

COMPARISON OF STOCHASTIC MORTALITY MODELS ENHANCED BY MACHINE LEARNING

BY

SULYMAN IYANDA

A Thesis Presented to the
COLLEGE OF GRADUATE AND INTERDISCIPLINARY STUDIES
KING FAHD UNIVERSITY OF PETROLEUM & MINERALS
DHAHRAN, SAUDI ARABIA

In Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

In

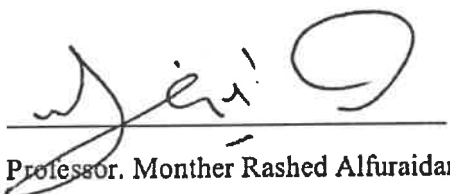
APPLIED STATISTICS

MAY 2024

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS
DHAHRAN- 31261, SAUDI ARABIA

DEANSHIP OF GRADUATE STUDIES

This thesis, written by **SULYMAN IYANDA** under the direction of his thesis advisor and approved by his thesis committee, has been presented and accepted by the Dean of Graduate Studies, in partial fulfillment of the requirements for the degree **MASTER OF SCIENCE IN APPLIED STATISTICS**.



Professor. Monther Rashed Alfuraidean
Department Chairman



Professor Suliman S. Alhomidan
Dean of Graduate Studies

Date



Dr. Mohammad Hafidz Omar
(Advisor)



Dr. Jimoh Olawale Ajadi
(Member)



Dr. Emmanuel Osita Afuecheta
(Member)



© SULTYMAN IYANDA

2024

DEDICATION

I dedicate this thesis to Almighty Allah (Subhanahu Wa Ta'ala سبحانه و تعالى) for His mercy and blessings in my life, as well as to my beloved father and my cherished late mother. I pray to Almighty Allah to grant her Aljanatul Firdaous (Aameen).

ACKNOWLEDGMENTS

Firstly, I praise, glorify, and appreciate Almighty Allah (S.W.A) for His mercy, blessings, full support, protection, and fortune in my life always. I deny none of Allah's grace; without His permission and help, this finding would not have been accomplished.

I am deeply indebted to my parents for their support - physically, financially, and spiritually - which has made this MS program a success. I pray to Almighty Allah, in His infinite mercy, to overlook any shortcomings of my late mother, whom I lost during this program. I also supplicate to Almighty Allah to grant her the highest position in paradise and to grant my father a long life filled with good health and prosperity.

My profound gratitude goes to my amiable thesis supervisor, **Dr. Mohammad Hafidz Omar**, for his unwavering supervision, contributions, and support - both academically and morally. May Almighty Allah continue to protect, guide, and increase him in beneficial knowledge. Additionally, I sincerely appreciate the contributions of Dr. Jimoh Ajadi and Dr. Afuecheta Osita Emmanuel to the successful completion of this thesis. I also extend my heartfelt thanks to faculty members Dr. Nasir Abbas and Prof. Muhammad Riaz for imparting quality knowledge to us. May Almighty Allah increase their intellect and bless them with long lives.

I must also express my gratitude to the chairman of mathematics, **Professor Monther Rashed Alfuraida**, the graduate coordinator, all staff, and the entire King Fahd University of Petroleum and Minerals (KFUPM) for the great scholarship opportunity they have provided me.

Lastly, I appreciate my three sisters for their prayers, love, and words of encouragement. I am also grateful to my Soulmate, all my friends, both near and far, as well as my family, mentors, guardians, and well-wishers, for their unwavering support, motivation, advice, and prayers. Thank you all very much. May Almighty Allah increase you all in goodness.

TABLE OF CONTENTS

ACKNOWLEDGMENTS.....	VII
TABLE OF CONTENTS	IX
LIST OF TABLES	XII
LIST OF FIGURES	XIII
LIST OF ABBREVIATIONS.....	XIV
ملخص الرسالة.....	3
CHAPTER 1	4
1.1 Introduction	4
1.2 Aims and Objectives	6
1.3 Mortality Models.....	6
1.3.1 Lee Carter (LC) Mode	7
1.3.2 Renshaw and Haberman (RH) Model	7
1.3.3 Age - Period - Cohort (APC) Model.....	8
1.3.4 Cairns – Blake – Dowd (CBD) Model.....	9
1.3.5 PLAT Model	10
1.4 Machine Learning (ML) Models	10
1.4.1 Decision Trees (DT)	11
1.4.2 Random Forest (RF)	13
1.4.3 Gradient Boosting (GB).....	14
1.5 Performance Criteria.....	15
1.5.1 likelihood function.....	15
1.5.2 Poisson Deviance	15
1.5.3 Akaike Information Criterion (AIC).....	16
1.5.4 Root mean square error (RMSE)	16
1.6 Organization of the findings	16
CHAPTER 2.....	18

LITERATURE REVIEW.....	18
2.1 Mortality Rates and its Models.....	18
2.1.1 The LC Model.....	20
2.1.2 Other Stochastic Mortality Models and their Variants	23
2.1.3 Comparison of Stochastic mortality models	28
2.2 Machine learning and mortality Models	32
2.3 Maternity	42
CHAPTER 3	49
COMPARATIVE ANALYSIS OF STOCHASTIC MORTALITY MODELS- ENHANCED BY MACHINE LEARNING.....	49
3.1 OVERVIEW	50
3.2 Methodology	52
3.2.1 Notation and Assumptions.....	52
3.2.2 Parametric Mortality Predictor Structures	54
3.2.3 Uniqueness Constraints	55
3.2.4 Description of datasets	56
3.2.5 Stochastic Mortality Models fitting.....	57
3.2.6 Diagnostic of the models	58
3.3 Machine Learning models.....	59
3.3.1 Model Notation and Assumptions.....	60
3.3.2 Decision Tree Machine Learning	62
3.3.3 Random Forest Machine Learning	63
3.3.4 Gradient Boosting Machine Learning.....	63
3.4 Results and Discussion	64
CHAPTER 4	70
ANALYZING MATERNAL MORTALITY WITH A ROBUST PARAMETRIC STOCHASTIC MORTALITY MODEL.	70
4.1 Introduction.....	70
4.2 Methodology	71
4.2.1 Maternal Mortality	71
4.2.2 Stochastic Mortality Models	72

4.2.3 Goodness of fit metric.....	72
4.3 Results and Discussion.....	73
CHAPTER 5.....	75
EXAMINING AGE-GENDER EFFECT ON LEE-CARTER MODEL USING GNM FRAMEWORK.....	75
5.1 Introduction	76
5.2 Literature Review	77
5.3 Methodology.....	82
5.3.1 Assumptions.....	82
5.3.2 Generalized non-linear models.	83
5.3.3 Goodness of fit	84
5.3.4 Original Lee-Carter Model.....	85
5.3.5 Extended Lee- Carter Model	87
5.4 Historical Dataset.....	88
5.5 Results and discussion	88
CHAPTER 6.....	94
CONCLUSION.....	94
REFERENCES	96
VITAE.....	108

LIST OF TABLES

Table 1 Stochastic Mortality Models and their uniqueness Constraints.	55
Table 2 Log-likelihood, Deviance, AIC, and RMSE statistics for stochastic mortality models. Ages 0 -100 and years 1950–2020, Switzerland.....	65
Table 3 Log-likelihood, Deviance, AIC, and RMSE statistics for stochastic mortality models. Ages 0 - 100 and years 1950–2020, Italy.	65
Table 4 Log-likelihood, Deviance, AIC, and RMSE statistics for stochastic mortality models. Ages 0 -100 and years 1950–2020, Belgium.....	65
Table 5 Log-likelihood, Deviance, AIC, and RMSE statistics for stochastic mortality models. Ages 0 -100 and years 1950–2020, Japan.	66
Table 6 Comparison of Stochastic mortality models with and without Machine Learning techniques, using Belgian mortality data.	68
Table 7 Comparison of Stochastic mortality models with and without Machine Learning techniques, using Belgian mortality data.	69
Table 8 Comparison of Stochastic mortality models with and without Machine Learning techniques, using Belgian mortality data.	69
Table 9 Belgian Maternal Mortality fits of year 1950.	73
Table 10 Description of variables/parameters.....	82
Table 11 Performance metrics for original Lee Carter fitted to each gender of Belgian historical dataset.....	89
Table 12 Performance metrics for extended Lee Carter fitted to combined genders of Belgian historical mortality dataset.....	92

LIST OF FIGURES

Figure 1 Decision Tree Structure (Image source: javatpoint)	12
Figure 2 Random Forest Structure (Image source: nvidia)	13
Figure 3 Pattern of logarithm of death rates according to age and time for Male(M) and Female(F) population of both Belgium and Switzerland.....	58
Figure 4 Pattern of logarithm of death rates according to age and time for Male(M) and Female(F) population of both Italy and Japan.	59
Figure 5 Ten parametric stochastic mortality models were fitted to the female gender using data from the datasets of four countries: Belgium, Switzerland, Italy, and Japan, all from the year 1950.	67
Figure 6 Ten parametric stochastic mortality models were fitted to the male gender using data from the datasets of four countries: Belgium, Switzerland, Italy, and Japan, all from the year 1950.....	68
Figure 7 The RHc, RHf, and FPLAT parametric stochastic models fit of Belgian historical maternal mortality rates for the year 1950.....	74
Figure 8 Original Lee Carter model parameter estimates plotted for male and female datasets in Belgian historical mortality data.	90
Figure 9 Parameter estimates for the Lee Carter model with age-gender effect, gender effect only, and no gender effect, applied to the combined male and female datasets from Belgian historical records.	91

LIST OF ABBREVIATIONS

Artificial neural networks (ANNs)	33
Chi-Squared Automatic Interaction Detection (CHAID)	10
Classification and Regression Tree (CART)	10
Classification and Regression Trees (CART)	35
flat cohort Renshaw-Haberman (RHf)	2
Generalized Linear Models (GLMs)	27
Generalized Nonlinear Models (GNMs)	27
Gradient Boosting (GB)	34
Hamiltonian Monte Carlo (HMC)	22
Lee-Carter (LC)	xv
non-flat cohort Renshaw-Haberman (RHc)	2
No-U-Turn sampler (NUTS)	22
Poisson Log-Normal Lee-Carter (PLNLC)	22
Quick, Unbiased, Efficient, Statistical Tree (QUEST)	10
Random Forests (RF)	33
Renshaw–Haberman (RH)	2
Value-at-Risk (VaR)	21

ABSTRACT

Full Name : Sulyman Iyanda

Thesis Title : Comparison Of Stochastic Mortality Models Enhanced By Machine Learning.

Major Field : Applied Statistics

Date of Degree: May, 2024

Stochastic mortality models are widely used in literature to examine death rates. Following the Lee-Carter (LC) model, several extensions have been developed, incorporating diverse features for accurate fits. These models analyze historical datasets to discern trends and predict future death rates. However, they often fit historical mortality datasets independently for each gender. Recently, Machine Learning techniques have been applied to improve model accuracy by identifying hidden dataset trends that mortality models have failed to capture. This study conducts a comparative analysis of ten common stochastic mortality models using historical datasets from four developed countries. Three outstanding models are identified and bolstered by Machine Learning techniques such as Decision Tree, Random Forest, and Gradient Boosting. Furthermore, a unique approach fits mortality models to a combined dataset of both genders, examining the influence of gender on death rates using generalized nonlinear model (gnm) framework of the Lee-Carter model that incorporates age-gender effects. The study evaluates model accuracy using seven goodness-of-fit metrics: Likelihood Function, Poisson Deviance, Akaike Information Criterion (AIC), Root Mean Square Errors (RMSE), Coefficient of Determination (R^2), Mean Absolute Error (MAE), and Mean Absolute

Percentage Error (MAPE). Applying Machine Learning to stochastic mortality models enhances fits, identifies model weaknesses and strengths, and aids in selecting appropriate stochastic models. The results of fitting mortality data for both genders combined demonstrate increased efficiency and parsimony.

ملخص الرسالة

الاسم الكامل: ادناي ناميلس

عنوان الرسالة: بي-لأ ملعتلا ةززعلا ةيناوشعلا تايفولا جدام ةنراقم

التخصص: قيبطنلا ءاصلا

تاريخ الدرجة العلمية: ن ورشعو ةعبرأو ن يفلأويام

عابتاب. تايفولا تلاءدم صحف تايبلا في عساو قباطن لء ةيناوشعلا تايفولا جدام مءخسء ءعونم تازيم نمضت يتلاو، تاءاتمالا نم ءياعلا ريوطت مت، (Lee-Carter (LC جءومء وبنلاو تاءاجتلا زيمتلا ءيخيراتلا تانايبلا تاءومجم ليحتب جداملا هءه موقت. ءقيء ءمءاملا ءيخيراتلا تايفولا تانايب تاءومجم بسانء ام ءبلا ءهنا، ءلء عمو. ءيلبقتسما تايفولا تلاءدمب نم جداملا ءقء ن يسحتل في لءا ملعتلا تاينقت قيبطنء مء، ءريءلا ءنولاً في ف. سنج لءل ءقسما لءشب هءه يرجء. ءهطاقتلا في تايفولا جدام تالشفي تلا ءيفءملا تانايبلا ءومجم تاءاجتا ءيءل لاء ءيخيراتلا تانايبلا تاءومجم مءءخسب ءيناوشعلا تايفولا ءعناش جدام ءرشعلا ءنراقم ءايءت ءسارءلا ءرجش لءم في لءا ملعتلا تاينقت عمو ءهءمءو ءزيمء جدام ءءلا ءيءل مء. ءمءقم ناءلب ءعبراً نم تايفولا جدام بسانب ءيرف ءهء لءانه، ءلء في لء ءولاء. ءرءلا زيزعتو ءيناوشعلا ءبلا ءارءلا راطا مءءخسب تايفولا تلاءدم في لء سنجلا ريثاء سرءيو، ن يسنجلا لاءل ءكرءشم تانايب ءومجم موقت. سنجلاو رءعلا تاريثاء نمضتي ءلا رءراكي لء جءومء (gnm) مءمءلا في طءلا ريء جءومءلا فارءناو، لءمءلا ءلاء: ءمءاملا ءءوء سبباقم ءعبس مءءخسب جءومءلا ءقء مبيءتب ءسارءلا لءامعو، (RMSE) رءءلا طسوتم عبرم ءاطءاو، (AIC) ءورءملا تاءولءملا رايءمو، ن وساءو يءوي. (MAPE) ءقءملا ءبسنلا أءخ يءعي و، (MAE) قءملا أءءلا طسوتمو، (R^2) ءيءل ءوقلاو ءعضلا طاقء ءيءل ءمءاملا ن يسحتل لء ءيناوشعلا تايفولا جدام في لءا ملعتلا قيبطنء تايفولا تانايب ءقباطم ءئاتء رهظتو. ءبسانملا ءيناوشعلا جداملا رايءءا في ءعاسيو، جءومءلا في لءبلاو ءءاقللا في ءءايز ءعمءم ن يسنجلا لاءل.

CHAPTER 1

1.1 Introduction

In the past several years, the reduction in mortality [1] due to infectious diseases, cardiovascular and circulatory disorders, and cancer has resulted in a sharp rise in life expectancy. Despite being a symbol of social progress, the increase in life expectancy presents a challenge faced by governments, private pension schemes, and life insurance companies due to its effect on health and pension expenditures [2]. Knowledge about future mortality rates is very important for both the government and private sectors, like life insurance businesses. The mortality rate is widely applicable in the fields of actuarial science, economics, and medicine. It is used to predetermine the country's economy, to obtain survival and death probabilities, and to project population of a country [3]. Most developed nations have seen sharp declines in mortality rates across the board. In contrast to 40 to 45 years in the middle of the eighteenth century and 60 to 65 years in the middle of the twentieth, the average life expectancy at the start of the twenty-first century is approximately 70 years old [4]. According to studies, the principal causes of death are cardiovascular diseases and cancer [5][6]. The COVID -19 pandemic was described as one of most prevalent cause of death after heart disease and cancer, as it caused most of deaths in 2020 [7], [8].

