, x}, \ldots \widehat{M}_{t_{n-1}+H-1, x}, \widehat{M}_{t_{n-1}+H, x}]$.

This method of mortality extrapolation of future mortality trends is based on both the first and the current rates, determined after each step of the estimation process [10]. It is important to note that numerical results in the

TABLE 1: Selected fitting and forecasting period.

Fitting ages (x_0, x_{k-1})	Fitting time (t_0, t_{k-1})	Forecasting period $(t_{m-1} + 1, x_{m-1} + H)$
20–100	2010–2019	2020–2030
20–100	2010–2019	2020–2030
20–100	2010–2019	2020–2030
60–100	2010–2019	2020–2030
60–100	2010–2019	2020–2030
60–100	2010–2019	2020–2030

TABLE 2: MAFE values of forecast errors for the duration 2000–2010 for ages 20–80 for the Lee–Carter model.

Fitting period	MAFE Gender	MAFE values							
		Ages 15–84 LC	Lee–Carter LC-Poisson	Random SEM	Coefficients MEM	Fixed EEM	Coefficients SEM	MEM	EEM
2010–2020	Male	0.5452	0.7450	0.4850	0.5354	0.43410	0.43508	0.4365	0.4315
	Female	0.3154	0.4125	0.3195	0.4329	0.2283	0.2545	0.2385	0.2225
2015–2010	Male	0.5768	0.7252	0.4496	0.5428	0.4456	0.4322	0.4345	0.4344
	Female	0.2745	0.3165	0.2598	0.2310	0.2234	0.22357	0.2145	0.21456
2018–2010	Male	0.5678	0.7322	0.4438	0.5325	0.4458	0.4014	0.4433	0.4328
	Female	0.3232	0.4415	0.2615	0.2982	0.21652	0.2267	0.2324	0.2265
Mean		0.4279 (7)	0.5022 (8)	0.3548 (4)	0.4086 (6)	0.3392 (2)	0.3359 (1)	0.3485 (5)	0.3469 (3)

TABLE 3: RMSFE values of forecast errors for the duration 2000–2010 for ages 20–80 for the Lee–Carter model.

Fitting period	MAFE Gender	MAFE values							
		Ages 15–84 LC	Lee–Carter LC-Poisson	Random SEM	Coefficients MEM	Fixed EEM	Coefficients SEM	MEM	EEM
2010–2020	Male	0.5445	0.7235	0.4546	0.5452	0.4321	0.4298	0.4267	0.4115
	Female	0.3250	0.4345	0.3128	0.4099	0.2238	0.2455	0.2430	0.2452
2015–2010	Male	0.5586	0.7348	0.4355	0.5122	0.4544	0.4325	0.4354	0.4359
	Female	0.2358	0.3132	0.2456	0.2223	0.2342	0.2252	0.2481	0.2239
2018–2010	Male	0.5765	0.7235	0.4434	0.5630	0.4321	0.4023	0.4239	0.4235
	Female	0.3091	0.4325	0.2546	0.3540	0.2258	0.2459	0.2128	0.2545
Mean		0.4279 (7)	0.5022 (8)	0.3548 (4)	0.4086 (6)	0.3392 (2)	0.3359 (1)	0.3485 (5)	0.3469 (3)

TABLE 4: MAFE values of forecast errors for the duration 2000–2010 for ages 20–80 for the CBD model.

Fitting period	MAFE Gender	MAFE values							
		Ages 20–80 LC	Lee–Carter LC-Poisson	Random SEM	Coefficients MEM	Fixed EEM	Coefficients SEM	MEM	EEM
2010–2020	Male	0.5285	0.7435	0.4765	0.5230	0.4349	0.4426	0.4214	0.4120
	Female	0.3127	0.4234	0.3142	0.4240	0.2234	0.2234	0.2398	0.2189
2015–2010	Male	0.5650	0.7105	0.4445	0.5187	0.4096	0.4186	0.4345	0.4453
	Female	0.2767	0.3156	0.2588	0.2292	0.2244	0.2243	0.2134	0.2090
2018–2010	Male	0.5888	0.7295	0.4348	0.5455	0.4301	0.4185	0.4143	0.4229
	Female	0.3098	0.4345	0.2326	0.3402	0.2456	0.2335	0.2243	0.2194
Mean		0.4279 (7)	0.5022 (8)	0.3548 (4)	0.4086 (6)	0.3392 (2)	0.3359 (1)	0.3485 (5)	0.3469 (3)

following section justify the methods applied efficiently in actuarial practice, especially in designing and pricing life assurance products and annuities sold to policyholders.

4. Data Analytics of Mortality Risk Modelling

4.1. Empirical Mortality Trend Demonstration. This section fits the models proposed in [1, 20] and the credibility regression models on Kenyan population mortality. After that, we evaluate the forecasting results using the two statistical rations, namely, mean absolute forecast error (MAFE) and the root mean of squared forecast error (RMSFE) measures.

Kenyan data have a limitation in historical mortality observation numbers (2010–2020) and are structured according to year, age, and sex. Moreover, in many life insurance datasets,

TABLE 5: RMSFE values of forecast errors for the duration 2000–2010 for ages 20–80 for the CBD model.

Fitting period	MAFE Gender	Ages 20–80 LC	Lee–Carter LC-Poisson	Random SEM	MAFE values Coefficients MEM	Fixed EEM	Coefficients SEM	MEM	EEM
2010–2020	Male	0.5565	0.7558	0.4865	0.5455	0.4418	0.4508	0.4426	0.4235
	Female	0.3365	0.4545	0.3285	0.4302	0.2325	0.2535	0.2436	0.2250
2015–2010	Male	0.5865	0.7258	0.4546	0.5250	0.4418	0.4208	0.4465	0.4494
	Female	0.2808	0.3239	0.2693	0.2328	0.2274	0.2276	0.2164	0.2160
2018–2010	Male	0.5605	0.7355	0.4438	0.5550	0.4181	0.4067	0.4392	0.4339
	Female	0.3182	0.4465	0.2642	0.3012	0.2266	0.2353	0.2316	0.2345
Mean		0.4279 (7)	0.5022 (8)	0.3548 (4)	0.4086 (6)	0.3392 (2)	0.3359 (1)	0.3485 (5)	0.3469 (3)

similar limitations do currently exist. This credibility regression can proficiently get the trends of underlying mortality data, particularly in those cases where a limited mortality experience during a specific period for a specific age is experienced but an elaborate experience for the complete age range (Kenyan mortality data). It makes regression methods easy to apply in several cases when dealing with larger mortality tables of developed countries [30].

Mortality evolution for the period 2010–2020 in Kenya is illustrated in Figures 1 and 2 for log $(m(t, x))$ and logit $(q(t, x))$. Both systematic mortality risk measures show linearity for discrete-time ages $x = 40, 60,$ and 80 of males and females (Figures 1 and 2) as well as females. Also, mean mortality decline has shown a vivid downward trend over time (right panels of Figures 1 and 2) for both males and females.

Remark 2. Figures 1 and 2 show the increase in the observed values of log $m(t, x)$ and logit $q(t, x)$ for the 2010 period in Kenyan, for males, females, and combined males and females at the ages 20, 30, 40, and 60. There is a general decrease in mortality between ages 25 and 30 before an upsurge of the trends in the remainder of the lifetimes of both males and females in Kenya.

4.2. Forecasting of Mortality Results. For purposes of numerical illustration in the following section, we applied the empirical data of age-specific mortality rates of $m(t, x)$ from 2010 to 2020 for both genders, such as males and females, from ages 20 years to 80 years. Besides, the age span choice as in similar studies [19], mainly when corresponding to a person's age up to the overall level of complete expectation in the developed countries.

To ensure robustness relative to the fitting range of data changes, we have applied two age and three-period spans when extracting forecasts for a 10-year ($H = 10$) forecasting horizon, as listed in Table 1. Specifically, in using the FC model, we used $M_y = I$ as proportions or weights.

The illustration of credibility regression methods, the LC model and the CBD model, has been implemented in the R program to use Poisson LC and CBD models; while fitting these methods, we have used the "Life Metrics," and "StMoMo" packages in the R statistical programming language [31].

As it retains the concept of linearity over all the available fitting period, the logarithmic transform $\widehat{M}_{t_{n-1},x}$ is used for all the age-specific mortality rates and the logit

TABLE 6: AIC and BIC values for the two models.

Model	AIC	BIC
Males		
Lee–Carter (LC)	2150.287	2893.905
Cairns–Blake–Dowd (CBD)	1643.186	1234.665
Females		
Lee–Carter (LC)	2045.857	2786.966
Cairns–Blake–Dowd (CBD)	1467.895	1124.617

transformation of $\widehat{M}_{t_{n-1},x} = \log itq(t, x) = q(x, t)/1 - q(x, t)$ for the given-year death probabilities.

When forecasting errors that were then calculated over the 10-year forecasting horizon when using MAFE and RMSFE, smaller values can indicate a much better performance when forecasting. We calculate the mean or average of MAFE and RMSFE values using the following equation:

$$
\begin{aligned}
\text{MAFE} = &\frac{1}{H \times (x_{k+1} - x_0 + 1)} \\
&\cdot \sum_{g=1}^{H} \sum_{x=0}^{x_k} \left| \widehat{m}(t_{n-1} + h, x) - m(t_{n-1} + h, x) \right| \times 100,
\end{aligned}
\tag{17}
$$

$$
\text{RMSFE} = \sqrt{\frac{1}{H \times (x_{k+1} - x_0 + 1)} \sum_{g=1}^{H} \sum_{x=0}^{x_k} (\widehat{m}(t_{n-1} + h, x) - m(t_{n-1} + h, x))^2 \times 100}.
\tag{18}
$$

Similarly, in the case of using $M_{t,x} = \log itq(t, x)$ as a given response variable, $m(t, x)$ should be replaced by the $q(t, x)$ as per equations (17) and (18), respectively.

When the forecast accuracy results at a percentage (%) scale, which are evaluated over the period 2010–2020. The values of MAFE and RMSFE values especially for fitting ages 20–80, which are using $M_{t,x} = \log m(t, x)$, are listed in Tables 2 and 3, respectively; whereas, the corresponding values for ages 60–100 with $M_{t,x} = \log m(t, x)$ are listed in Tables 4 and 5, respectively. It is important to note that the CBD model has been included only for comparisons when fitting ages 55–84, which has particularly been designed for the higher ages.

For both genders, males and females, the results are tabulated in Tables 2 and 3 when fitting ages 20–60 and Tables 4 and 5 for ages 60–80, thus indicating that, for every fitting period, credibility regression models outperform LC and CBD models for both error measures.
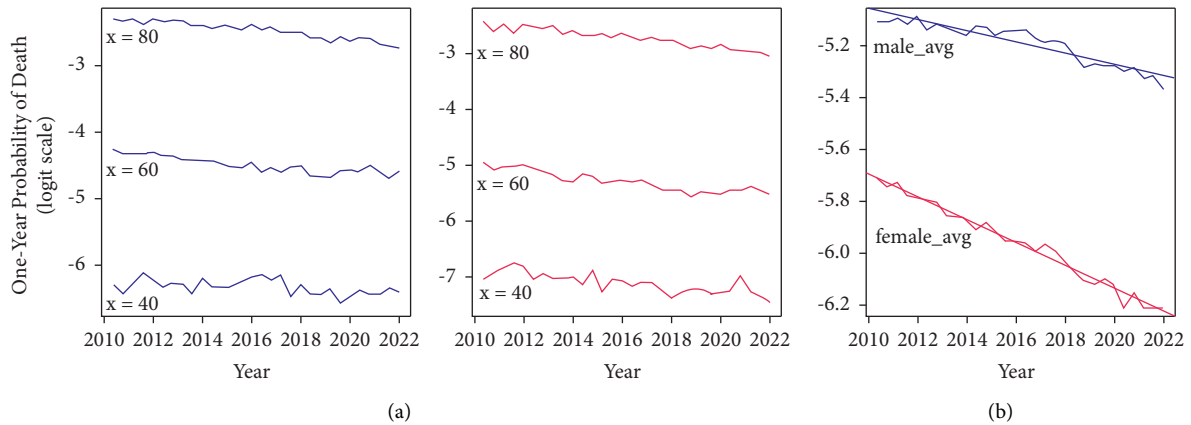
(a)