Several parametric death models have been proposed to fit the death rates. The LC is the pioneer and the most common stochastic mortality models [9][10]. Later,

Renshaw–Haberman (RH), and CBD were introduced as extensions of LC model. However, there have been variants and combinations of these basic models. The RH model has two versions, which are the flat cohort Renshaw-Haberman (RHf) and the non-flat cohort Renshaw-Haberman (RHc). Also, a model based on age, period, and cohort (APC) without modulating parameters and both the full and reduced Poisson latent models have been employed in the literature. CBD and its three variants are constrained to older ages only, specifically pensioner ages. Nonetheless, most of the models faced the problem of unique solution of the parameter estimation. For that reason, they required constraints to address the issue. While certain limitations are more logical than others, the selection of constraints is rather arbitrary [11]. Mathematical equations of these models are highlighted in subsection 1.3.

Recently, there has been literature that employed machine learning to study mortality rate unlike the previous years when there was limited application of the machine learning techniques in demography. Today, machine learning is widely used in actuarial science as well as computer science and other fields of study. Computer learning is the use of artificial intelligence through several data-optimized algorithms samples or prior knowledge. Machine learning is particularly helpful in identifying underlying correlations and patterns present in data sets. - those that are unknown and unidentifiable by the models, although it may not be able to explain everything.

This enhanced our understanding of processes and enables us to automatically classify data sets and forecast future events based on existing data.

1.2 Aims and Objectives

This study intends to perform a comparative analysis of parametric stochastic mortality models enhanced by machine learning approaches to augment the performance of the selected death models.

The research objectives are as follows:

- i. To fit stochastic mortality models for various countries, while considering the full range of age groups, choose the model that performs best based on four performance criteria.
- ii. To employ machine learning techniques to boost the efficiency of the first four best models.
- iii. To utilize the best stochastic mortality model identified in objective a) for maternal mortality datasets.
- iv. To develop a stochastic mortality model that includes gender as a factor.

1.3 Mortality Models

Stochastic mortality models are often used to fit either the initial death rate or central death rate. When calculating the mortality rates, one must consider a count variable in the numerator, which is often represented as a Poisson or binomial variable. Here we considered the central mortality rate denoted by $\varphi_{a,p}$ for the age a , individual calendar year or period p , and year-of-birth or cohort $c = p - a$. It is calculated by:

$$\varphi_{a,p} = \frac{\pi_{a,p}}{\psi_{a,p}} \quad (1.1)$$

Where $\pi_{a,p}$ is the realization of death count of people of age group a , and period p and $\psi_{a,p}$ denotes central exposure of age group a , at period p .

1.3.1 Lee Carter (LC) Mode

The LC model was proposed by [1] to extrapolate mortality trends and age patterns. It is the most widely used and original mortality model that gives an understandable and empirical explanation of the change in death [2]. It is defined with two age parameters $\kappa_a^{(1)}$ and $\kappa_a^{(2)}$, a random period parameter ρ_p and a random error term $\varepsilon_{a,p}$.

$$\ln \varphi_{a,p} = \kappa_a^{(1)} + \kappa_a^{(2)} \rho_p + \varepsilon_{a,p}. \quad (1.2)$$

For the model identifiability, the following constraints are necessary [3].

$$\sum_a \kappa_a^{(2)} = 1 \text{ and } \sum_p \rho_p = 0.$$

The fact that the LC model is one of the parsimonious parametric stochastic models does not give the impression of being the most suitable for mortality data. Versions of the LC model, that outperformed it, have been developed in previous studies; for example, the Renshaw and Haberman stochastic model incorporates the effect of cohort, which is the effect of individual cohort. However, some of these expansions are excessive in terms of the quantity of parameters. The Lee Carter model provides a poor fit when it is used to fit historical data of a country where a cohort effect is present. Throughout this work, the Lee and Carter model will be the LC model.

1.3.2 Renshaw and Haberman (RH) Model

The RH model was developed to address the problem of cohort effect facing the LC model. [4] provided the first proposition that took the cohort effect into account. The model offers a noticeably better fit to the historical data for countries where a cohort effect has previously been noted. However, it is not very robust. Furthermore, the correlation structure is negligible for the majority of the simulated

death rates even with the addition of a stochastic factor in the model for the cohort effect. The simulated cohort parameters mainly pertain to the elderly ages towards the conclusion of the projection, particularly when a wide age span is employed. The two versions of the Renshaw and Haberman models will be called RHc model for the non-flat cohort model and the RHf model for the flat cohort model in this study. They are defined, respectively, as follows:

$$\ln \varphi_{a,p} = \kappa_a^{(1)} + \kappa_a^{(2)} \rho_p + \kappa_a^{(3)} \delta_c + \varepsilon_{a,p}. \quad (1.2a)$$

In the second version, $\kappa_a^{(3)} = 1$.

$$\ln \varphi_{a,p} = \kappa_a^{(1)} + \kappa_a^{(2)} \rho_p + \delta_c + \varepsilon_{a,p}. \quad (1.3b)$$

Where δ_c represents random cohort effect with $c = p-a$. For model's identifiability, the following constraints are required:

$$\sum_a \kappa_a^{(2)} = 1, \sum_a \kappa_a^{(3)} = 1, \sum_p \rho_p = 0, \text{ and } \sum_c \delta_c = 0.$$

1.3.3 Age - Period - Cohort (APC) Model

A simplified version of RH called age-period-cohort was introduced by [5] afterward. It solves the problem of robustness of RH model and defined as:

$$\ln \varphi_{a,p} = \kappa_a + \rho_p + \delta_c + \varepsilon_{a,p}, \quad (1.4)$$

with identifiability constraints:

$$\sum_p \rho_p = 0, \sum_y c \delta_c = 0, \text{ and } \sum_c \delta_c = 0.$$

The APC model is not commonly applied in actuarial mortality literature, despite having along history in the sociology and medical domains [6].

1.3.4 Cairns – Blake – Dowd (CBD) Model

There are two factors used by the CBD model to calculate mortality rates: the one that exhibits a similar cause on mortality rates at all ages and the one that only has a different effect on the dynamics of death rates at older ages. These two factors are modelled as a two-dimensional random walk with drift [7]. The CBD model is initially distinct from the LC and RH models because it employs a different methodology that assumes a direct correlation between the logit of death rates and age for every period. The basic CBD model has neither a cohort effect nor a static age function. Also, it is based on no parameter constraint, as it does not face any problem of identifiability. The use of a more accurate CBD model for the elderly or those with ages of over 50 years is affected by this linearity assumption [8]. The model of CBD is defined as:

$$\ln \varphi_{a,p} = \rho_p^{(1)} + \rho_p^{(2)}(\bar{a} - a) + \varepsilon_{a,p}. \quad (1.5a)$$

Where $\bar{a}$ represents an average of the ages. The three variants of CBD model are as follows:

$$\ln \varphi_{a,p} = \rho_p^{(1)} + \rho_p^{(2)}(\bar{a} - a) + \delta_c + \varepsilon_{a,p}. \quad (1.5b)$$

It uses two of the constraints of APC model, these are:

$$\sum_c \delta_c = 0 \text{ and } \sum_c c \delta_c = 0.$$

$$\ln \varphi_{a,p} = \rho_p^{(1)} + \rho_p^{(2)}(\bar{a} - a) + \rho_p^{(3)}[(\bar{a} - a)^2 - \hat{\sigma}^2] + \delta_c + \varepsilon_{a,p}, \quad (1.5c)$$

where $\hat{\sigma}^2$ is the variance of the age considered and the model constraints on:

$$\sum_c \delta_c = 0, \sum_c c \delta_c = 0, \text{ and } \sum_{c^2} c \delta_c = 0.$$

$$\ln \varphi_{a,p} = \rho_p^{(1)} + \rho_p^{(1)}(\bar{a} - a) + (\omega - a)\delta_c + \varepsilon_{a,p} \quad (1.5d)$$

where ω is a constant parameter. The model uses the constraint,

$$\sum_c \delta_c = 0,$$

to avoid identifiability problems.

1.3.5 PLAT Model

The plat model was proposed by [9] in 2009. The model integrated key features from both the APC and CBD models to create a comprehensive approach. It possesses four stochastic factors, viz: $\rho_p^{(1)}$, $\rho_p^{(2)}$, $\rho_p^{(3)}$, and δ_c . It is defined by.

$$\ln \varphi_{a,p} = \kappa_a + \rho_p^{(1)} + \rho_p^{(2)}(\bar{a} - a) + \rho_p^{(3)}a^* + \delta_c + \varepsilon_{a,p}, \quad (1.6a)$$

where a^* is the maximum value between $(\bar{a} - a)$ and zero, i.e., $a^* = \max[(\bar{a} - a), 0]$. The following constraints are considered for the model, to avoid the problem of identifiability.

$$\sum_c \delta_c = 0, \sum_c c\delta_c = 0, \text{ and } \sum_p \rho_p^{(3)} = 0.$$

The reduced PLAT model also possesses all the advantageous features, and it can be obtained by removing the fourth term of the full model. It is only applicable to higher ages [9].

$$\ln \varphi_{a,p} = \kappa_a + \rho_p^{(1)} + \rho_p^{(2)}(\bar{a} - a) + \delta_c + \varepsilon_{a,p}, \quad (1.6b)$$

1.4 Machine Learning (ML) Models

Machine Learning comprises a collection of algorithms that enable computers to process and categorize new data considering previously provided information (training data). Computers performed the actions on their own, using mathematical models that are created by machine learning algorithms, without

explicit programming. ML can be classified into unsupervised and supervised learning [10]. In supervised learning methods, e.g., naive Bayes, logistic and multiple regression, classification and regression trees, and others, the algorithm learns the rules from the information that is provided to the machine. Contrarily, unsupervised learning algorithms, such as association rules and clustering, do not consider an output measure; instead, they establish patterns among diverse variables to categorize records showing resemblances. [11]. ML algorithms offer the potential to enhance the accuracy of mortality models by uncovering hidden trends and factors that traditional mortality models may overlook [19]. In our study, we have utilized three commonly employed Machine learning methods like Decision Tree (DT), Random Forest (RF), and Gradient Boosting (GB) to assess the strengths and weaknesses of the models, thereby improving their effectiveness.

1.4.1 Decision Trees (DT)

Decision Trees machine learning have similar features as a normal tree[12]. They are well known supervised classification[13]. Because of their high prediction performance and ease of interpretation, decision trees are frequently used in machine learning [14][15]. They are trees with features displayed at each node, decisions displayed at each branch, and outcomes displayed at each leaf.

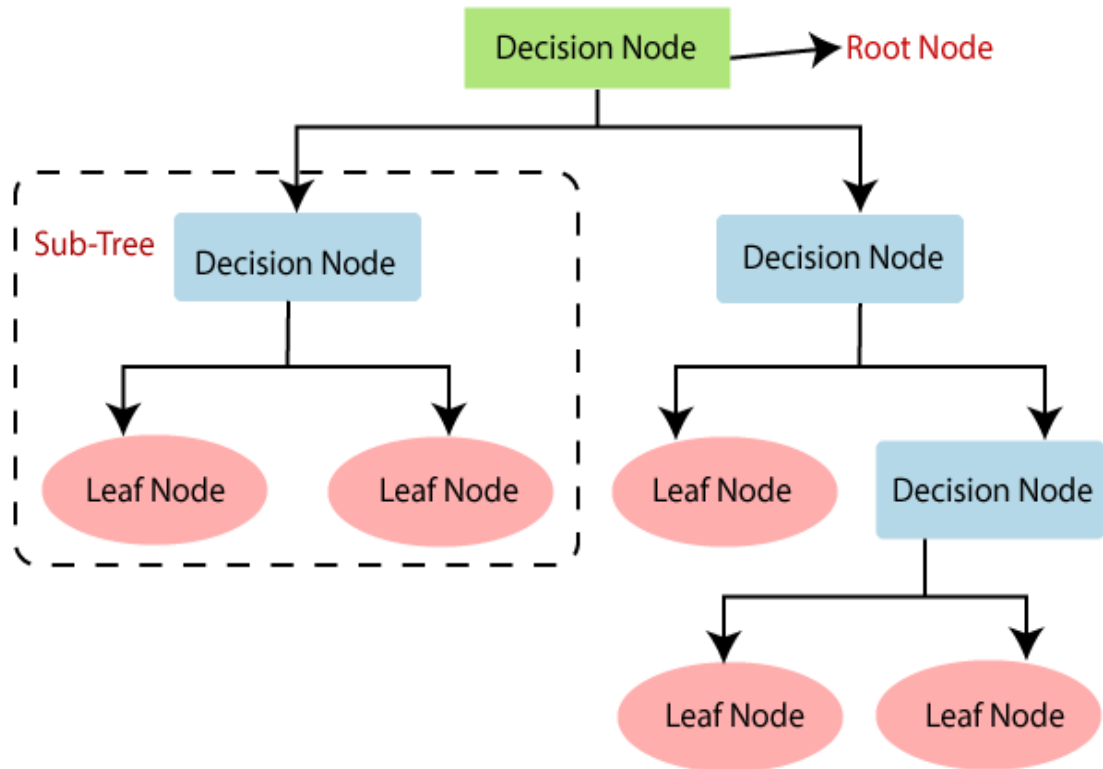


Figure 1 Decision Tree Structure (Image source: javatpoint)

Decision trees can be applicable to both Numerical (discrete and continuous) and categorical data sets [16]. They rely on dividing the feature space X using a sequence of binary splits, with each tree serving as a summary of the splitting rules applied to segment the predictor space. Once the entire feature space is recursively divided into a defined number of basic regions, one can predict the response for a particular observation by averaging the training observations within that observation's region. Decision Tree has various algorithms which include Classification and regression tree (CART), Quick, Unbiased, Efficient, Statistical Tree (QUEST), Chi-Squared Automatic Interaction Detection (CHAID) etc.

1.4.2 Random Forest (RF)

Random forests are also referred to as ensembles of decision trees . To improve prediction accuracy, it combines many decision trees using an ensemble learning approach [17]. They belong to the realm of supervised machine learning and are utilized for various prediction tasks, requiring minimal parameter adjustments. Reducing bias and preventing overfitting are two of RF's main advantages[12]. Apart from its user-friendly nature, the technique is renowned for its accuracy and capability to handle high-dimensional feature spaces and small sample sizes effectively. It can potentially handle big real-world systems and is readily parallelizable [18].

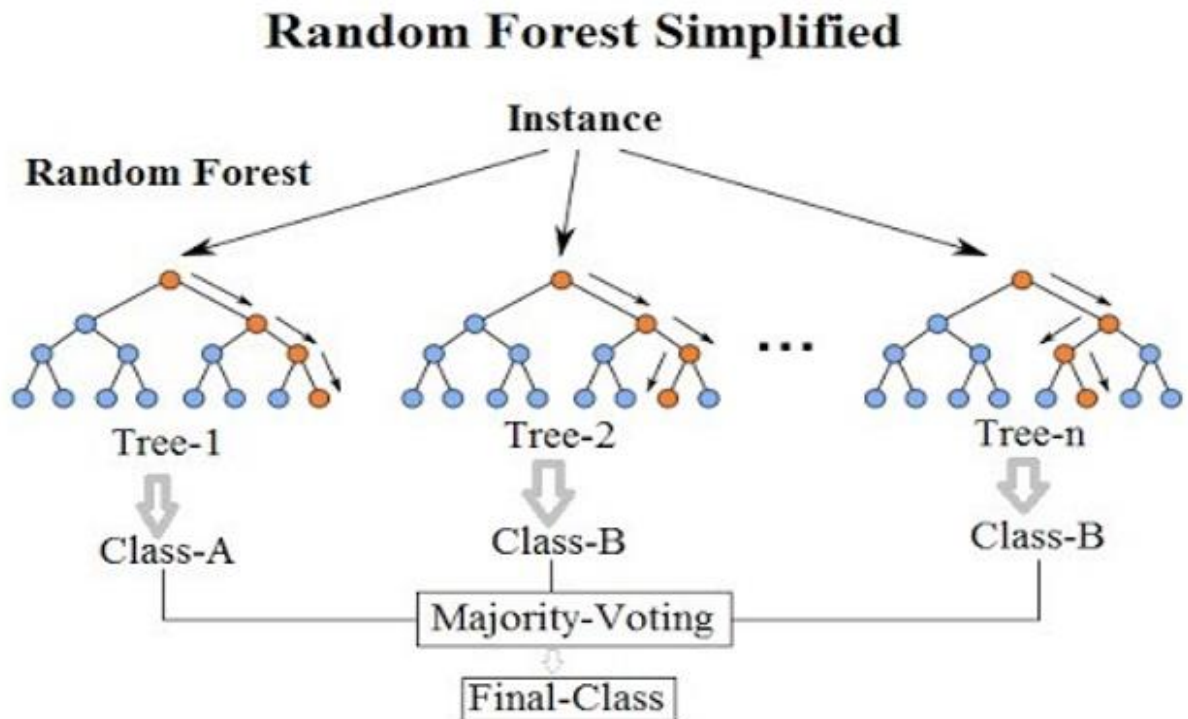


Figure 2 Random Forest Structure (Image source: nvidia)

During the training phase, a supervised classification algorithm known as random forest generates multiple decision trees and then during the testing phase, it outputs the class that garners the highest number of votes among these trees [19]. Each decision tree within Random Forest produces a class prediction, and the prediction chosen most frequently among all trees becomes the model's final prediction. Randomization is used by the random forest to generate many decision trees. For classification problems, voting is used to combine the output of these trees, and for regression problems, averaging is used to combine the output [20].

1.4.3 Gradient Boosting (GB)

A ML approach called gradient boosting has been investigated in several settings. It has been applied to classification tasks, comparing various implementations on real-world datasets, including XGBoost, LightGBM, and CatBoost [21]. It uses a scalable gradient boosting approach and is based on ensemble trees[22]. Through additive training methods, the predictions made by a group of "weak" learners are merged to create a robust "strong" learner, a technique referred to as boosting [23]. A regularization term plus a loss function make up the cost objective function that XGBoost is intended to maximize. Also, it is considered among the top machine learning frameworks for handling tasks involving both regression and classification. To provide a more precise estimate of the response variable, the learning processes of gradient-boosting machines are constantly fitting new models. The main concept underlying this algorithm is to build new base-learners that have the

highest possible correlation with the ensemble's negative loss function gradient [24]. Finding a gradient boosting variant that strikes the ideal mix between efficiency, dependability, and usability has been the aim [21].

1.5 Performance Criteria

To calibrate the competence of the models, we employed four different criteria. This is to ensure the efficiency of the best model based on these measures and not just one.

1.5.1 likelihood function

The likelihood function, assumed to follow a Poisson distribution, is considered as the first criterion employed in this study. The likelihood function for the Poisson distribution is defined as:

$$l(\varphi_{a,p}, \hat{\varphi}_{a,p}) = \sum_{a,p} \beta_{a,p} \{ \varphi_{a,p} \log \hat{\varphi}_{a,p} - \hat{\varphi}_{a,p} - \log \varphi_{a,p} \}, \quad (1.7)$$

Where $\beta_{a,p}$ is the binary weight for data cell. It takes 1 if a data cell is available and takes 0 if otherwise.

1.5.2 Poisson Deviance

Deviance is the second measure of model performance considered in this study. It presents the difference between the twice of Poisson log-likelihood of the estimated values of $\varphi_{a,p}$ and the observed values. It is calculated by the following formula:

$$\Delta = 2l(\varphi_{a,p}) - 2l(\hat{\varphi}_{a,p}), \quad (1.8)$$

where $l(.)$ represents the log likelihood function.

1.5.3 Akaike Information Criterion (AIC)

Akaike Information Criterion (AIC) is considered next, for measuring the accuracy of the model. It is often expected that the more the parameters of the models, the better the fit when assessing the accuracy of the models. Therefore, it is essential to penalize the higher parameterized models by adding the penalty function to the maximum likelihood criteria to justify the model(s) with the best fit. Several kinds of criteria that handled such situations have been discussed in the literature, one of them is AIC. Its equation is highlighted as follow:

$$AIC = 2r - 2\ln(\hat{l}) \quad (1.9)$$

where r and $\hat{l}$ represent the number of estimated parameters and maximized value of likelihood function respectively.

1.5.4 Root mean square error (RMSE)

The root mean square error (RMSE) will also be employed. RMSE is a metric which indicates the average distance between the observed values and the fitted values of mortality rates. It is given as:

$$RMSE = \sqrt{\frac{\sum(\varphi_{a,p} - \hat{\varphi}_{a,p})^2}{n}}, \quad (1.10)$$

where n represents the dimension of the data.

1.6 Organization of the findings

The other chapters of this study are arranged chronologically in this ways: in chapter two, the literature review is discussed; chapter three entails the comparative analysis of parametric stochastic mortality models enhanced by machine learning; chapter

four discusses the fitting of maternal mortality model using the best stochastic mortality model and in chapter five, a proposed extension of Lee-Carter model, which incorporates gender as a factor, is discussed and finally, we summarize and conclude in chapter six.

CHAPTER 2

LITERATURE REVIEW

This chapter is structured into three parts. The first part of this chapter delves into mortality and its diverse models, while the second part provides an overview of machine learning models. Lastly, the third section provides a specific focus on maternity.

2.1 Mortality Rates and its Models

The growing recognition of mortality improvements in nations with reliable data points to an underlying stochastic process driving these changes. This recognition emphasizes the need for a comprehensive understanding of death rate trends over time to facilitate accurate forecasting and projection of these underlying changes into the future. Governments, professionals, insurance and actuarial experts, and individuals are primarily motivated by the need because they want to be able to finance their retirement [25]. There is requirement for an appropriate model of death rates that captures the dynamics of the data and forecasts them with enough accuracy to be helpful to estimate the costs of extending life. Over an extended period, numerous parametric mortality models have been proposed to analyze the changes in death rates. They are fitted using either the generalized (non)linear package [26] or StMoMo package [27] in R. Among the different methods used for fitting and predicting mortality rates, the LC model [1] is notably the most utilized [28][29].

Lee and Carter presented a novel technique for forecasting mortality levels and age patterns. Their forecasting method merges a time series approach with a straightforward approach to handle mortality age distribution. Specifically, the approach presents the logarithm of age-specific death rates over time as a combination of an age component independent of time and a bilinear term. This bilinear term results from multiplying a time-varying parameter, reflecting the overall level of mortality, by an age-specific component.

The age-specific component indicates the speed at which mortality increases or decreases at each age due to changes in overall mortality levels. In the study of [30], retirement expenditures are examined using Lee Carter's mortality forecasting methodology. The paper specifically demonstrates how life expectancy forecasts can be successfully performed using R, offering valuable actuarial applications for social security and annuity insurances. The study entails the use of life contingency and demographic packages, conducting mechanical analyses designed to serve pedagogic goals.

[31] examined the impact of COVID-19 by analyzing mortality statistics from mortality monitoring programs in several European nations, utilizing accessible all-cause mortality data. Their research employed two methodologies to quantify the increased mortality attributed to COVID-19: Remaining after Seasonal Time Series Decomposition (RSTS) and the deviation from the predicted value (DEV) during corresponding periods. The study involved a comparison of various policy responses to COVID-19 and the prediction of excess mortality across different age and gender groups. Moreover, the research revealed that the increase in mortality surpasses

deaths officially attributed to COVID-19 in each of the five nations. Their findings suggest that to varying degrees among nations, the impact of COVID-19 on death was greater than officially reported deaths.

2.1.1 The LC Model

LC model was specifically designed to fit and forecast age-specific death rates associated to individual age and period of U.S. In contrast to the approach used by [1], which employed singular value decomposition (SVD) technique to obtain the least square estimation, [32][33], and [34] applied the binomial, Poisson, and negative binomial frameworks of the LC model in their respective works. All these frameworks of the LC model calculate the parameters by using the maximum likelihood approach, which enables the assessment of parameter accuracy. The work of [2] elucidated the basic Lee-Carter methodology and its forecasting strengths. The work delved into the constraints of the technique alongside its recent expansions, implementations, and methodological advancements. Furthermore, it provided an overview of how the method is employed in broader stochastic population projections and stochastic financial predictions concerning the US Social Security system.

In a recent study, [35] present the methodology and provide an ex-post evaluation of the original LC projection for U.S. death, which was performed three decades ago. They discuss most well-known expansions of the LC method, addressing specific issues and providing an analysis of previous assessments and comparisons. Their focus is on evaluating the effectiveness of different expansions. [36]

investigate residuals and potential future trends by expanding and employing LC methods to death data from Spain.

A comparative analysis of simulation techniques for evaluating risk in mortality rate forecasts and related estimations of life expectancy and annuity values in period and cohort frameworks is included in [37]'s work. The original LC model was extended and employed in the work by [38] to forecast death rates by using a new class of time series models called Generalized Autoregressive Score (GAS) or Dynamic Conditional Score (DCS) models. Notably, the research suggests five probability models—Poisson, binomial, negative binomial, Gaussian, and beta—that are based on the GAS framework and are applied to estimate the Lee–Carter parameters in a single, integrated step to dynamically forecast the mortality rates. The study's conclusion is that the LC model's negative binomial extension is the most effective at predicting death rates among the probability models. An extensive analysis of the link between the pattern of economic expansion as measured by the GDP and the latent trend in mortality dynamics was demonstrated by [39] . The Lee-Carter framework is then expanded by adding GDP as a second component next to the latent factor, which improves forecast interpretability and fit. [40] highlight various hidden characteristics of the Lee-Carter mortality forecasting method, demonstrating that it is a special case of a simpler and less frequently biased random walk with drift model. They further establish that the age profile forecast from both approaches, whether predicting forward or backward in time, will eventually deviate from any given baseline and become less smooth and unrealistic after a certain point.

The problem of estimating mortality and life expectancy in the event of a structural change is taken into consideration in the research work by [3]. The study shows how, within the framework of the LC model, new improvements in statistical testing for structural changes may be used to derive adequately described models for the overall mortality rate. Specifically, the results of tests for a shift in the mortality index's trend and the existence of a unit root are utilized to determine the most suitable forecasting models. Additionally, almost all the countries examined had structurally different rates of fall in overall mortality, particularly in male populations, according to the study.

The uncertainty surrounding future mortality trends significantly impacts an insurance company's risk profile in the annuity industry. Accurately predicting future survival patterns is challenging, especially when considering retirement ages and the effects of random fluctuations, coupled with the rectangularization phenomena. The correlation among ages is a crucial aspect in mortality modeling, [41], [42]. Another significant factor influencing forecasts is the "cohort-period effect". Specifically, the mortality experience of industrialized nations over the 20th century implies a significant age–time interaction, with two primary trends impacting various age groups at various times. This suggests that investigating the dependence structure, that is, the inter-period correlation across time, is worthwhile. [42] adjusted the mortality forecasting process to incorporate the dependence structure of nearby data in the population, improving model performance and forecasting capacity. Their expanded methodology was achieved by modifying the Lee–Carter context's mechanism for assessing projection uncertainty and by

substituting a custom bootstrap for a standard one. [43] provide a thorough examination of the uniqueness problems in age/time death models to comprehend and subsequently address them. The study goes on to explain how these identification problems occur, how to select identification schemes that support the explanation of the models' demographics, and how to predict the models so that future predictions are independent of the arbitrary selections made when identifying the historical parameters derived from historical data.

There is a consensus that incorporating trends from a larger population could enhance mortality projections for countries. [44] expanded the bilinear LC model to incorporate a three-way framework, introducing an extra component in decomposing the log-death rates. This three-way LC model was then applied to multiple nations. Their analysis, performed prior to extracting the rank-one components using the LC model, aids in identifying the sources of variance in the raw death data. Additionally, the method is utilized to derive a customized time death trend parameter for subgroups categorized by age. According to their findings, the suggested approach produces a better logical explanation of death for age-specific subgroups.

2.1.2 Other Stochastic Mortality Models and their Variants

There is a sharp rise in the concern of cohort effects in death dataset when there are more research concentrating on expanding the LC model to include year-of-birth parameters. Nevertheless, certain research indicates that these models lack resilience to modifications in the data or fitting technique, hence restricting their use for several uses. There has been speculation that an unresolved uniqueness problem could be

the cause of these robustness issues. The study in [45] thoroughly examined the reliability of cohort expansions in the LC model and the combination of algorithms used to fit it. Their research established the existence of this kind of identifiability challenge and suggests an extra approximation identifiability constraint that addresses several of the issues raised.

A completely Bayesian implementation of the APC-improvement (APCI) model with over dispersion is presented by [46] and is contrasted with the cohort-based Lee–Carter model. This implementation is then compared with the LC-cohort model. The study reveals that naive prior specification can lead to misleading inferences, prompting the authors to propose Laplace prior as an elegant solution. Additionally, they employ model averaging to account for model uncertainty. According to their research, the APCI model provides a better forecast and fit for data covering England and Wales from 1961 to 2002. The study by [47] presented a comprehensive implementation of the APCI model as a fully stochastic model. Their study sheds light on various intriguing aspects of the APCI model, containing the parameters suitable for smoothing and the extent to which the model surpasses other related models in terms of fitting data. While the improved fit is notable, it raises concerns about increased capital requirements for insurers in a value-at-risk (VaR) framework. The paper delves into an exploration of this phenomenon by examining the density of VaR simulations. The research described in [48] tackles the complexity of mortality modeling by integrating cohort factors into a state-space methodology through an innovative approach. The study shows that cohort factors can be seamlessly incorporated within the state-space framework, despite being indexed by the year of

birth rather than the calendar year. Bayesian inference for cohort models within a state-space framework is then established using an efficient Markov chain Monte Carlo sampler. This allows for the measurement of parameter uncertainty in cohort models and the subsequent mortality predictions, which are used to build life tables and expectancy calculations. The study's findings reveal the presence of cohort patterns in certain countries considered, emphasizing the significance of incorporating cohort factors to capture these phenomena. This highlights the benefits of establishing cohort models within the state-space.

The study by [49] introduces a parallel methodology based on generalized linear modeling, offering a reinterpretation of the model that form the basis of the LC method for death rates forecasting. Residual plots were used in the study to assess the efficiency and different methods for life expectancy forecasts are compared. The study concludes by comparing several approaches to life expectancy prediction. Instead of relying solely on the first set of single value decomposition vectors, like well-known LC (Gaussian) approach to death projection, [50] explored the possibility of developing mortality forecasts using the first two sets of vectors. The study presents and compares forecasts from three applications with those made using two related methods based on generalized linear and bilinear models with Poisson error structures. [51] provide a cohort effect and an overdispersion component in their adaptation of the Poisson LC model for death rate prediction. The research calculated the estimation of model parameters and predictions of mortality rates utilizing Bayesian analysis and the No-U-Turn sampler (NUTS), which is an extension of Hamiltonian Monte Carlo (HMC). Specifically, they use real mortality

data from multiple European nations to compare the prediction accuracy of the proposed model and the Poisson Log-Normal-LC (PLNLC) model and then reached the conclusion that the proposed model outperformed the PLNLC model.

The Lee-Carter modeling framework is modified by (Renshaw & Haberman, 2006) through the incorporation of cohort effects. This adjustment is accomplished by introducing a larger class of generalized, parametric, non-linear models that enable the modeling and forecasting of age-specific cohort effects. Similarly, [53] presented an extension of LC modeling technique for mortality fitting, projecting, and examining projecting robustness. It has been broadened to encompass an age-varying cohort index alongside the traditional age-adjusted period index. For the male death experience in England and Wales, life expectancy and annuity value projections are compared using both the age-adjusted cohort index and without it, at pensioner ages and for different eras. A stochastic mortality model was proposed by [9]. The model incorporates features from existing models while addressing their limitations. Particularly, the developed model exhibits a perfect fit to historical data, is applicable across the entire age range, incorporates cohort effects, features a non-trivial (yet not overly complex) correlation structure, and exhibits no robustness issues. Despite these advantages, the model remains relatively simple in structure. In addition, the research outlines how to include parameter uncertainty into the model and provides a risk-neutral pricing version of the model. In the study by [54], a new method for modeling subnational mortality is proposed. This method uses individual factors such as sex, geographical locations, and socioeconomic status, instead of focusing directly on subpopulations. The approach is considered

parsimonious as it requires fewer parameters to explain mortality. The research presents a broad Bayesian estimation method that effectively works with many age–period–cohort mortality models suggested in existing literature. It utilizes Markov chain Monte Carlo methods by integrating deterministic filtering with adaptive Metropolis steps, making it suitable for complex cases. Ultimately, the study shows that this approach, when paired with an appropriate mortality model, produces sensible, accurate, and consistent mortality predictions that align well with the data. In a study conducted by [55], a model proposed based on the mortality surface's breakdown into a polynomial foundation. The research demonstrates the application of cross-validation and regularization techniques to produce a parsimonious and logical predictive model for death projection. Furthermore, the study assesses the potential impact of COVID-19-like effects on predictions made using both classical and employed approaches. Essentially, predictions of mortality rates are generally more stable than those based on a cohort effect, and regularized models exhibit better projection performance and stability than the so-called P-spline model. [56] proposes straightforward stochastic models that are well-suited for numerical analysis of cash flows related to mortality. The statistical research supports the established relationship between old age mortality and GDP, with implications for various financial market sectors. These suggested models offer a clear and quantitative representation of these linkages. Special attention is given to the long-term progression of death rates, a crucial consideration of life insurance field. [57] examine how the post-age-60 death curve has changed in the UK and how this transformation has influenced the cost of longevity risk—a risk associated with long-

term improvements in overall mortality. The procedure involves a two-factor stochastic model for the temporal development of the curve. Notably, the second component influences mortality-rate dynamics at higher ages significantly more than at lower ages, while the first factor affects death-rate dynamics uniformly across the entire ages. [58] evaluate whether estimates of various metrics capture variations between models.