(b)

FIGURE 3: A linear trend of the recorded logit $q(t, x)$ of the period 2010–2020 in Kenya, for both males (a) and females (b) at the ages of 40, 60, and 80.
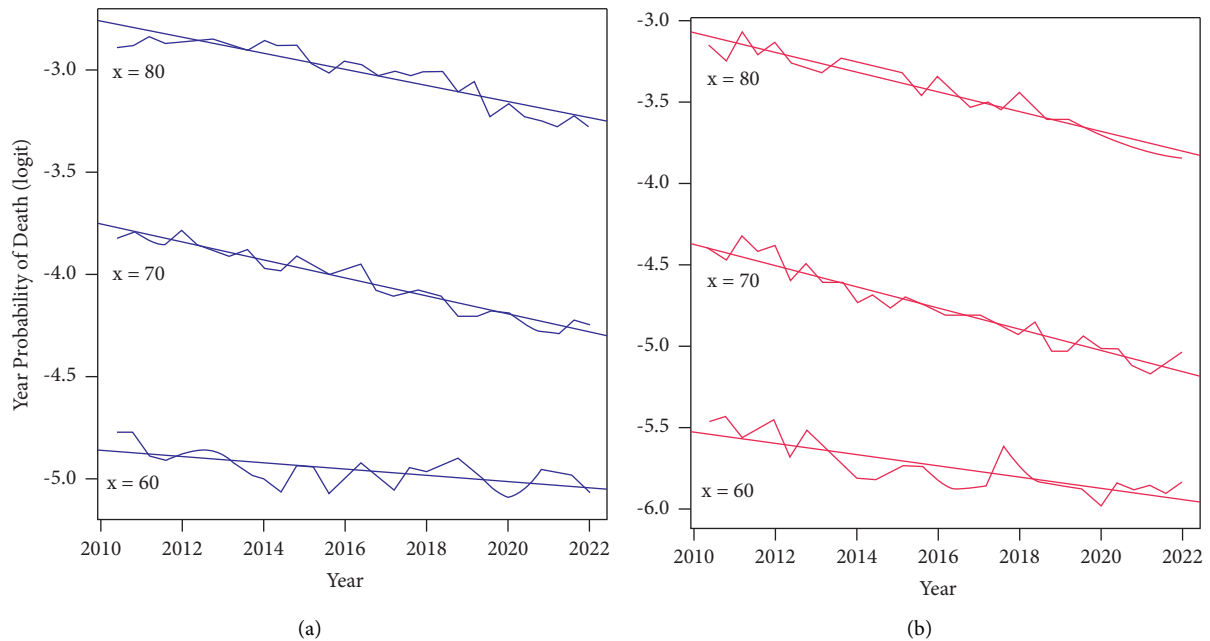

(a)

(b)

FIGURE 4: AFE values against the age of logit $q(2000 + h, x), h = 1, \ldots, 10$ between the actual rates and the rates produced from the best performing models without credibility for males (a) and females (b) over the period of 2010–2020, fitted to pension persons of ages 60–80.

The mean values are provided in the last rows of every measure's suitability over the whole period. To be more precise, for specific ages 15–84, the FC-MEM and FC-SEM produce, thus making the smallest mean of MAFE and RMSFE. In contrast, for ages 60–80, RC-MEM performs better in the mean on both measures, indicating that it predicts higher ages based on the recent mortality trends.

Table 2 lists the MAFE values of forecast errors for the duration 2000–2010 for ages 20–80 for the Lee–Carter Model.

Table 3 lists the RMSFE values of forecast errors for the duration 2000–2010 for ages 20–80 for the Lee–Carter model.

Table 4 lists the MAFE values of forecast errors for the duration 2000–2010 for ages 20–80 for the CBD model.

Table 5 lists the RMSFE values of forecast errors for the duration 2000–2010 for ages 20–80 for the CBD model.

Remark 3. From Tables 2–5, the MAFE and RMSFE values are calculated based on the Lee–Carter and CBD models, especially for prediction values. They show that the fitting period and credibility regression models outperform LC and CBD models for both measures of errors, making them vital for predicting the mortality models.

4.3. Model Selection. From the two models, we will look at the best among the two in systematic mortality risk modelling. We do model selection using the log-likelihood and the number of the effective parameters along with AIC and BIC values of the mortality models for males and females. Table 6 lists the AIC and BIC values for both models.