The study generates nine sets of indicator forecasts by combining three models with three block-bootstrap samples, with each sample comprising 50 observations. Moreover, functional ANOVA was utilized to evaluate the estimated mortality indicators. Following that, the models and block-bootstrap techniques are employed to death data from Spain. The outcomes showcase the impacts of model choice, block-bootstrap technique, and their interaction on all mortality indicators.

2.1.3 Comparison of Stochastic mortality models

The work by [8] considered eight different stochastic mortality models to investigate the improvements in death rates and examined them quantitatively. They utilized England and Wales and US datasets for males aged sixty to eighty-nine, assessing models based on criteria such as consistency and robustness relative to historical data. For the two countries used, no single model outperformed the others; that is, no model could be taken out as being the best in all aspects. Yet, each model displayed specific strength. The family of generalized APC death models, encompassing most recent models of stochastic mortality projection, such as LC and CBD models, is defined within framework of generalized linear and nonlinear models (Turner et al., 2012). The work introduces a package of R, called StMoMo

[27] which unifies the architecture of the generalized APC family. The package offers methods for estimating and fitting stochastic mortality models, produces mortality projections, and metrics for examining model's accuracy. Lastly, A comparative study of the package implementation and various death models applied to historical datasets of England and Wales was performed.

In the same vein, [29] compared basic mortality models to adjusted mortality models, to see how perfect the adjusted versions handled the inadequacies observed in some basic models, focusing on pensioner ages and a wider range capturing working ages for both sexes. Some stochastic mortality models, including the APC model, were excluded for not fitting the age-period-cohort pattern correctly. Short-to-medium-term efficiency of five expansions of the LC model for death prediction is compared in [60]. The findings indicate that all variants and expansions outperform the original LC method when predicting the log of mortality rates. However, the study emphasizes that efficiency in log mortality rates does not necessarily translate into efficiency in life expectancy. Essentially, the methods do not significantly differ in forecasting accuracy for life expectancy.

Similarly, [61] discussed the implementation of the APC-Improvement (APCI) model, which shares attributes with the LC and APC models, for death projection. Their work assessed how well the APCI model performs compared to others with similar structural properties. However, the model demands higher capital under the Value at Risk (VaR) assessment of longevity trend risk. For further ideas on VaR assessment, the readers are recommended to refer to Richards et al. (2021).

In the realm of fitting and forecasting mortality, generalized linear models (GLMs) and generalized nonlinear models (GNMs), efficiently express stochastic mortality models. In a comprehensive study by [62], eight distinct mortality model structures were compared. The R language was employed to delineate these structures, providing algebraic descriptions with the advantage of clarity and simplicity. Among the eight models under consideration, six were fitted using the `glm` approach, while the remaining two models such as the Lee–Carter and Renshaw–Haberman, employed `gnm` due to the presence of bilinear (that is, the product of two parameters) terms. The `gnm` package [26] in R software permits the direct fitting of models with such specifications that is very similar to that of the `glm` function. The models employed Poisson structures with log and logit link functions for the force of mortality, as well as binomial structures with logit and complementary log-log link functions for the rate of mortality within each structure. Furthermore, the study delved into the feasibility of projecting using these models. Notably, the incorporation of the effects of year-of-birth was found to enhance overall model fit. However, it was acknowledged that incorporating cohort effects could pose challenges in forecasting with these models.

Moreso, the study conducted by [63], a back testing approach was employed to compare the LC and CBD mortality models. The research specifically focused on contrasting the effectiveness of medium-length and short-run forecasts using three distinct back testing methodologies. The findings revealed that neither model captured the improved shock on mortality in the male population over the examined time. Additionally, the study indicated that Lee-Carter model forecasts tend to be

more accurate for the data, while CBD projections are generally dependable for ages over 75. The performance of six stochastic mortality models is demonstrated by [64], using England & Wales male mortality data. The study leverages the framework of each model to derive multiple residual series projected to be independently and identically distributed (i.i.d.) with a standard normal distribution under the null hypothesis of model efficiency. Their approach assesses the model's fit quality using standard tests for evaluating i.i.d. standard normal predictions. The study by [25] evaluates several stochastic mortality models by assessing the fitting quality and examining the residuals to ensure they exhibit normality. Their comprehensive analysis encompasses 30 nations and reveals that nearly all the residuals lack normality. Furthermore, the study identifies indications of residual structure that the models are unable to adequately reflect.

[65] delved into an exploration of the diverse sources of uncertainty impacting the valuation of a distinctive financial instrument known as a q-forward contract, which is intricately linked to longevity risk. This specialized contract involves the exchange of a predetermined mortality rate, agreed upon at the contract's initiation, for the actual mortality rate observed in a specific population upon the contract's maturity. The establishment of this fixed rate is a critical step in the pricing of a q-forward contract. The research systematically dissects the three primary sources of uncertainty, rigorously investigating their individual impacts on the pricing dynamics. These sources encompass the pricing methodology itself, the time window employed for rate estimation, and the choice of the underlying mortality rate model. By scrutinizing each of these facets, the study provides valuable insights

into the nuanced factors influencing the pricing of q-forward contracts in the realm of longevity-based financial instruments.

2.2 Machine learning and mortality Models

Machine learning entails the process of identifying and classifying hidden patterns in data to extract valuable knowledge. There are six common task classes in data mining: identification of anomalies, learning association rules, clustering, classification, regression, and summarization. Classification, a method of machine learning, is employed to forecast group membership for data instances. [14] offers a comprehensive overview of machine learning approaches, supplemented by a detailed analysis of select algorithms. [66] provides an overview of fundamental classification methods, including Bayesian networks, k-nearest neighbor classifier, and decision tree induction. The study offer a comprehensive analysis of various data mining categorization strategies. Among various data mining techniques, Decision Tree stands out as one of the most widely used. Decision trees adopt the divide and conquer method as a foundational learning approach. Essentially, a decision tree is structured like a flow chart, where each inner node represents a "test" on a feature, each branch signifies the outcome of the test, and each leaf node denotes the class label. In their work, [13] delves into the features, benefits, and drawbacks of several decision tree algorithms, which includes ID3, C4.5, and CART. Similarly, various decision tree algorithms have been developed based on accuracy and effectiveness. It is crucial for us to understand how to employ the optimal algorithm in each case when making decisions. [16] examined three distinct Decision Tree algorithms, including ID3, C4.5, and CART.

In [67], advancements over previous techniques, such as the first regression tree algorithm, are surveyed, and the main concepts underlying several significant algorithms are briefly reviewed. Nowadays, classification trees can fit models such as kernel density, closest neighbor, and others in the partitions and split the dataset using linear splits on subsets of variables. Additionally, regression trees can fit almost all types of traditional statistical models, including logistic, Poisson, least-squares, quantile, logistic, and models for longitudinal and multiple response data. Through research and the review of recently published papers focusing on quantitative and qualitative research in real-time problem-solving, particularly predictive analytics in multidisciplinary fields, [10] clarify various new updates and conclusions related to the most common machine learning algorithms. The study thoroughly evaluates and investigates all relevant qualitative and quantitative features to understand and explain the algorithms' correctness, robustness, and dependability. Additionally, it discusses the impromptu origin of these algorithms. Furthermore, the report outlines possible future uses of artificial intelligence and machine learning algorithms, as well as their roles in automation and holistic development, encompassing both technological and humanitarian aspects.

In the study by [68], machine learning is employed, specifically regression tree boosting machine, to evaluate the deficiencies and accuracy of both LC and Renshaw-Haberman models. This effort resulted in improvements by considering individual age, period and cohort. Similarly, [69] developed neural networks that duplicate the structure of the individual LC models for fitting the original LC and the Poisson-LC models on multiple populations. The method allows joint fitting of

the historical mortality data from forty different countries simultaneously. [11] examined the effectiveness of machine learning techniques on various mortality models, emphasizing the significance of machine learning in optimizing both estimation and forecasting. They utilized decision trees, random forest, and gradient boosting. Empirical evidence has consistently shown that tree boosting is an exceptionally effective predictive modeling technique, yielding outstanding results across various domains. Traditionally, MART has been the favored method for boosting trees. However, in recent times, XGBoost, another tree boosting technique, has gained prominence by winning numerous machine learning competitions. [70] delves into the comparison between XGBoost and the more conventional MART. The study highlights that XGBoost utilizes the Newton boosting method, contrasting with MART's gradient boosting approach. Moreover, XGBoost incorporates several minor adjustments that enable it to carefully manage the bias-variance tradeoff. The paper also explores the regularization methods employed by both approaches and their respective impacts on the resulting models.

The work of [71] employed , machine learning to minimize values of the metrics considered in the work to justify the best of the mortality models. The study discussed that residuals contain information that mortality models may be unable to capture. However, their study was limited to a decision-tree optimization approach and considered a few basic death models. To determine the risk variables associated with death through the analysis of survival data, doctors, researchers, and biostatisticians have traditionally relied on certain regression approaches, most notably the Cox model. With the broader distribution of computing power, there has

been an increase in the utilization of techniques requiring more sophisticated mathematics. In the current era of machine learning and "big data," survival analysis has expanded methodologically. [20] investigates Random Forest, a regression tree method that achieves a high level of prediction accuracy by employing bootstrap aggregation and randomization of predictors. To compare and ascertain which model performs best on the same data, the study builds two models: a random forest model and a Cox model.

The application of gradient boosting in the context of classification is examined in [21]. The study compared Bayesian optimization with randomized search, specifically emphasizing hyperparameter optimization using the Tree-structured Parzen Estimator. In addition to runtime and tuning time, conventional classification accuracy criteria were used to evaluate their methods' performance. Additionally, proper statistical testing has been used to validate the acquired results. Furthermore, efforts were made to identify a gradient boosting version that balances dependability, efficiency, and usability. The advancements in the most current theoretical and methodological aspects of random forests are reviewed by [18]. The focus is on the mathematical principles that underpin the algorithm, specifically focusing on parameter selection, resampling, and variable significance measures. The primary principles are meant to be easily accessible to non-experts through this review. The study conducted by [15] focused on the key concepts, benefits, and drawbacks of four popular supervised machine learning techniques: support vector machine, artificial neural network, decision tree, and K-nearest-neighbor. The study used sensitivity and specificity measurements to assess the performance and carried

out practical application to compare the methodologies' performance. In summary, the study discovered that no single metric can fully capture the classifier's performance, nor is there a single classifier that can meet every need.

More adaptable tree-based techniques are presented by [72] as an alternative to generalized linear models (GLM), which are the industry standard for pricing non-life insurance products. GLM requires an in-depth prior understanding of the data's structure. The article first discusses the theoretical aspects of these approaches before applying them to address the issue of forecasting a policyholder's claim frequency. According to the study, the regression tree methodology is a helpful tool for creating homogeneous sub-portfolios with similar risk characteristics. Furthermore, more advanced tree ensemble approaches like bagging and random forest offer comparable alternatives to GLM models in terms of predictive power. [73] utilized historical data from Malaysia to develop several Lee-Carter (LC) based models for forecasting death rates. In addition to the LC-ARIMA model, the study employed a hybrid LC model that integrated machine learning models from artificial neural networks (ANNs) and random forests (RF). Subsequently, death rates were predicted using the model that produce best representation of the data. Remarkably, the study identified the LC-ANN and LC-ARIMA models as the most accurate for predicting male and female death rates in Malaysia, respectively. Furthermore, their research indicated that the LC-ANN model outperformed in predicting death rates in nations with less accessible and efficient healthcare systems, while the LC-ARIMA model excelled in nations with longer life expectancies.

A powerful family of machine-learning algorithms, known as gradient boosting machines, has demonstrated notable effectiveness in a variety of real-world applications. These algorithms can be easily tailored to meet specific application requirements, such as learning various loss functions. [74] provided a tutorial on gradient boosting approaches with a particular emphasis on the machine learning components of modeling. The study first addresses model complexity management and then provides insights into three real-world applications of gradient boosting. Data scientists extensively employ XGBoost, a scalable end-to-end tree boosting technique, to attain cutting-edge outcomes in a range of machine learning challenges [75]. The work proposes a weighted quantile sketch for approximate tree learning and a unique sparsity-aware algorithm tailored for sparse data. Essentially, it leverages knowledge of data compression, shaping, and cache access patterns to develop a scalable tree boosting system. For both regression and classification tasks, gradient boosting (GB) constructs a projection model as an ensemble of weak projection models, typically decision trees. Among the various GB techniques, XGBoost (eXtreme Gradient Boosting) stands out as a widely used framework available in multiple programming languages, including C++, Python, and Java. It aims to provide a "Scalable, Portable and Distributed Gradient Boosting Library." [76] introduce gradient boosting techniques and discuss a theoretical subset of the XGBoost framework implemented within the JADBIO framework, drawing comparisons between the two implementations. The gradient boosting approach within JADBIO incorporates a practical subset of XGBoost's features.

To enhance the scalability of tree boosting techniques, [77] and [75] introduce XGBoost, a reliable distributed machine learning system. Designed to withstand faults in distributed environments, the system is optimized for rapid parallel tree generation. XGBoost can process tens of millions of samples on a single node, and, with distributed computing, it can handle billions of samples. A practical analysis of the operation of three innovative gradient boosting variants—XGBoost, LightGBM, and CatBoost—in terms of training speed, generalization performance, and hyper-parameter configuration is presented by [78]. Notably, the study conducts a comprehensive comparison of the three variations—random forests, gradient boosting, and both—using models that have been meticulously adjusted beyond their initial settings. According to the findings, while LightGBM is the fastest approach, it may not be the most accurate, whereas CatBoost demonstrates superior generalization accuracy. Additionally, XGBoost ranks second in both training speed and accuracy. Lastly, the study delves into the impact of hyper-parameter adjustment in LightGBM, CatBoost, and XGBoost.

Random forests, an enhancement of classification and regression trees (CART), are known for their speed, adaptability, and reliability in mining high-dimensional data. They perform well even when faced with numerous features and limited observations [18], and they can effectively handle censoring for various types of results, comprising continuous, categorical, and time-to-event data. Although they are commonly utilized, and they have good performance in many applications, the theoretical underpinnings of random forests remain poorly understood [79]. [80] surveys several applications of random forests, such as survival data estimation,

probability estimation, and categorization. Additionally, the paper discusses the effects of different feature selection methods, including random selection, no selection, and a combination of deterministic and random selection, on the performance of random forests. Random forests are constructed by aggregating various tree predictors, wherein each tree's projections depend on values sampled randomly from a shared distribution[81]. A complete analysis of [17] random forests model, closely resembling the original algorithm, is provided by [82]. According to the study, the process is consistent and sparsity-adaptive, where the rate of convergence is determined by the number of strong features rather than the quantity of noisy variables. Random forests are experimentally examined in [83]. The study's primary contributions are to investigate two traditional variable selection problems and offer insights into the behavior of the variable importance index based on random forests. Using the random forests score of relevance, the technique ranks the explanatory factors and introduces them stepwise, incrementally.

The user needs to adjust several hyperparameters for the random forest (RF) algorithm. [84] reviewed the impact of these parameters on prediction performance and variable importance measures. While RF generally performs well with default hyperparameter values provided in software packages, its performance can be further improved by adjusting these parameters. The study demonstrates the use of model-based optimization (MBO), a well-known tuning technique. Additionally, to simplify the process, the study introduces the tuneRanger R package, which automates RF tuning with MBO. Current clinical prediction models typically focus solely on assessing mortality risk at birth, without considering the length of hospital

stay. In contrast, (Lee et al., 2021) introduced an innovative approach by incorporating the length of hospital stay alongside conventional mortality risk factors. Utilizing the best logistic regression model and the Clinical Risk Index for Babies II (CRIB-II) score, they quantified mortality risk based on both traditional means and clinical criteria. A subset of the cohort was allocated for training a random forest (RF) model, while the remaining infants were reserved for testing purposes. Impressively, the RF model exhibited a sensitivity of 88% and an AUC of 0.93, surpassing the performance of both the CRIB-II and logistic regression models in mortality prediction. This suggests that the RF model outperforms traditional approaches in accurately predicting mortality risk among newborns.

To enhance the accuracy of mortality predictions, [86] introduce an innovative approach that combines two-dimensional P-spline with random forest techniques. Initially, their method utilizes the random forest evaluator to evaluate the discrepancy between actual death counts and the estimates generated by a specific mortality model, thereby diagnosing the limitations of the LC model. Subsequently, during the forecasting phase, the random forest evaluator is smoothed and projected using a two-dimensional P-spline. This process emphasizes the importance of evaluating demographic consistency in the outcomes. The study rigorously assesses the model's accuracy through an out-of-sample test. [87] utilized the Random Forest classification algorithm to produce up-to-date mortality estimations for extensive tree populations spanning 53 species and 48 ecological sectors. The findings reveal notable recent mortality rates in the Southern and Middle Rocky Mountains, largely linked to increased mortality rates in species like *Abies lasiocarpa*, *Populus*

tremuloides, Pinus contorta, and Pinus albicaulis. Furthermore, the findings suggest that further investigation into susceptible species could prove beneficial for future modeling or prediction of broad forest mortality. Additionally, it is inferred that notable mortality occurrences in particular species may not necessarily impact dynamics across all ecosystems. Although the National Early Warning Score (NEWS) has demonstrated its reliability as a tool for identifying clinical deterioration in hospital wards, there is ongoing debate regarding its applicability in prehospital environments. Contemporary machine learning algorithms can demonstrate better performance than conventional statistical analysis when projecting short-term death. Using prehospital vital indicators, [88] compares the accuracy of NEWS's mortality prediction with that of random forest machine learning. The random forest models underwent training and validation through ten-fold cross-validation. In adult prehospital patients, a random forest algorithm trained with NEWS characteristics outperformed classical NEWS in predicting one-day death.

Reducing the number of variables needed for forecasting is often the main goal of prediction modeling, aiming to enhance efficiency and reduce the workload associated with data collection. [89] assesses different random forest variable selection methods based on the area under the receiver operating curve, the number of variables, computational durations, and prediction error rates. The study compares random forest variable selection techniques across different datasets, including those with binary outcomes, numerous predictors, and imbalanced outcomes. It also compares various method types, such as test-based versus

performance-based methods, and normal random forest versus conditional random forest methods. The Random Forest base classifier, which is essentially a decision tree, is primarily designed using a randomization method of features and data, along with a split criterion that quantifies knowledge gained using traditional precise probabilities. One of the drawbacks of this classifier is its poor performance on datasets with class noise. In response, [90] proposed a new single decision tree model by modifying the Random Forest base classifier with a new criterion. The Credal Random Forest is a novel approach that is comparable to the Random Forest, with its base classifier being the model combined with the randomization process of features.

2.3 Maternity

The impacts of maternity extend significantly to the family, community, and the entire country. In a rural Indian tertiary care teaching hospital, as investigated by [91], the study assesses the causes of maternal death over a five-year period and examines the maternal mortality ratio (MR). The research employs a retrospective, hospital-based analysis of 256 maternal fatalities occurring during this period. Notably, the findings reveal that 60.92% of women succumbed within 12 hours of admission, highlighting delays in reaching tertiary care hospitals, a crucial factor influencing maternal outcomes. The use of R software and the microdata of sibling survival histories taken from the Demographic and Health Surveys (DHS) to implement direct estimation of adult and maternal mortality was defined in full in [92]. Specifically, a set of R scripts was created for the purpose of Population Division. Apart from approximating the death rates and probabilities of dying for

males and females between the ages of 15 and 49, the scripts also generate estimates for maternal mortality rates and the percentage of female deaths among maternal deaths, which are utilized to depict the risks of maternal death for women during pregnancy and childbirth. Furthermore, the scripts enable the calculation of standard errors linked to each adult and maternal mortality statistics produced.

In Queensland, only a small percentage of mothers receive continuous care during their maternity episodes. The Office of the Chief Nursing and Midwifery Officer, entrusted with improving death services, assumed the role of assessing and enhancing the Maternity Models of Care and Workforce. To address this issue, [93] developed a decision-making framework (DMF) aimed at improving the continuity of maternity care models. The study involved constructing a prototype, testing it, and adjusting based on feedback from metropolitan, regional, and rural facilities. The i-DMF and associated tools have been successfully implemented at three remote sites. Among these, three sites have established midwifery continuity-of-carer models specifically for Aboriginal and Torres Strait Islander women, while another is working towards establishing a local rural birth service. One site has utilized the tool for quality assurance of its existing midwifery continuity model. The study by [94] aims to ascertain the pattern of maternal death over the past eight years. They conducted a retrospective analysis of all maternal deaths that occurred at the Edward Francis Small Teaching Hospital between January 1st, 2007, and December 31st, 2014. The study performed case abstraction using a pre-structured questionnaire based on the World Health Organization (WHO) definition of maternal mortality. [95] assessed the impact of five selected approaches on the

response rate to pilot surveys conducted before the comprehensive National Maternity Survey in England. Women who were three months postpartum and randomly selected from birth registrations participated in the cross-sectional population-based questionnaire surveys, which constituted the pilot national maternity surveys. Notably, the first pilot survey (pilot 1, $n = 1000$) experienced a lower response rate than anticipated. The study not only identifies new approaches for engaging populations in survey research but also employs randomized controlled trials to evaluate the effectiveness of each method. Despite the promotion of Millennium Development Goal (MDG) 5 aiming for a substantial reduction by 2015, maternal mortality rates have shown an upward trend. [95] analyzed data from a public hospital in Kumasi, Ghana, to model maternal mortality using the Box-Jenkins approach via the univariate version of time series autoregressive integrated moving average (ARIMA). The results suggested that, based on the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC), the ARIMA (1, 1, 1) model was the most suitable for modeling and forecasting monthly maternal cases. The study's results emphasized a notable increase in maternal death instances from 2000 to 2011.

In the examination of reproductive-age mortality, [96] focused specifically on maternal fatalities in Jordan between 2007 and 2008 among women aged 15 to 49. The study aimed to pinpoint the reasons behind maternal mortality, calculate the maternal death rate per 100,000 live births among Jordanian women, and compare these results with a previous study encompassing the years 1995–1996. The results of the study revealed that out of 1,406 detected fatalities among women of

reproductive age, 76 were maternal deaths, occurring within 397,588 live births, resulting in a maternal mortality ratio of 19.1 deaths per 100,000 live births. Furthermore, the study observed a decline in maternal fatalities nationwide. Interestingly, the maternal mortality ratio of 19.1 deaths per 100,000 live births for 2007–2008 demonstrated a significant 53.9% decrease over the 12 years since the 1995–1996 report, nearly reaching the 75% reduction target suggested by Millennium Development Goal 5. [97] conducted an examination of maternal death ratios at Kumasi's Okomfo Anokye Teaching Hospital for the years 2000–2010. To predict and forecast maternal mortality ratios (MMR), the study explores the applicability of the Box-Jenkins Approach to time series autoregressive integrated moving average (ARIMA). Notably, the study utilizes the AIC as a diagnostic tool for the model. The results of the analysis reveal that, while the hospital's maternal mortality ratio (MMR) remained relatively stable, it exhibited a concerning average quarterly MMR of 967.7 per 100,000 live births. This figure is approximately twice the national ratio of 451 per 100,000 live births.

To present a model-based technique for calculating maternal mortality at the subnational level and demonstrate how to apply it to the estimation of MMRs and MMRs in each of Bangladesh's 64 districts. Maternal mortality at the subnational level was estimated by [98] using an empirical Bayesian prediction approach based on the spatial distribution of such factors. Monitoring and forecasting maternal mortality rates (MMRs) in regions like Africa is crucial for health strategy planning by governments, international organizations, and non-governmental organizations (NGOs), aligning with the United Nations Sustainable Development Goals (SDG)

for the period 2015 to 2030, which prioritize the reduction of maternal mortality. [99] employ the autoregressive integrated moving average (ARIMA) model combined with point regression to assess trends. Additionally, the study incorporates the Holt method and the machine-learning Prophet Forecasting Model. The findings reveal a notable average annual percentage change (APC) of -2.6% in MMR declines across Africa. Despite this observed trend towards decline, the maternal mortality rate (MMR) in Africa remains relatively high. At Christian Medical College & Hospital, a tertiary care facility, [100] conducted a study to assess Maternal Mortality Rate (MMR), examine various factors influencing it, and explore potential interventions for preventing maternal mortality over the past five years. The study involved a detailed examination of each maternal death record, considering sociodemographic characteristics, contributing factors, methods of treatment, and final outcomes. Notably, slum regions in both rural and urban areas accounted for over 90% of maternal mortality cases. Furthermore, 10.4% of cases received no primary care, and 57.8% of cases received care from unskilled delivery attendants. The study identifies three key strategies to reduce maternal mortality: ensuring adequate prenatal care, developing effective methods for preventing and managing unintended pregnancies, and improving societal attitudes toward expectant mothers.

The study by [101] outlines crucial factors to consider for improving our understanding of the quality of maternity care. Achieving high quality care involves providing a baseline level of care to all expectant mothers and their newborns, as well as addressing the needs of those requiring a higher level of care. The ultimate

goals include achieving optimal medical outcomes for both mother and child, ensuring satisfaction among users and providers, and maintaining sound managerial and financial performance. While all these factors are significant, determining the best approach to prioritize efforts for enhancement depends on cultural norms, expectations, and resource availability. [102] conducted a retrospective analysis of 73 maternal deaths that occurred between January 1, 2012, and December 31, 2013, to investigate the maternal death rates at Assamese Rural Medical College. The study revealed that during the investigation period, there were 73 maternal fatalities out of 10,291 live births, resulting in a Maternal Mortality Ratio (MMR) of 709.35. Many maternal deaths occurred between 12 and 24 hours after admission, with a concentration in the 18 to 23 age range. The maternal mortality ratio (MMR), representing the number of maternal fatalities per 100,000 live births in a population, is a crucial indicator. The United Nations Maternal Mortality Estimation Inter-Agency Group (UN MMEIG) periodically publishes country-specific MMR estimates to track advancements in reducing maternal deaths. These estimations are utilized to assess national and regional progress toward achieving Millennium Development Goal (MDG) 5, which sought a 75% decrease in the MMR from 1990 to 2015. [102] devised a Bayesian model for maternal mortality estimation, building upon the UN MMEIG's multilevel regression model. By integrating an ARIMA time series model that accounts for fluctuations in the rate of change in the MMR, the proposed model investigates trends influenced by the data. To investigate the frequency, underlying causes, and temporal distribution of maternal deaths over the preceding decade (1990–1999), maternal mortality ratios (MMR) were calculated

by [103]. Confidence intervals around these values were derived using the Poisson assumption. Specifically, a non-linear regression model was employed to identify the optimal temporal trajectory for MMR throughout the research decade. Among 171,621 deliveries, a total of 4,154 maternal fatalities occurred, resulting in an MMR of 2,420 deaths per 100,000 individuals—one of the highest globally. The findings suggested that implementing effective interventions targeting deaths from eclampsia, ruptured uterus, and anemia could potentially halve maternal mortality at the study site.

CHAPTER 3

COMPARATIVE ANALYSIS OF STOCHASTIC MORTALITY MODELS-ENHANCED BY MACHINE LEARNING.

This chapter focuses on achieving the first two objectives of the study which are:

- a) To fit stochastic mortality models for various countries, considering the full range of age groups, and select the best-performing model based on four performance criteria.
- b) To employ machine learning techniques to enhance the performance of the top four models.

We will compare ten different stochastic mortality models across four countries for both genders. Then, we will evaluate the performance of these models and identify the model(s) that best fit the historical death rate datasets across age ranges, genders, and countries considered in the study. Specifically, we will analyze data from age 0 to 100 and the years 1950 to 2020, ensuring consistency throughout the analysis.

We also plan to utilize machine learning techniques to identify weaknesses in the Parametric Stochastic Model, thereby enhancing the performance of the robust model(s) across selected age groups, genders, and countries using historical data from Belgium. This data set includes detailed records of all deaths within specific years and age ranges. It's important to note that among the models compared, only the APC model demonstrates linearity in parameters, whereas the CBD model and

its variations lack age-specific parameters. We will then proceed to present the results for both genders across all four countries.

3.1 OVERVIEW

A brief overview of the current literature on the comparison of stochastic mortality models are enumerated below.

In actuarial science, it is commonly acknowledged that a clear understanding of historical patterns in mortality rates is essential for accurate modeling and future forecasts of underlying changes, particularly when it comes to pricing and reserving for annuity and pension products. An alarming focus on risk management and measurement for insurers and pension funds using stochastic mortality models has been observed over the last decade. Additionally, declines have been mostly noted in the populations of numerous wealthy nations [29]. The understatement of premiums and reserves for annuity and pension products could have catastrophic repercussions for the financial institutions involved if these downward trends are not properly factored in. In 1992, Lee and Carter [1], presented an approach that offers a novel method for estimating the rate and age distribution of mortality. The approach that was initially developed to analyze mortality rates in the USA has found extensive use in analyzing historical data from various other countries. Various modifications and extensions to this model, which incorporate diverse assumptions about the data, constraints on parameters, and suitable methodologies, have enhanced the model's accuracy in fitting historical data and its ability to forecast future trends. Their forecasting strategy combines a straightforward

approach with a statistical time series method to address the age distribution of mortality. Specifically, the approach involves taking a time series of logarithms of age-specific mortality rates, adding a bilinear term to an age-specific component that remains constant over time. This bilinear term is the result of multiplying a time-varying parameter representing the total mortality rate by an age-specific component. The Lee–Carter model is the most well-known extrapolative technique for predicting death rates. Following the Lee–Carter model, other models have been developed; these are considered extensions of the original Lee–Carter model. These extensions include the Age-period-cohort model, the Cairn Blakes Dowd model, and its variations, and a cohort extension of the Lee-Carter model 1992 [49].

Several parametric stochastic mortality models have been compared in the literature from different perspectives. [8] conducted a quantitative comparison of eight different stochastic mortality models. Their study evaluated the models based on their consistency and robustness relative to historical data, with each model demonstrating specific strengths. The family of generalized age-period-cohort stochastic mortality models is defined within the framework of generalized linear and nonlinear models, as highlighted by [59]. Similarly, [29] conducted a comparative study that compared the basic Lee-Carter mortality model to its extensions to examine the effectiveness of the improvements made. The work of [104] emphasized the superior performance of all variants and expansions of the original Lee-Carter method when forecasting the logarithm of mortality rates.

Furthermore, [61] discussed the implementation of the Age-Period-Cohort-Improvement (APCI) model, which shares features with both the Lee-Carter and APC

models, for death projection. Additionally, the study conducted by [62] compared eight distinct mortality models.

3.2 Methodology

In this section, we highlight the methods employed in this chapter.

3.2.1 Notation and Assumptions

Here, we assume that the death rate is similar within age groups and time but varies among different age groups and periods. Independent death counts $\pi_{a,p}$, for different individual features considered, such as age, gender, and year of birth, are assumed. Furthermore, the death counts are assumed to follow a Poisson distribution with a parameter proportional to exposures. [58],[68],[6].

$$\pi_{a,p} \sim Pois(\varphi_{a,p}, \psi_{a,p}) \quad (3.1)$$

In addition, random error terms are assumed to be independently and identically distributed with equal variances [40], [105]. However, the assumption of error terms having constant variances is not realistic [34]. We are considering ten different stochastic mortality models in this study; therefore, we will maintain consistency with the notation throughout. Table 1 below presents the notation for the parameters, features, and their descriptions. The central death rate, denoted as $\varphi_{a,p}$ for age group a in period (calendar year) p , is defined as the ratio of death counts $\pi_{a,p}$ for age group a during period p to the central exposure-to-risk $\psi_{a,p}$ for age a at period p . It is mathematically given as:

$$\varphi_{a,p} = \frac{\pi_{a,p}}{\psi_{a,p}} \quad (3.2)$$

Table 3.1 Notation of parameters, features (variables) and their descriptions.