The AIC and BIC values for the two models show that Cairns–Blake–Dowd (CBD) model is the better model
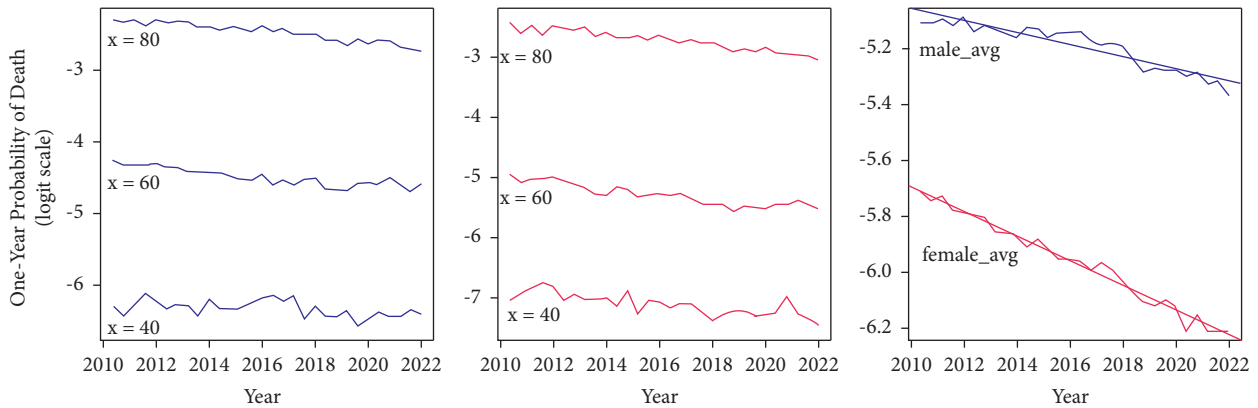
FIGURE 5: Intercept and slope estimates of logit$q\,(2000 + h, x)$ for $h = 1, 2, \ldots, 10$ and ages $x = 60$ for males and $x = 80$ for females, with credibility for FC-MEM and RC-MEM and without credibility for the mortality models.
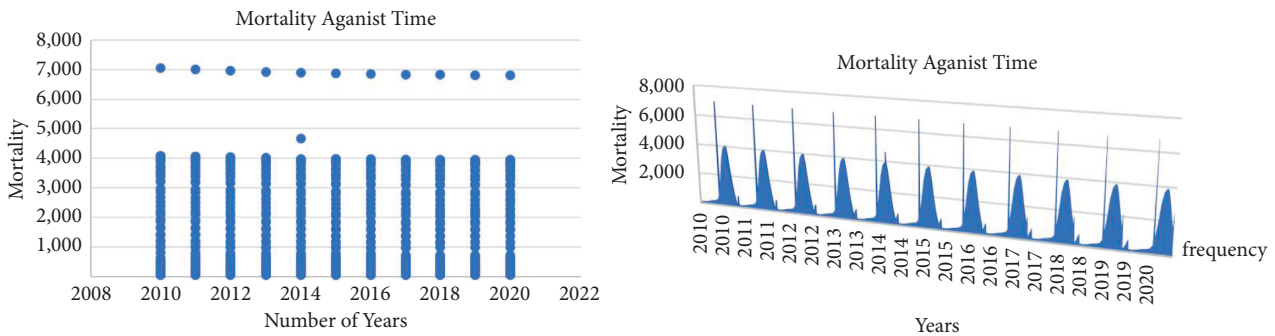

FIGURE 6: Poisson incorporated credibility regression modelling of mortality for Kenyan males.

compared to the values of Lee–Carter (LC) model. Thus, the Akaike Information Criteria and Bayesian Information Criteria help in evaluating and selecting the best model of the two with the model. The model fits well for females and males from the AIC and BIC values, showing lower when compared to the males. In addition, it proves lower values of systematic mortality risk when modelling.

Using the Bühlmann credibility approach in mortality modelling is essential when the country is experiencing limited data for modelling and predicting systematic mortality risk. Ultimately, this is helpful during the correct pricing of the life assurance products sold in the market for the consumers.

4.4. Effects of Credibility Approach on Stochastic Mortality Models.
In the previous subsection, we used the proposed credibility regression methods when estimating the actual mortality trends for a specific age through the proportion of the mortality trends for this age and the mean trend over a much wider group of ages, which provide more information:

$$\text{AFE}_x = |\text{logit}\,(\widehat{q}\,(2000 + h, x) - q\,(2000 + h, x))| \times 100\%. \quad (19)$$

Remark 4. Figures 3 to 5 show the intercept as well as slope estimates of logit$q\,(2000 + h, x)$ for $h = 1, 2, \ldots, 10$ and ages $x = 60$ for males and $x = 80$ for females, with credibility for FC-MEM and RC-MEM and without credibility for the

mortality models, making the incorporated credibility more accurate in predicting the mortality models from the two mortality risk models, namely, Lee–Carter and CBD models.

On the effect of the credibility on the trends, we have the models listed below.

The Poisson incorporated credibility regression modelling of mortality, especially for the populations with finite data, can be determined from the Poison distribution as follows:

Figure 6 shows the Poisson incorporated credibility regression modelling of mortality for Kenyan males, which shows how deaths of males have been trending over the last decade based on the two mortality models, namely, the Lee–Carter and CBD models.

Figure 7 shows the Poisson incorporated credibility regression modelling of mortality for Kenyan females, which shows how deaths of males have been trending over the last decade based on the two mortality models, namely, Lee–Carter and CBD models.

Figure 8 shows the Poisson incorporated credibility regression modelling of mortality for Kenyan males and females. It also shows how deaths of males have been trending over the last decade based on the two mortality models, namely, Lee–Carter and CBD models.

Remark 5. Figures 6 to 8 illustrate Poisson incorporated credibility regression modelling of mortality for both Kenyan males and females, which is critical in modelling the
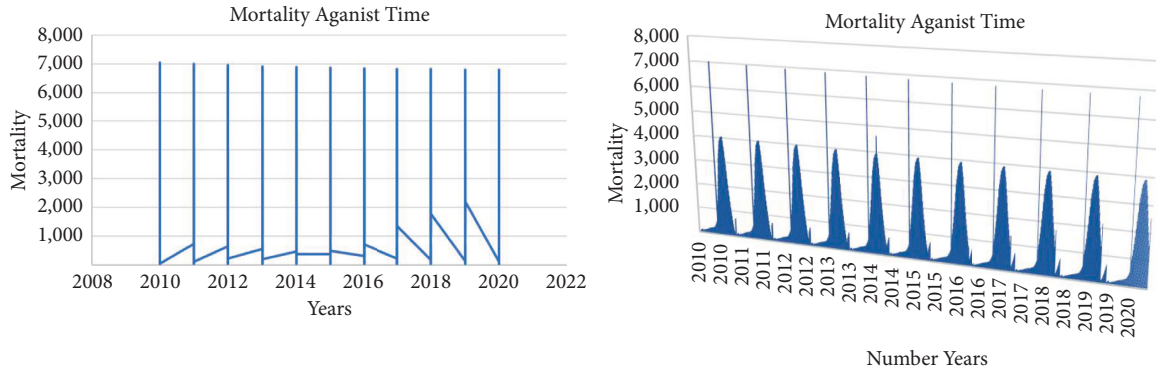
FIGURE 7: Poisson incorporated credibility regression modelling of mortality for Kenyan females.
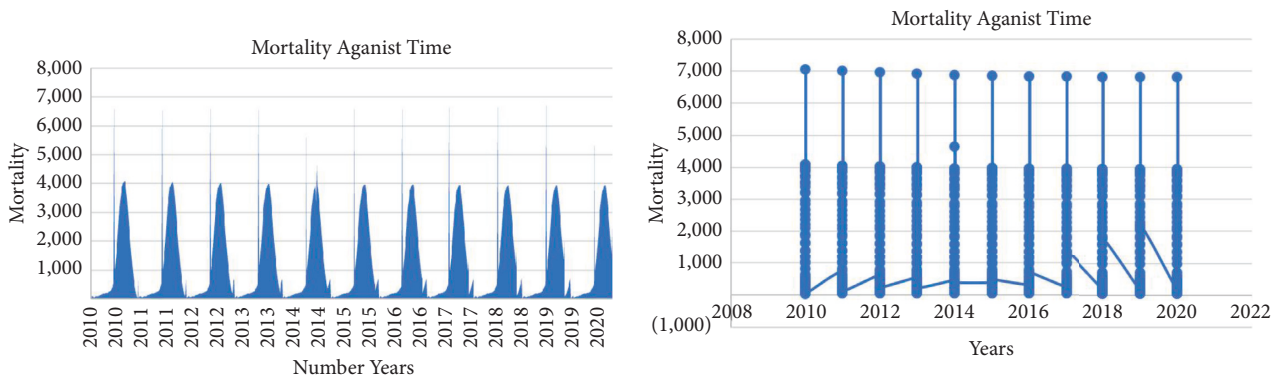

FIGURE 8: Poisson incorporated credibility regression modelling of mortality for both Kenyan males and females.

TABLE 7: Trends for mortality models.

Trends					
Male			Female		
Model	Trend	Slope	Model	Trend	Slope
Actual	−4.803	−0.027	Actual	−5.58	−0.039
LC	−4.795	−0.016	LC	−5.53	−0.025
CBD	−4.800	−0.017	CBD	−5.54	−0.028
FC-MEM	−4.803	−0.019	FC-MEM	−5.57	−0.031

life assurance products sold in the Kenyan market. The deaths have decreased from the data since people live longer than expected, thus there is a massive challenge for the financial institutions making these life annuities and assurance products.

5. Application of the Bühlmann Credibility Approach

[32] proposed a Bühlmann credibility approach that can be used when forecasting mortality rates for both males and females in the three most developed countries: Japan, the UK, and the US. Our model improves its model by modelling it as an incorporated Poisson distribution.

Definition 3. Let the general regression model be $Z_y = (1, 1, \ldots 1)\prime$ and $W_y = I$ for $y = y_0, y_1, \ldots y_{k-1}$, then it

is easy to estimate the parameters of the regression model as follows:

$$\widehat{s}^2 = \frac{1}{(y_{k-1} - y_0 + 1)(t_{n-1} - t_0)} \sum_{y=y_0}^{y_{k-1}} \sum_{t=t_0}^{t_{n-1}} (Q_{t,y} - \overline{Q_y}),$$

$$\widehat{b} = \frac{1}{(y_{k-1} - y_0 + 1)} \sum_{y=y_0}^{y_{k-1}} \overline{Q_y},$$

$$\widehat{U} = \frac{1}{(y_{k-1} - y_0)} \sum_{y=y_0}^{y_{k-1}} (\overline{Q_y} - \overline{Q}) - \frac{\widehat{s}^2}{(t_{n-1} - t_0 + 1)},$$

$$K = [(t_{n-1} - t_0 + 1)]\widehat{U} [\widehat{s}^2 + (t_{n-1} - t_0 + 1)]^{-1},$$

(20)

where $\widehat{U}$ and K are mortality index and Bühlmann credibility factor estimates, respectively.

TABLE 8: MAFE values of forecast errors during the period 2010–2020 for ages 20–60.

Fitting period	MAFE for ages 20–60	Bühlmann methods	Regression	Methods	RC	Regression	Methods	FC	
	Gender	EW	MW	SEM	MEM	EEM	SEM	MEM	EEM
2010–2020	Male	0.2818	0.2801	0.1697	0.1544	0.1655	0.1733	0.1633	0.1752
	Female	0.1116	0.1117	0.0979	0.1058	0.1012	0.0949	0.0936	0.0948
	Male	0.2604	0.2753	0.1595	0.1585	0.1567	0.1637	0.1630	0.1648
2010–2020	Female	0.1102	0.1103	0.0938	0.1022	0.1091	0.0889	0.0966	0.0896
	Male	0.2826	0.2710	0.1679	0.1670	0.1643	0.1721	0.1706	0.1708
2010–2020	Female	0.1145	0.1120	0.1003	0.1007	0.1055	0.0958	0.0982	0.0962
	Mean	0.1936 (8)	0.1934 (7)	0.1313 (2)	0.1315 (4)	0.1337 (6)	0.1314 (3)	0.1309 (1)	0.1318 (5)

TABLE 9: RMSFE values of forecast errors during the period 2010–2020 for ages 20–60.

Fitting period	RMSFE for ages 20–60	Bühlmann methods	Regression	Methods	RC	Regression	Methods	FC	
	Gender	EW	MW	SEM	MEM	EEM	SEM	MEM	EEM
2010–2020	Male	0.5976	0.5938	0.3160	0.3170	0.3079	0.3307	0.3317	0.3359
	Female	0.1116	0.1117	0.0979	0.1058	0.1012	0.0949	0.0936	0.0948
	Male	0.2604	0.2753	0.1595	0.1585	0.1567	0.1637	0.1630	0.1648
2010–2020	Female	0.1102	0.1103	0.0938	0.1022	0.1091	0.0889	0.0966	0.0896
	Male	0.2826	0.2710	0.1679	0.1670	0.1643	0.1721	0.1706	0.1708
2010–2020	Female	0.1145	0.1120	0.1003	0.1007	0.1055	0.0958	0.0982	0.0962
	Mean	0.2936 (8)	0.3282 (7)	0.2357 (3)	0.2338 (4)	0.12434 (6)	0.2318 (1)	0.2323 (2)	0.2354 (5)

TABLE 10: MAPFE values of forecast errors during the period 2010–2020 for ages 20–60.