Parameters/ variables	Descriptions
a	Individual age
$\bar{a}$	Average of individual age
p	Period (or calendar year)
c	Individual year of birth (cohort)
n	Number of observations
$\varphi_{a,p}$	Central death rate for age a during period p
$\hat{\varphi}_{a,p}$	Estimated central death rate for age a during a calendar -year p
$\pi_{a,p}$	the total number of deaths observed in a population in period p at age group a on their last birthday.
$\hat{\pi}_{a,p}$	Expected total death counts estimated by a certain mortality model
$\psi_{a,p}$	central exposure-to-risk for age a at period p (i.e. number of lives at risk)
$\kappa_a^{(1)}$	Average of log of mortality rates of age group a – Age effect
$\kappa_a^{(2)}$	Changes in age-specific mortality patterns in response to period effects
$\kappa_a^{(3)}$	Changes in age-specific mortality patterns in response to cohort effects
ρ_p	Random period parameter
$\varepsilon_{a,p}$	Random error term
δ_c	Random cohort effect, a function of the year of birth
$\lambda_{a,p}$	Parametric predictor
ω	a constant parameter
$\rho_p^{(1)}, \rho_p^{(2)}, \rho_p^{(3)}$	Random period parameter
$\beta_{a,p}$	Binary weight for data cell
τ_y	ratio of observed death counts to the estimated values
$\hat{\tau}_{ML}$	The Predicted values of τ_{ML}
$\hat{\varphi}_{ML}(\mathbf{y})$	Enhanced mortality rates by a certain machine Learning model

3.2.2 Parametric Mortality Predictor Structures

There have been development of new stochastic mortality models after the first proposed stochastic mortality model by [1]. These developments are highlighted in the works of [106], [107], [104], [57], [52], [108], [8], [9], [109] etc.

The stochastic models exhibit substantial differences due to various factors, including the sources of randomness impacting mortality improvements across different ages, assumptions regarding the smoothness of age and period dimensions, the inclusion or exclusion of cohort effects, and the estimation technique employed. This study examines both versions of cohort extension of the Lee and Carter (1992) by [52]. Similarly, we considered Cairns Blake Dowd and all its three extensions and the two specifications of Plat (2009) model. We named each of the models as follow: LC for Lee and Carter (1992) model, RHc for the first version of Renshaw and Haberman (2006), that is a non-flat cohort extension of LC model, RHf for second extension of Renshaw and Haberman (2006) model with a flat cohort (i.e., $\kappa_a^{(3)} = 1$), APC for Age-Period-Cohort model, PLAT (2009) model and its variant, addressed the shortcomings of the aforementioned models while combining their strengths. They are represented as FPLAT and RPLAT for the full and reduced versions, respectively.

In Table 1 below, the ten model structures are tabulated together with their identifiability constraints. Each of these models has been discussed in the previous chapters, particularly in one. Among the ten predictors in this study, only LC, RHc, and RHf predictors are non-linear in the parameters whereas the remaining six predictors possess linear nature. For the constraints of the three nonlinear predictors,

there is room for introducing other constraints. Notably, every model in our study captured the period effects. Similarly, all the models incorporate cohort effects except LC and CBD.

3.2.3 Uniqueness Constraints

Uniqueness problems can arise in certain stochastic mortality models when there is no impact on the mortality rates $\varphi_{a,p}$ despite rescaling and shifting the parameters [110]. For example, if we multiply all historical ρ_p values by a constant b and simultaneously divide all $\kappa_a^{(2)}$ values by b , the rescaling of parameters in the model does not affect the fitted $\hat{\varphi}_{a,p}$. Similarly, when the random-walk model is refitted, the ρ_p 's values changed, but the fitted $\hat{\varphi}_{a,p}$ values remain the same. To address this kind of problem, uniqueness constraints are introduced. Essentially, the selection of constraints is somewhat arbitrary, even though certain options are more logical than others [8]. It is worth noting that all models under consideration, except for CBD, encountered uniqueness issues. This required the imposition of uniqueness constraints, as indicated in the last column of Table 1.

Table 1 Stochastic Mortality Models and their uniqueness Constraints.

S/No	Predictors (Names)	Constraints
1.	$\lambda_{a,p} = \kappa_a^{(1)} + \kappa_a^{(2)} \rho_p + \varepsilon_{a,p}$ (LC)	$\sum_a \kappa_a^{(2)} = 1$; $\sum_p \rho_p = 0$.
2.	$\lambda_{a,p} = \kappa_a^{(1)} + \kappa_a^{(2)} \rho_p + \kappa_a^{(3)} \delta_c + \varepsilon_{a,p}$. (RHc)	$\sum_a \kappa_a^{(2)} = 1$; $\sum_a \kappa_a^{(3)} = 1$; $\sum_p \rho_p = 0$; $\sum_c \delta_c = 0$.
3.	$\lambda_{a,p} = \kappa_a^{(1)} + \kappa_a^{(2)} \rho_p + \delta_c + \varepsilon_{a,p}$. (RHf)	$\sum_a \kappa_a^{(2)} = 1$; $\sum_p \rho_p = 0$; $\sum_c \delta_c = 0$.
4.	$\lambda_{a,p} = \kappa_a + \rho_p + \delta_c + \varepsilon_{a,p}$ (APC)	$\sum_p \rho_p = 0$; $\sum_y c \delta_c = 0$; $\sum_c \delta_c = 0$.

5. $\lambda_{a,p} = \rho_p^{(1)} + \rho_p^{(2)}(\bar{a} - a) + \varepsilon_{a,p}$. (CBD) —
 6. $\lambda_{a,p} = \rho_p^{(1)} + \rho_p^{(2)}(\bar{a} - a) + \delta_c + \varepsilon_{a,p}$. $\sum_c \delta_c = 0; \sum_c c\delta_c = 0$.
(M6)
 7. $\lambda_{a,p} = \rho_p^{(1)} + \rho_p^{(2)}(\bar{a} - a) + \rho_p^{(3)}[(\bar{a} - a)^2 - \hat{\sigma}^2] + \delta_c + \varepsilon_{a,p}$ (M7) $\sum_c \delta_c = 0; \sum_c c\delta_c = 0; \sum_{c^2} c\delta_c = 0$
 8. $\lambda_{a,p} = \rho_p^{(1)} + \rho_p^{(1)}(\bar{a} - a) + (\omega - a)\delta_c + \varepsilon_{a,p}$ (M8) $\sum_c \delta_c = 0$
 9. $\lambda_{a,p} = \kappa_a + \rho_p^{(1)} + \rho_p^{(2)}(\bar{a} - a) + \rho_p^{(3)}a^* + \delta_c + \varepsilon_{a,p}$, $\sum_c \delta_c = 0; \sum_c c\delta_c = 0; \sum_p \rho_p^{(3)} = 0$
where $a^* = \max[(\bar{a} - a), 0]$. (FPLAT)
 - 10 $\lambda_{a,p} = \kappa_a + \rho_p^{(1)} + \rho_p^{(2)}(\bar{a} - a) + \delta_c + \varepsilon_{a,p}$ $\sum_c \delta_c = 0; \sum_c c\delta_c = 0; \sum_p \rho_p^{(3)} = 0$.
(RPLAT)
-

3.2.4 Description of datasets

The historical mortality datasets of four different countries are obtained from the Human Mortality Database [111]. The data are arranged in rectangular array formats, in which the rows represent age groups a_i ($i = 1, 2, \dots, k$), and a_k is the maximum age, an individual can attain. The columns of the rectangular array represent calendar years p_h ($h = 1, 2, \dots, H$), and p_H denotes the maximum calendar year considered. One rectangular array for death counts, D_{at} and a separate one for central exposure E_{at} , for male, female, and total separately, for each county. The four countries considered are: Japan, Belgium, Italy, and Switzerland. Each country recorded the information on deaths for different periods of time, therefore, we have limited both age groups and calendar year to 0 to 100 years and 1950 to 2020 respectively. Each of the four datasets consists of 101 rows and 71 columns, resulting in 7171 data points.

3.2.5 Stochastic Mortality Models fitting

For the comparative study, we utilize an R package called StMoMo [27] to fit all the parametric mortality models. As mentioned earlier in the previous section, we assume a Poisson distribution of death counts in all cases. Although a variety of link functions could be considered, we find it convenient to pair the Poisson distribution with the log link function. Therefore, we employ the "log link" function for easy comparison of the models under consideration. The estimation of parameters for stochastic mortality models involves maximizing the log-likelihood of the model. When considering the Generalized Age-Period-Cohort (GAPC) mortality models, the mathematical expression for the model's log-likelihood is represented by:

$$l(\varphi_{a,p}, \hat{\varphi}_{a,p}) = \sum_{a,p} \beta_{a,p} \{ \varphi_{a,p} \log \hat{\varphi}_{a,p} - \hat{\varphi}_{a,p} - \log \varphi_{a,p} \}, \quad (3.3)$$

with the estimated rates

$$\hat{\varphi}_{a,p} = E_{a,p} f^{-1}(\kappa_{1,a} + \sum_j \kappa_{2,a}^j \rho_p^j + \kappa_{3,a} \delta_c). \quad (3.4)$$

For $j = 1, \dots, N$, by the model, where f^{-1} is the inverse of the *link function* used and $\beta_{a,p}$ represent the binary weight for data cell. The usual iterative process in mortality literature to maximize log-likelihood is Newton-Raphson specific to each model [6]. An accurate model fitted to the historical datasets tends to provide quality mortality projection. However, the model best fits the data does not imply quality of the forecast in some other cases [61]. The choice of ω for M8 depends on the selected age range [29], which, in this case, is limited to the age range covered by the available data, that is $\omega = 100$.

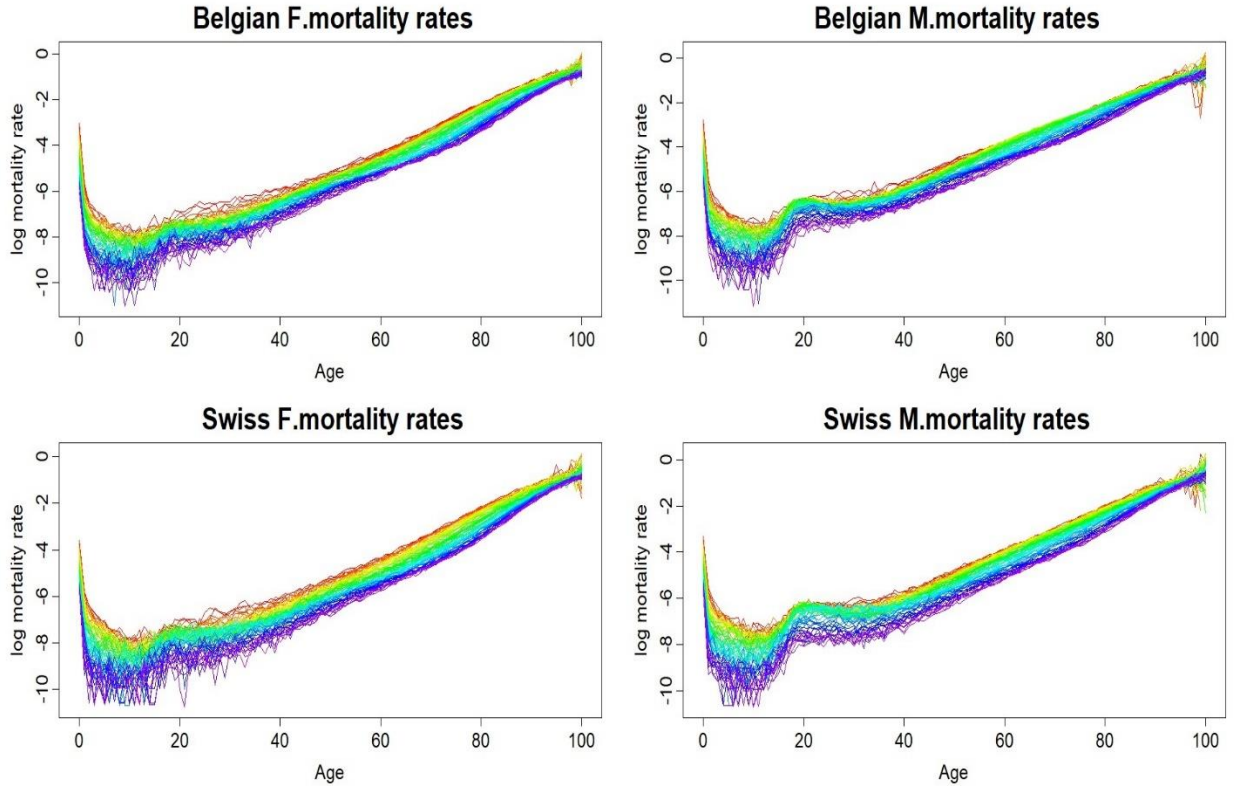


Figure 3 Pattern of logarithm of death rates according to age and time for Male(M) and Female(F) population of both Belgium and Switzerland.

3.2.6 Diagnostic of the models

Models with more parameters are more likely to produce a better fit compared to models with fewer parameters. Examining the goodness of fit of models is essential for selecting the best model among others. To address the issue of models taking advantage of more parameters over others with fewer parameters, it has become a tradition to modify the maximum likelihood criterion by including a penalty function so that models with more parameters are penalized. Several such penalty functions have been proposed in the literature, for example, the Akaike Information Criterion (AIC), which is one of the most common choices. Three other metrics—Log-likelihood, Deviance, and Root Mean Square Error—are employed together with AIC. AIC allows the evaluation of multiple models and the selection of the one that

strikes the best balance between fitting and complexity. The values of the four metrics are shown in Tables 3.3a, 3.3b, 3.3c, and 3.3d for all the 10 models, for each gender separately, and for the four countries under study.

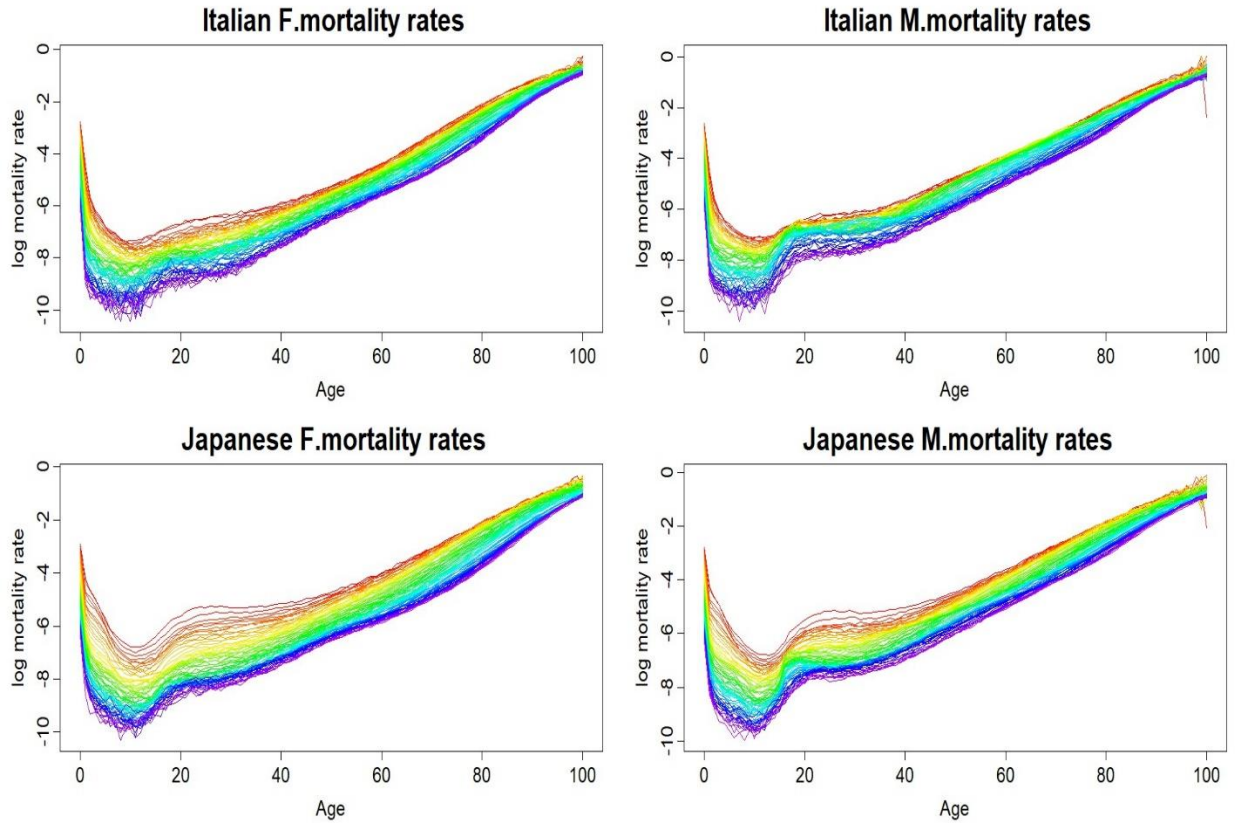


Figure 4 Pattern of logarithm of death rates according to age and time for Male(M) and Female(F) population of both Italy and Japan.