Fitting period	MAPFE for ages 20–60	Bühlmann methods	Regression	Methods	RC	Regression	Methods	FC	
	Gender	EW	MW	SEM	MEM	EEM	SEM	MEM	EEM
2010–2020	Male	0.2818	0.2801	0.1697	0.1544	0.1655	0.1733	0.1633	0.1752
	Female	0.1116	0.1117	0.0979	0.1058	0.1012	0.0949	0.0936	0.0948
	Male	0.2604	0.2753	0.1595	0.1585	0.1567	0.1637	0.1630	0.1648
2010–2020	Female	0.1102	0.1103	0.0938	0.1022	0.1091	0.0889	0.0966	0.0896
	Male	0.2826	0.2710	0.1679	0.1670	0.1643	0.1721	0.1706	0.1708
2010–2020	Female	0.1145	0.1120	0.1003	0.1007	0.1055	0.0958	0.0982	0.0962
	Mean	0.2879 (8)	0.3192 (7)	0.1313 (2)	0.1315 (4)	0.1337 (6)	0.1314 (3)	0.1309 (1)	0.1318 (5)

TABLE 11: MAFE values of forecast errors during the period 2010–2020 for ages 20–80.

Fitting period	MAFE for ages 20–80	Bühlmann methods	Regression	Methods	RC	Regression	Methods	FC	
	Gender	EW	MW	SEM	MEM	EEM	SEM	MEM	EEM
2010–2020	Male	0.4319	0.4204	0.3926	0.3614	0.3852	0.3914	0.3643	0.3906
	Female	0.2023	0.1948	0.2060	0.1960	0.2082	0.2053	0.1914	0.2051
	Male	0.3880	0.4116	0.3472	0.3550	0.3463	0.3535	0.3589	0.3524
2010–2020	Female	0.1777	0.1847	0.1841	0.1922	0.1940	0.1794	0.1888	0.1813
	Male	0.4494	0.4369	0.3550	0.3521	0.3524	0.3599	0.3545	0.3568
2010–2020	Female	0.2004	0.1975	0.1940	0.1939	0.1936	0.1921	0.1950	0.1938
	Mean	0.3083 (8)	0.3077 (7)	0.2798 (3)	0.2752 (1)	0.2800 (4)	0.2803 (6)	0.2755 (1)	0.2801 (5)

TABLE 12: RMSFE values of forecast errors during the period 2010–2020 for ages 60–100.

Fitting period	RMSFE for ages 60–100	Bühlmann methods	Regression	Methods	RC	Regression	Methods	FC	
	Gender	EW	MW	SEM	MEM	EEM	SEM	MEM	EEM
2010–2020	Male	0.6493	0.6313	0.5604	0.5056	0.5429	0.5640	0.5166	0.5615
	Female	0.2830	0.2738	0.2842	0.2690	0.2839	0.2867	0.2642	0.2857
	Male	0.5822	0.6191	0.4878	0.4966	0.4784	0.5065	0.5069	0.5022
2010–2020	Female	0.2544	0.2614	0.2581	0.2682	0.2725	0.2507	0.2630	0.2528
	Male	0.6763	0.6566	0.4967	0.4956	0.4886	0.5149	0.5021	0.5034
2010–2020	Female	0.2806	0.2768	0.2692	0.2695	0.2683	0.2660	0.2708	0.2678
	Mean	0.4543 (8)	0.4532 (7)	0.3928 (4)	0.3841 (1)	0.3892 (3)	0.3982 (6)	0.3873 (2)	0.3956 (5)

TABLE 13: MAPFE values of forecast errors during the period 2010–2020 for ages 60–100.

Fitting period	MAPFE for ages 60–100	Bühlmann methods	Regression	Methods	RC	Regression	Methods	FC	
	Gender	EW	MW	SEM	MEM	EEM	SEM	MEM	EEM
2010–2020	Male	11.44	11.21	11.38	11.00	11.45	11.15	10.76	11.17
	Female	11.92	11.66	11.98	11.77	12.43	11.58	11.32	11.63
	Male	10.58	11.04	10.54	10.76	10.79	10.33	10.62	10.39
2010–2020	Female	11.08	11.34	10.97	11.30	11.54	10.61	11.11	10.78
	Male	11.78	11.53	10.62	10.49	10.80	10.34	10.49	10.54
2010–2020	Female	11.86	11.77	11.39	11.20	11.35	11.18	11.42	11.38
	Mean	11.45 (8)	11.42 (7)	11.15 (5)	11.09 (4)	11.40 (6)	10.87 (1)	10.92 (2)	10.98 (3)

The systematic mortality risk from the Bühlmann credibility estimates becomes

$$\widehat{Q}_{t_{n-1}+1,y} = \widehat{K}\overline{Q_y} + (1 - \widehat{K})\overline{Q}, \qquad (21)$$

all for the values of $y = y_0, y_1, \ldots y_{k-1}$ and $\overline{Q_y}$ and $\overline{Q}$ are for the mortality risk of population estimates for finite and infinite data respectively.

Contrary to the credibility regression approaches that aims at capturing the central death rate, $m(t, x)$ downward trend as in s over t, application of the Bühlmann credibility approach eliminates it during modelling.

For the same reason, Tsai and Lin [19] did an application of the Bühlmann credibility model on a specified time series of mortality rate that changes as opposed to flat levels of the mortality rate, such that, $Q_{t,x} = lnm(t, x) - lnm(t - 1, x)$, for $t_1, t_2 \ldots, t_{n-1}$. Thereafter, they applied two strategies during the estimation of $Q_{t+h,x}, h = 2, 3, 4, \ldots, H$. The first strategy expands the fitting window (EW strategy) by a year, similar to the EEM regression method, as described in Section 4, and the second one moves the fitting window (MW) through a year, similar to the MEM regression method.