3.3 Machine Learning models

In recent years, studies have leveraged machine learning approaches to improve the accuracy of mortality rate estimates derived from parametric mortality models. Some notable works that have employed these approaches include [68], [11], [71], [86], [69], and [73]. Machine learning models can uncover trends and factors in

historical datasets that traditional mortality models may overlook, thus enhancing the performance of mortality models in fitting historical mortality data. Commonly utilized machine learning algorithms for this purpose include Decision Tree, Random Forest, Artificial Neural Network (ANN), K-nearest neighbor (KNN), Support Vector Machine (SVM), and Gradient Boosting (GB). Unlike the study by [11], which utilized only two mortality models and relied solely on Decision Tree machine learning models, this study delves into ten stochastic mortality models.

This study incorporates three machine learning algorithms—Decision Tree (DT), Random Forest (RF), and Gradient Boosting (GB)—to enhance the performance of parametric stochastic mortality models. Drawing from previous research [86], these algorithms are employed to refine the fitting accuracy of the best-performing model(s). Specifically, this involves refining the ratio between observed death counts and the estimated values derived from the model that best aligns with the historical datasets from the four countries under examination. By scrutinizing the quality of mortality estimates derived from stochastic mortality models in this way, it becomes possible to comprehensively evaluate the model's strengths and weaknesses, thereby improving its performance. This evaluation is conducted with respect to individual features such as age, period, cohort, and gender.

3.3.1 Model Notation and Assumptions

The following qualitative variables will be considered: individual age denoted by " a ," gender denoted by " g ," period (or calendar year) denoted by " p ," and year of birth (or cohort) denoted by " c ." Everyone is assigned the feature $\mathbf{y}$, where $\mathbf{y}$ is a vector containing all the considered features, represented as $\mathbf{y} = (a, g, p, c)$ in the

space $Y = A \times G \times P \times C$. Where feature spaces $A = \{0, \dots, a_k\}$, $G = \{females, males\}$, $P = \{p_1, \dots, p_H\}$, and $C = \{c_1, \dots, c_m\}$, with their levels enclosed in the parentheses. As stated earlier, the death counts, π_y follow Poisson distribution. This is expressed mathematically as:

$$\pi_y \sim Pois(\tau_y \cdot \hat{\pi}_y), \text{ for } \tau_y \equiv 1 \text{ and } \hat{\pi}_y = \hat{\phi}_{a,p} \cdot \psi_{a,p}$$

The $\tau_y \equiv 1$ signifies that the given mortality model precisely matches the crude rates. However, in practical scenarios, a mortality model might either overestimate ($\tau_y \leq 1$) or underestimate ($\tau_y \geq 1$) the crude rates [86].

Here we denote $\hat{\pi}_y$ as the estimated expected death counts using a given stochastic mortality model (i.e. the product of estimated death rates by a particular stochastic mortality model and the central exposure-to-risk, $\psi_{a,p}$), and $\hat{\phi}_{a,p}$ as the estimated central death rates.

Additionally, we introduce τ_y , which represents the ratio between observed death counts and the estimated values derived by a certain model and given as:

$$\frac{\pi_y}{\hat{\pi}_y} \sim \text{age} + \text{period} + \text{cohort} + \text{gender}, \quad (3.5)$$

This ratio, τ_y is estimated by each machine learning model using the four features: age, period, time of birth (cohort), and agender. Equation 3.5 is estimated using the three machine learning techniques and we denote the calibration by τ_{ML} . Then, τ_{ML} is predicted using the R command “predict,” and the predicted estimate is represented by $\hat{\tau}_{ML}$. Subsequently, the enhancement of mortality rates through the

machine learning is achieved by multiplying $\hat{\tau}_{ML}$ with $\hat{\varphi}_{a,p}$, which represent the estimated central death rates from a stochastic mortality model. That is,

$$\hat{\varphi}_{ML}(\mathbf{y}) = \hat{\tau}_{ML} \hat{\varphi}_{a,p}, \quad \text{for } \mathbf{y} \in Y. \quad (3.6)$$

Here, we will employ three performance metrics to examine and compare the three applied machine techniques. These metrics include Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Coefficient of Determination (R^2). RMSE formula is discussed in chapter one. The coefficient of determination (R^2) gives the percentage of variability in death counts that can be attributed to factors like age, gender, calendar year, and birth year.

$$R^2 = 1 - \frac{\sum_{a,p} (\varphi_{a,p} - \hat{\varphi}_{ML}(\mathbf{y}))^2}{\sum_{a,p} (\varphi_{a,p} - \bar{\varphi}_{a,p})^2} \quad (3.7)$$

where $\bar{\varphi}_{a,p}$ is the average of the observed death counts.

MAE evaluates model accuracy by quantifying the absolute variance between the predicted values and observed values of death counts. Its mathematical expression is given as:

$$MAE = \frac{1}{n} \sum_{a,p} |\varphi_{a,p} - \hat{\varphi}_{ML}(\mathbf{y})| \quad (3.8)$$

3.3.2 Decision Tree Machine Learning

Decision Tree (DT) technique is based on partition of the feature space $\mathbf{Y}$, through a sequence of binary splits, and set of splitting rules used to segment the predictor

space. For *DT*, an R package called ‘rpart’ [112], is utilized to calibrate the ratio τ_y , and it is structured as follows:

```
 $\tau_{DT} <- \text{rpart}(\tau_y \sim \text{age} + \text{year} + \text{cohort} + \text{gender}, \text{data} = \text{data}, \text{method} = "poisson",$   
 $\text{cp} = \text{cpvalue}).$ 
```

Here, *cp* is the cost complexity parameter, allowing the regression tree algorithm to perform a specific number of iteration steps. The *cpvalue* is chosen as $2e^{-3}$ following the work of [68].

3.3.3 Random Forest Machine Learning

The Random Forest (RF) method is a regression tree technique that utilizes bootstrap aggregation and predictor randomization to achieve a high level of predictive accuracy. It is employed to estimate the ratio τ_y , using the R package “*randomForest*” [113], [114]. This method is valued for its simplicity, accuracy, and effectiveness in handling small sample sizes and high-dimensional feature spaces [18]. During each split in the tree, its algorithms consider only a randomized subset of predictors [11]. However, to manage computational costs effectively, it is essential to select the optimal number of trees for the procedure. This number must strike a balance: it should be sufficient to explain a significant portion of the variance and achieve a low mean squared error (MSE) without being excessively large.

3.3.4 Gradient Boosting Machine Learning

Gradient Boosting method, utilizing the R package “*gbm*” [115] and its associated R command *gbm*(·), is used to calibrate the ratio τ_y . This approach employs an adaptive learning algorithm to iteratively combine weak predictors, aiming to

minimize in-sample loss. When utilizing the `gbm` package, it is crucial to select the key parameters such as the number of trees (*n.trees*), cross-validation folds (*cv.folds*), tree depth (*interaction.depth*), and learning rate (*shrinkage*). The choice of the number of trees is particularly important, as it impacts training set error and computational costs. Too few trees lead to overfitting and too many increasing computational resources. Cross-validation folds, usually five-fold cross-validation, are chosen based on dataset size to balance testing accuracy and data utilization. The interaction depth parameter plays a significant role in determining the model's complexity and predictive performance. It dictates the maximum nodes per tree or level of variable interactions allowed, thereby influencing how the model handles complex relationship within the data [11].

3.4 Results and Discussion

We applied the ten models to historical datasets from Switzerland, Italy, Belgium, and Japan, with the results presented in Tables 2, 3, 4, and 5, respectively. Our comparative analysis of these models relies on four key metrics to evaluate their performance. All models were fitted to the age range 0-100 and the period 1950-2020, with assessment metrics including Log Likelihood (L), Deviance (Dev), Akaike Information Criterion (AIC), and Root Mean Square Error (RMSE), where lower values indicate better fit. Interestingly, none of the considered models consistently demonstrates the best fit across all four countries and genders. However, the optimal fit is observed among three models, each excelling in fitting historical datasets. These three models exhibit the best fit for different genders across various countries, as highlighted in Tables 2, 3, 4, and 5 Specifically, RHc demonstrates the

best fit for the female gender in Belgium and Italy, as well as for the male gender in Japan. Similarly, RHc shows the best fit for the female dataset in both Switzerland and Japan, and for the male dataset in Italy. On the other hand, the PLAT model exhibits the best fit for male datasets in Switzerland and Belgium.

Table 2 Log-likelihood, Deviance, AIC, and RMSE statistics for stochastic mortality models. Ages 0 -100 and years 1950–2020, Switzerland.

Models	Np	Goodness of Fit Metrics									
		Log-likelihood		Deviance		BIC		AIC		RMSE	
		F	M	F	M	F	M	F	M	F	M
RHc	540	-27082.67	-29565.53	8648.33	11387.26	58959.35	63925.07	55245.34	60211.06	21.2782	26.3262
RHf	441	-27012.77	-28356.73	8508.24	8967.41	57940.64	60628.57	54907.53	57595.46	18.3770 (1)	18.9093 (2)
LC	271	-27771.12	-30293.16	10025.21	12848.75	57948.12	62992.21	56084.24	61128.32	20.1359 (3)	23.06573
RPLAT	409	-29035.21	-30236.96	12559.34	12729.19	61701.44	64104.93	58888.42	61291.91	24.2672	22.8613 (3)
CBD	142	-222963.35	-225260.58	400453.69	402797.47	447187.4	451781.8	446210.7	450805.2	95.9013	95.5737
M6	311	-100278.69	-126549.42	154867.31	205259.09	203318.4	255859.8	201179.4	253720.8	64.8677	80.6014
M7	381	-70829.64	-86408	96000.8	124988.18	145041.7	176198.4	142421.3	173578	52.9319	66.4992
M8	311	-82704.2	-102408.52	119709.69	156994.33	168169.4	207578	166030.4	205439	54.707	71.6694
FPLAT	479	-27094.76	-28039.01	8668.93	8331.17	58441.98	60330.48	55147.52	57036.01	18.6875 (2)	17.887 (1)
APC	340	-30807.18	-32089.46	16098.32	16430.89	64632.82	67197.38	62294.36	64858.92	29.45855	27.77897

Table 3 Log-likelihood, Deviance, AIC, and RMSE statistics for stochastic mortality models. Ages 0 - 100 and years 1950–2020, Italy.

Models	Np	Goodness of Fit Metrics									
		Log-likelihood		Deviance		BIC		AIC		RMSE	
		F	M	F	M	F	M	F	M	F	M
RHc	540	-37442.51	-46720.82	13854.37	30183.23	79679.03	107563	75965.02	94521.64	69.88237(1)	117.739(3)
RHf	441	-40768.22	-46803.16	20505.79	30347.92	85451.55	97521.43	82418.44	94488.32	95.6833(2)	101.1453(1)
LC	271	-49763.09	-84490.83	38495.54	105723.3	101932.1	171387.6	100068.2	169523.7	151.2138	248.0754
RPLAT	409	-68302.11	-79008.77	75573.57	94759.13	140235.2	161648.6	137422.2	158835.5	182.1018	175.9149
CBD	142	-3000078	-2995763	5939125	5928268	6001416	5992788	6000439	5991811	1315.659	1469.854
M6	311	-1391332	-1773137	2721633	3483016	2785425	3549035	2783286	3546896	1106.56	1357.592
M7	381	-740129.7	-970498.6	1419229	1877739	1483642	1944380	1481021	1941759	860.0123	1116.205
M8	311	-991451.2	-1312633	1921872	2562008	1985663	2628027	1983524	2625888	927.2254	1210.512
FPLAT	479	-44505.79	-49784	27980.93	36309.53	93264.04	103820.4	89969.57	100525.9	112.179(3)	111.6571(2)
APC	340	-74426.8	-95493.35	87822.96	127728.3	151872.1	194005.2	149533.6	191666.7	204.7337	252.1048

Table 4 Log-likelihood, Deviance, AIC, and RMSE statistics for stochastic mortality models. Ages 0 -100 and years 1950–2020, Belgium.

Models	Np	Goodness of Fit Metrics									
		Log-likelihood		Deviance		BIC		AIC		RMSE	
		F	M	F	M	F	M	F	M	F	M
RHc	540	-62983.81	-49584.45	60502.75	31827.71	130761.6	103962.9	127047.6	100248.9	265.8835(3)	145.5171(1)
RHf	441	-53105.67	-64264.23	40746.47	61187.27	110126.4	132443.6	107093.3	129410.5	207.9211(1)	264.3412(2)
LC	271	-154644.67	-102235.15	243856.4	137129.11	311695.2	206876.2	309831.3	205012.3	426.7961	358.0528
RPLAT	409	-108748.8	-106843.53	152032.73	146345.88	221128.6	217318.1	218315.6	214505.1	380.2086	365.5948
CBD	142	-3788904.41	-3786959.32	7512343.95	7506577.46	7579069	7575179	7578093	7574203	2070.994	2287.14
M6	311	-1472532.95	-2055794.03	2879601.04	4044246.87	2947827	4114349	2945688	4112210	1560.24	1986.98
M7	381	-962103.86	-1170353.64	1858742.86	2273366.1	1927590	2344090	1924970	2341469	1282.467	1618.513
M8	311	-1079260.81	-1459848.55	2093056.75	2852355.92	2161283	2161283	2159144	2920319	1282.727	1691.182
FPLAT	479	-57477.58	-71219.84	49490.3	75098.51	119207.6	146692.2	115913.2	143397.7	258.2702(2)	280.4921(3)
APC	340	-181803.29	-172682.48	298141.71	278023.78	366625	348383.4	364286.6	346045	441.3585	505.1344

Table 5 Log-likelihood, Deviance, AIC, and RMSE statistics for stochastic mortality models. Ages 0 -100 and years 1950–2020, Japan.

Models	Np	Goodness of Fit Metrics									
		Log-likelihood		Deviance		BIC		AIC		RMSE	
		F	M	F	M	F	M	F	M	F	M
RHc	540	-29031.94	-30167.96	8359.48	8547.94	62857.89	65129.94	59143.88	61415.93	24.8522(1)	24.9701 (1)
RHf	441	-29328.32	-31018.55	8952.24	10249.12	62571.75	65952.22	59538.64	62919.11	25.2131(3)	27.789 (3)
LC	271	-31642.23	-38819.66	13580.05	25851.34	65690.34	80045.21	63826.45	78181.32	36.3997	61.9753
RPLAT	409	-34199.29	-35849.51	18694.17	19911.04	72029.59	75330.04	69216.57	72517.02	41.25488	42.1738
CBD	142	-410994	-426124.4	772283.6	800460.9	823248.6	853509.5	822272	852532.9	184.9207	197.6327
M6	311	-203111.8	-274914.2	356519.1	498040.3	408984.5	552589.3	406845.5	550450.3	141.793	185.4242
M7	381	-128850.6	-166069.8	207996.8	280351.7	261083.6	335522.1	258463.2	332901.7	112.6105	152.3154
M8	311	-159976.9	-215359.8	270249.5	378931.7	322714.9	433480.6	320575.9	431341.6	118.3455	166.1929
FPLAT	479	-29344.88	-30814.55	8985.3	9841.12	62942.22	65881.57	59647.76	62587.1	24.9206 (2)	26.4800(2)
APC	340	-35418.01	-37913.2	21131.62	24038.41	73854.47	78844.85	71516.01	76506.4	47.1829	54.1233

In addition, none of the models demonstrates robustness across all countries and genders. All these models have a high number of parameters, except for the CBD and LC models, which have the fewest estimated parameters. Both models are relatively parsimonious, although their metrics indicate poor fitting to the historical data. Figures 5 and 6 below display the fitted values of all ten models for both male and female genders compared to the observed death rate (CR) of the historical datasets from the four countries.