Similarly, we have done a comparison of the predicting performance in between the Bühlmann and the credibility regression approaches on Kenyan data [32]. To be consistent with the Bühlmann modelling structure [19]. Also, the age fitting spans between 20–100 and 60–100, which were chosen, and forecast errors were also evaluated under the mean MAPFE values denoted by MAPFE_{mean} as follows:

$$\text{MAPFE}_{\text{mean}} = \frac{1}{H(y_{k-1} - y_0 + 1)} \sum_{h=1}^{H} \sum_{y=y_0}^{y_{k-1}} \left| \frac{\widehat{m}(t_{n-1} + h, y) - m(t_{n-1} + h, y)}{m(t_{n-1} + h, y)} \right| * 100\%. \qquad (22)$$

It is important to note that the error values for every gender were evaluated by fitting $Q_{t,x}, s$ on period 2010–2020 and 2020–2050.

A comparison of mean values of MAFE, RMSFE, and MAPF results between Bühlmann and credibility regression methods is given for both genders in Table 4 for ages 20–60 and Table 5 for ages 50–90. The results indicate that credibility regression methods produce the smallest MAFE, RMSFE, and MAPFE values for the most selected fitting periods for both age spans. The FC-MEM method has the optimum expected performance per the MAFE and MAPFE values for ages 20–80. In contrast, the RC-MEM method seems more excellent in capturing future mortality trends for older ages 60–80.

We note that the smallest values in expected values are produced by different regression methods, which depend on the measures used. Such inconsistencies are always expected due to the nature of MAFE, RMSFE, and MAPFE formulas. The concept was discussed, as pointed out by Tsai and Yang [21].

Credibility regression techniques have proved vital for mortality datasets of a comparatively short historical time or period of recordings (for finite data). They have efficiently captured the underlying mortality trends for a specific age from all the information gathered from populations of many other periods.

This research study has proposed the mortality modelling approach that embeds it within the framework of credibility regression. Using Kenyan data for illustration, credibility regression approaches have led to better SMR forecasts for males and females. Comparing MAFE and RMSFE measures to the conventional LC and CBD models, they fit well during the credibility approach modelling.

Finally, we used a structure proposal of credibility regression mortality when dealing with randomly varying coefficients on a particular case with fixed coefficients. Thus, as explained in previous section five, we have presented extrapolation methods for each credibility approach when estimating the future mortality rates. From the analysis, we have concluded that, in aggregate, credibility modelling methods performed better than the LC and CBD mortality models. For forecasting accuracy during the study, we have seen a better analysis for the whole age fitting span, with fixed coefficients credibility methods that performed much better on average when compared to the RC-MEM, which makes a better analysis.

6. Conclusion and Practical Applications

The proposed credibility regression model have resulted in better SMR forecasts based on MAFE and RMSFE measures. Credibility regression thus offers excellent forecasting performance when applied to datasets of developing countries (having a relatively small population), such as Kenyan mortality data. As a proposal for further research, the forecasting comparison between many other countries' datasets would be a good avenue of explorative data analysis.

During the comparison, the Bühlmann credibility approach was applied to the Kenyan life table in Section 5. The credibility regression methods yielded excellent SMR forecasts based on MAFE, MAPFE, and RMSFE measures. Moreover, credibility regression methods had an excellent predicting performance.

We applied the proposed model to the datasets of other countries (with a relatively small population) for a few selected fitting periods for many countries with developed structures in mortality modelling. Additional prediction comparison between datasets of other countries has been left out for future research work. Our numerical illustration yielded relevant results and proved the credibility modelling approach, including those that could contribute to future systematic mortality risk projection studies.

On policy recommendation, the government and policymakers should model the systematic mortality risks correctly since it will always determine the amount of money in terms of benefits payable to the policyholders whenever they are in financial need from the life assurance and annuities bought. Ensuring that the affordable premiums are payable makes it an avenue for many policyholders who may now purchase these products sold by the pension, government, and insurance companies in developing countries.

The results of this study will help better systematic mortality risk modelling, which is important in actuarial modelling and pricing. Many pension and insurance firms will improve their estimates, which ultimately helps enhance the pricing rates of the actuarial rates sold in the Kenyan markets. By making these premiums affordable, policyholders can buy more products while reducing the burden of deaths that many Africans experience while living in many sub-Saharan countries such as Kenya.

Appendix

Tables 7–12 show the SMR values of MAFE, RMSFE, and MAPFE measures of forecast errors during the different periods for ages for both males and females when modelling the mortality rates. The authors note that incorporating the Bühlmann credibility approach helps enhance the accuracy and prediction of future mortality risks from the lower levels of errors than classical models.

Table 7 lists the MAFE values of forecast errors during the period 2010–2020 for ages 20–60 for the two models under the credibility approach.

Table 8 lists the RMSFE values of forecast errors during the period 2010–2020 between the ages 20 and 60 for both males and females.

Table 9 lists the MAPFE values of forecast errors during the period 2010–2020 for ages 20–60 for males and females. It shows how high levels of precision are achieved under the credibility approach.

Table 10 lists the MAFE values of forecast errors during the period 2010–2020 for ages 20–80 for males and females.

Table 11 lists the RMSFE values of forecast errors during the period 2010–2020 for ages 60–100 for both males and femalesTable 13.

Table 12 lists the MAPFE values of forecast errors during the period 2010–2020 for ages 60–100 for both males and females when modelling the systematic mortality rates.

Data Availability

The data used to support the findings of this study can be obtained from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] R. D. Lee and L. R. Carter, "Modeling and forecasting US mortality," *Journal of the American Statistical Association*, vol. 87, no. 419, pp. 659–671, 1992.

[2] K. Dowd, A. J. G. Cairns, D. Blake, G. D. Coughlan, and M. Khalaf-Allah, "A gravity model of mortality rates for two related populations," *North American Actuarial Journal*, vol. 15, no. 2, pp. 334–356, 2011.

[3] A. J. G. Cairns, D. Blake, D. Kevin et al., "Bayesian stochastic mortality modelling for two populations," *ASTIN Bulletin: The Journal of the IAA*, vol. 41, no. 1, pp. 29–59, 2011a.

[4] H. Booth and L. Tickle, "Mortality modelling and forecasting: a review of methods," *Annals of Actuarial Science*, vol. 3, no. 1-2, pp. 3–43, 2008.

[5] J. Tomas and F. Planchet, "Prospective mortality tables: taking heterogeneity into account," *Insurance: Mathematics and Economics*, vol. 63, pp. 169–190, 2015.

[6] H. Bühlmann, "Experience rating and credibility," *ASTIN Bulletin*, vol. 4, no. 3, pp. 199–207, 1967.

[7] Y. Salhi, P.-E. Thérond, and J. Tomas, "A credibility approach of the Makeham mortality law," *European Actuarial Journal*, vol. 6, no. 1, pp. 61–96, 2016.

[8] M. R. Hardy and H. H. Panjer, "A credibility approach to mortality risk," *ASTIN Bulletin*, vol. 28, no. 2, pp. 269–283, 1998.

[9] F. van Berkum, K. Antonio, and M. Vellekoop, "A Bayesian joint model for population and portfolio-specific mortality," *ASTIN Bulletin*, vol. 47, no. 3, pp. 681–713, 2017.

[10] C. C.-L. Tsai and L. Tzuling, "Incorporating the B'uhlmann credibility into mortality models to improve forecasting performances," *Scandinavian Actuarial Journal*, vol. 5, pp. 419–440, 2017b.

[11] A. Olivieri and E. Pitacco, "Life tables in actuarial models: from the deterministic setting to a Bayesian approach," *AStA Advances in Statistical Analysis*, vol. 96, no. 2, pp. 127–153, 2012.

[12] C. Pedroza, "A Bayesian forecasting model: predicting US male mortality," *Biostatistics*, vol. 7, no. 4, pp. 530–550, 2006.

[13] Algamal and Y. Zakariya, "Diagnostic in Poisson regression models," *Electronic Journal of Applied Statistical Analysis*, vol. 5, no. 2, pp. 178–186, 2012.

[14] Z. Y. Algamal and M. M. Alanaz, "Proposed methods in estimating the ridge regression parameter in Poisson regression model," *Electronic Journal of Applied Statistical Analysis*, vol. 11, no. 2, pp. 506–515, 2018.

[15] N. K. Rashad and Y. A. Zakariya, "A new ridge estimator for the Poisson regression model," *Iranian Journal of Science and Technology Transactions: Science*, vol. 43, no. 6, pp. 2921–2928, 2019.

[16] A. Alkhateeb and Z. Algamal, "Jackknifed Liu-type estimator in Poisson regression model," *Journal of the Iranian Statistical Society*, vol. 19, no. 1, pp. 21–37, 2020.

[17] K. Antonio, A. Bardoutsos, and W. Ouburg, "Bayesian Poisson log-bilinear models for mortality projections with multiple populations," *European Actuarial Journal*, vol. 5, no. 2, pp. 245–281, 2015.

[18] X. Luan, "A pseudo non-parametric buhlmann credibility approach to modeling mortality rates," 2015.

[19] 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, 2017a.

[20] A. J. G. Cairns, B. David, and D. Kevin, "A two-factor model for stochastic mortality with parameter uncertainty: theory and calibration," *Journal of Risk & Insurance*, vol. 73, no. 4, pp. 687–718, 2006.

[21] C. C. L. Tsai, S. Yang, and S. Yang, "A linear regression approach to modeling mortality rates of different forms," *North American Actuarial Journal*, vol. 19, no. 1, pp. 1–23, 2015.

[22] H. L. Shang, H. Booth, and R. J. Hyndman, "Point and interval forecasts of mortality rates and life expectancy: a comparison of ten principal component methods," *Demographic Research*, vol. 25, pp. 173–214, 2011.

[23] J. Odhiambo, P. Ngare, and P. Weke, "B'uhlmann credibility approach to systematic mortality risk modeling for sub-Saharan Africa populations (Kenya)," *Research in Mathematics*, vol. 9, no. 1, Article ID 2023979, 2022.

[24] J. Odhiambo, P. Weke, and P. Ngare, "A Deep learning integrated cairns-blake-dowd (CBD) sytematic mortality risk model," *Journal of Risk and Financial Management*, vol. 14, no. 6, p. 259, 2021.

[25] A. J. Cairns, D. Blake, K. Dowd, G. D. Coughlan, D. Epstein, and M. Khalaf-Allah, "Mortality density forecasts: an analysis of six stochastic mortality models," *Insurance: Mathematics and Economics*, vol. 48, no. 3, pp. 355–367, 2011b.

[26] N. Brouhns, M. Denuit, and J. K. Vermunt, "A Poisson logbilinear regression approach to the construction of projected lifetables," *Insurance: Mathematics and Economics*, vol. 31, no. 3, pp. 373–393, 2002.

[27] C. Hildreth and J. P. Houck, "Some estimators for a linear model with random coefficients," *Journal of the American Statistical Association*, vol. 63, no. 322, pp. 584–595, 1968.

[28] H. Li and Y. Lu, "A Bayesian non-parametric model for small population mortality," *Scandinavian Actuarial Journal*, vol. 2018, pp. 605–628, 2018.

[29] H. Bühlmann and A. Gisler, *A course in credibility theory and its applications*, Springer Science & Business Media, Berlin, Germany, 2006.

[30] Y. Salhi and P.-E. Thérond, "Age-specific adjustment of graduated mortality," *Astin Bulletin*, vol. 48, no. 02, 2018.

[31] Team R Core, *R: A Language and Environment for Statistical Computing* R Foundation for Statistical Computing, Vienna, Austria, 2017.

[32] M. M. D. SchinzingerEdo and M. C. Christiansen, "A multivariate evolutionary credibility model for mortality improvement rates," *Insurance: Mathematics and Economics*, vol. 69, pp. 70–81, 2016a.