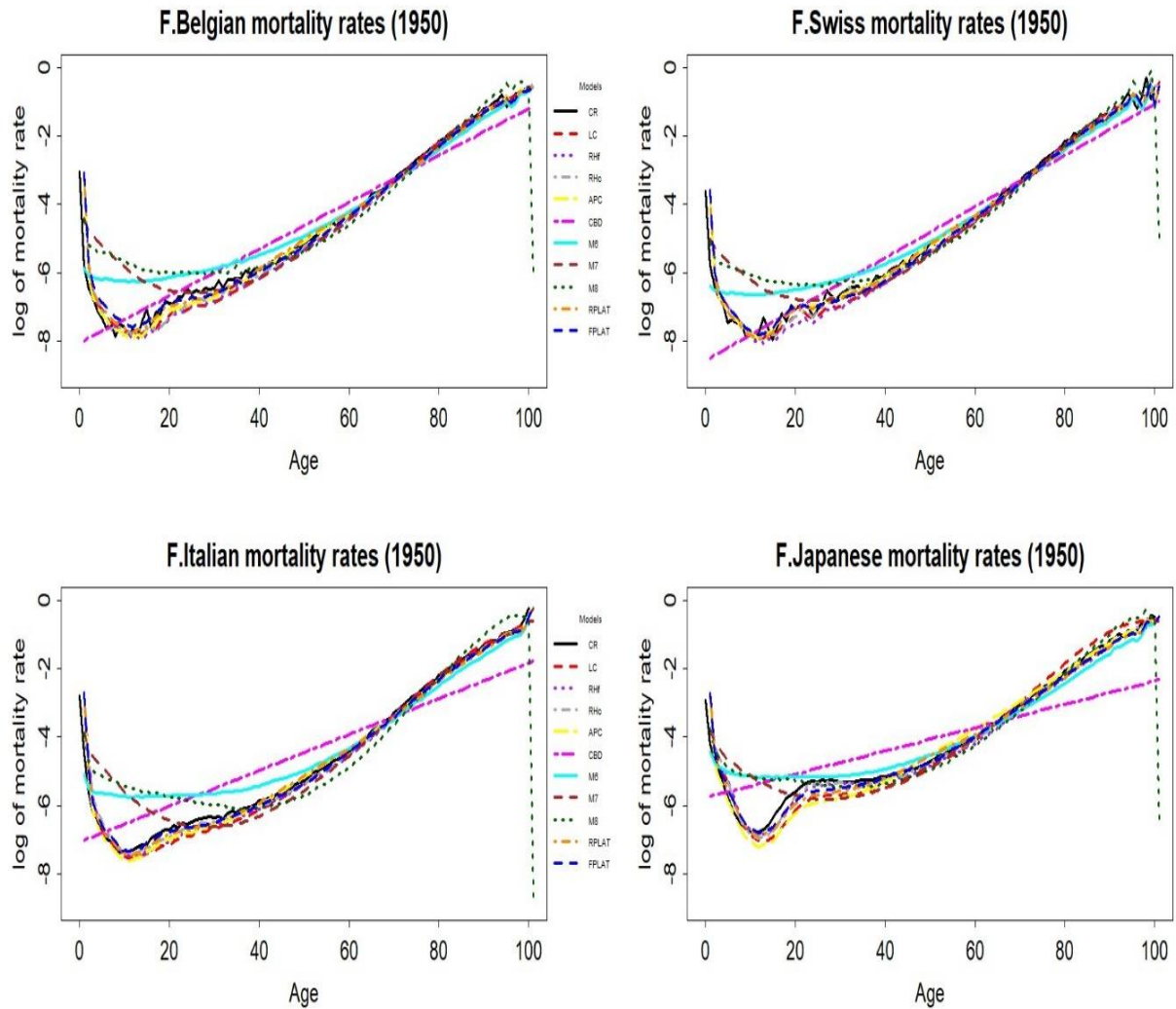


Figure 5 Ten parametric stochastic mortality models were fitted to the female gender using data from the datasets of four countries: Belgium, Switzerland, Italy, and Japan, all from the year 1950.

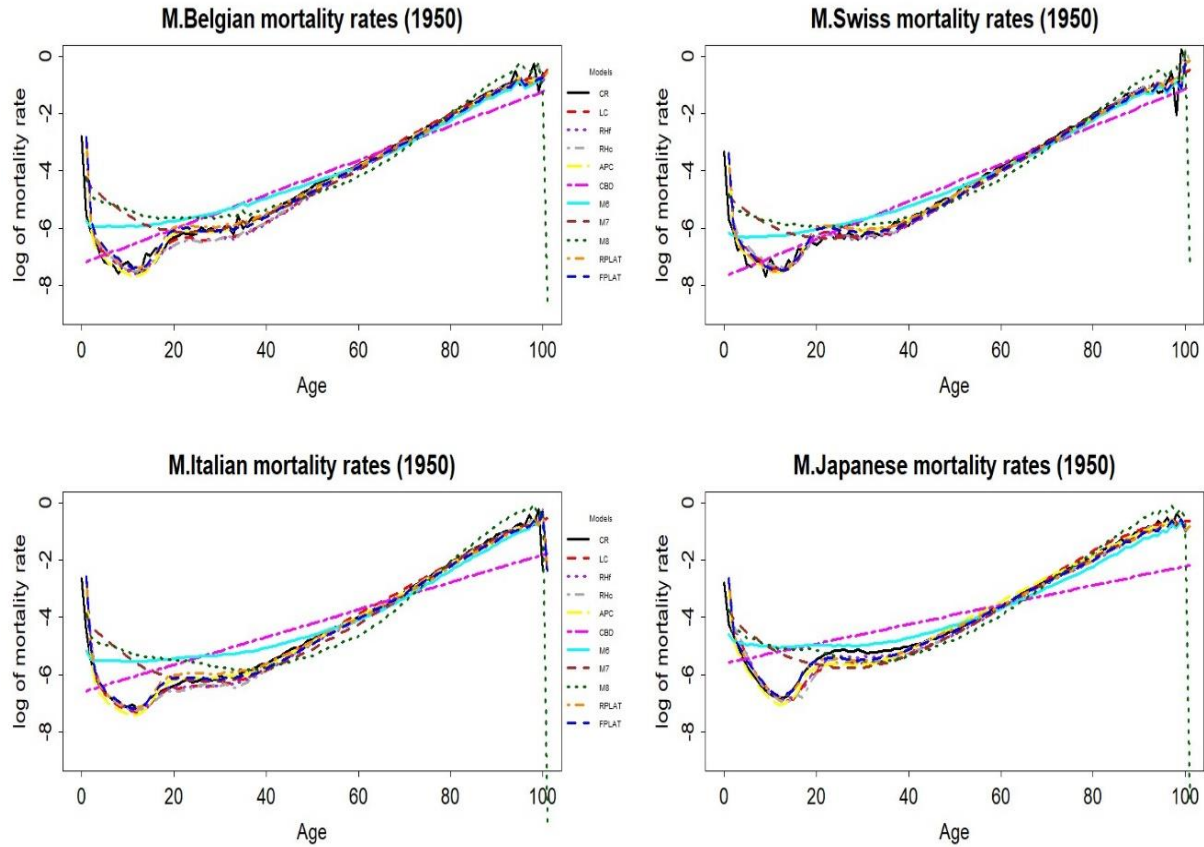


Figure 6 Ten parametric stochastic mortality models were fitted to the male gender using data from the datasets of four countries: Belgium, Switzerland, Italy, and Japan, all from the year 1950.

Tables 6, 7, and 8 below display the metric results after the application of machine learning models used to enhance the three stochastic models. Among these machine learning models; Random Forest displayed the most significant improvement.

Table 6 Comparison of Stochastic mortality models with and without Machine Learning techniques, using Belgian mortality data.

Models	Genders	No ML			DT-ML		
		RMSE	MAE	R ²	RMSE	MAE	R ²
RHf	Female	25.2131	15.9543	0.9986	23.6559	14.5670	0.9988
	Male	27.7890	18.1893	0.9981	26.7422	15.7422	0.9983
RHc	Female	24.8522	15.6127	0.9987	23.6878	14.3039	0.9988
	Male	26.6875	18.7289	0.9976	26.5086	16.1612	0.9982
FPLAT	Female	24.9206	15.9006	0.9987	23.5874	14.4796	0.9988
	Male	26.4800	17.4507	0.9982	25.6596	16.7645	0.9985

Table 7 Comparison of Stochastic mortality models with and without Machine Learning techniques, using Belgian mortality data.

Models	Genders	No ML			RF-ML		
		RMSE	MAE	R ²	RMSE	MAE	R ²
RHf	Female	25.2131	15.9543	0.9986	12.0084	7.3834	0.9997
	Male	27.7890	18.1893	0.9981	14.9865	9.1499	0.9994
RHc	Female	24.8522	15.6127	0.9987	18.1214	10.6870	0.9993
	Male	26.6875	18.7289	0.9976	15.1568	9.0838	0.9994
FPLAT	Female	24.9206	15.9006	0.9987	12.7053	7.7686	0.9997
	Male	26.4800	17.4507	0.9982	15.1821	9.6249	0.9994

Before incorporating machine learning, the RHc model appeared to perform better with female mortality datasets, while the FPLAT model seemed more suitable for male datasets. However, after applying machine learning techniques, we observed a reduction in RMSE and MAE values, as well as an increase in the R² value across all three models. Specifically, a significant decrease in both RMSE and MAE, along with a substantial increase in the R² value, was noticed for the Random Forest model. The RHf model exhibited the overall best fit, with the lowest RMSE and MAE values and the highest R² value after employing machine learning, ultimately surpassing both the RHc and FPLAT models.

Table 8 Comparison of Stochastic mortality models with and without Machine Learning techniques, using Belgian mortality data.

Models	Genders	No ML			GB-ML		
		RMSE	MAE	R ²	RMSE	MAE	R ²
RHf	Female	25.2131	15.9543	0.9986	22.4311	14.9897	0.9989
	Male	27.7890	18.1893	0.9981	24.1227	15.4068	0.9985
RHc	Female	24.8522	15.6127	0.9987	23.5356	14.3243	0.9988
	Male	26.6875	18.7289	0.9976	25.1203	16.2813	0.9983
FPLAT	Female	24.9206	15.9006	0.9987	23.7944	14.4053	0.9988
	Male	26.4800	17.4507	0.9982	23.9520	15.2280	0.9986

CHAPTER 4

ANALYZING MATERNAL MORTALITY WITH A ROBUST PARAMETRIC STOCHASTIC MORTALITY MODEL.

We have compared ten stochastic mortality models in the previous chapter. So, this chapter is dedicated to the research objective which is to employ the best stochastic mortality model(s) identified in objective a) for maternal mortality datasets.

4.1 Introduction

The maternal mortality rate represents the proportion of maternal deaths occurring within a specific period relative to the total person-years lived by women aged 15 to 49 in a population during that period [116]. Maternal mortality has garnered greater attention as a significant global health concern in recent decades, notably through its inclusion as a vital indicator in various international conferences. This inclusion has led to improved data availability over the past ten years. Different methods for measuring maternal mortality are in use globally [117]. In some countries where accurate mortality and fertility records are lacking, census questionnaires are commonly utilized. These questionnaires typically include one or more inquiries aimed at identifying deaths among women of reproductive age who were pregnant or within six weeks of delivery at the time of their death, providing valuable insights into maternal mortality rates. In 2020, an estimated 9.8% of deaths among women aged 15 to 49 were attributed to maternal causes, with Nigeria projected to have the highest global maternal mortality rate [118].

4.2 Methodology

Here, we employ the models that best fit historical datasets of Belgium as discussed in Chapter Three, that is RHc model, to analyze maternal mortality among women of reproductive age (15 to 49) during the period of 1950.

4.2.1 Maternal Mortality

Maternal fatalities counts refer to the instances of mothers' deaths within a particular population within a defined timeframe, often spanning one year [118].

The maternal mortality ratio (MMR) is expressed as the count of maternal deaths within a specified timeframe per 100,000 live births during that period. Consequently, it measures the likelihood of maternal death in relation to the total number of live births. It can be calculated using,

$$MMRatio = \frac{\text{Number of maternal deaths}}{\text{Number of live births}} \quad (4.1)$$

Maternal mortality Rates ((MMRate) refers to the death of a woman due to causes related to or exacerbated by pregnancy or its management. This includes deaths occurring during pregnancy or within 42 days after its termination, excluding unintentional or incidental causes, irrespective of the pregnancy's duration or location. It is determined by dividing the number of maternal deaths by the person-years lived by women of reproductive age within a population. This metric reflects both the likelihood of maternal death per pregnancy and the population's fertility rate. The formula for calculating MMRate is given as follows:

$$MMRate = \frac{\text{Number of maternal deaths}}{\text{Number of women of reproductive age 15 – 49 years}} \quad (4.2)$$

4.2.2 Stochastic Mortality Models

The three outstanding models from chapter three, RHc, RHf and FPLAT models which properly fitted the historical datasets of both genders of the four countries used are applied to historical maternal mortality data of Belgium. The mathematical expressions of these models are given respective as follows:

$$\ln\varphi_{a,p} = \kappa_a^{(1)} + \kappa_a^{(2)}\rho_p + \kappa_a^{(3)}\delta_c + \varepsilon_{a,p}, \quad (4.1)$$

in the second version, $\kappa_a^{(3)} = 1$.

$$\ln\varphi_{a,p} = \kappa_a^{(1)} + \kappa_a^{(2)}\rho_p + \delta_c + \varepsilon_{a,p}. \quad (4.2)$$

With the identifiability constraints,

$$\sum_a \kappa_a^{(2)} = 1, \sum_a \kappa_a^{(3)} = 1, \sum_p \rho_p = 0, \text{ and } \sum_c \delta_c = 0$$

$$\ln\varphi_{a,p} = \kappa_a + \rho_p^{(1)} + \rho_p^{(2)}(\bar{a} - a) + \rho_p^{(3)}a^* + \delta_c + \varepsilon_{a,p}, \quad (4.3)$$

where $a^* = \max[(\bar{a} - a), 0]$.

4.2.3 Goodness of fit metric

In addition to the model performance metrics discussed in Chapters One and Three, we have incorporated another metric, Mean Absolute Percentage Error (MAPE), to evaluate the goodness of fit of the stochastic mortality models used to fit maternal mortality data in Belgium. MAPE assesses the absolute percentage error of the model by comparing actual maternal mortality rates with the values predicted by the model. Its mathematical expression is given below:

$$MAPE = \frac{1}{n} \sum_{a,p} \left| \frac{M_{a,p} - \hat{M}_{a,p}}{M_{a,p}} \right| \times 100, \quad (3.7)$$

Here, $M_{a,p}$ denotes the actual maternal mortality rates of age group a during period p , while $\hat{M}_{a,p}$ represents the fitted values obtained from a specific parametric mortality model.

4.3 Results and Discussion

In the comparative study of parametric stochastic mortality models presented in Chapter Three, the RHc model emerged as the top performer among ten models in fitting overall Belgian female mortality. It maintained its superiority, ranking among the top three models with the lowest values for goodness-of-fit metrics as shown in the Table 9 below.

Table 9 Belgian Maternal Mortality fits of year 1950.

Models	Metrics			
	RMSE	MAE	MAPE	R-square
RHc	0.6646556	0.4976029	0.119086	0.9799533
RHf	0.6809935	0.5095931	0.1231374	0.9789557
FPLAT	0.6887647	0.5129403	0.1231392	0.9784727

The evaluation in this chapter considered four metrics: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Coefficient of Determination (R^2). Among these, the RHc model demonstrated the best fit to historical maternal deaths in Belgium, exhibiting the lowest values for RMSE, MAE, and MAPE, and the highest value for R^2 . Figure 7 below presents the

plot of observed maternal mortality rates and the fit of the three stochastic mortality models.

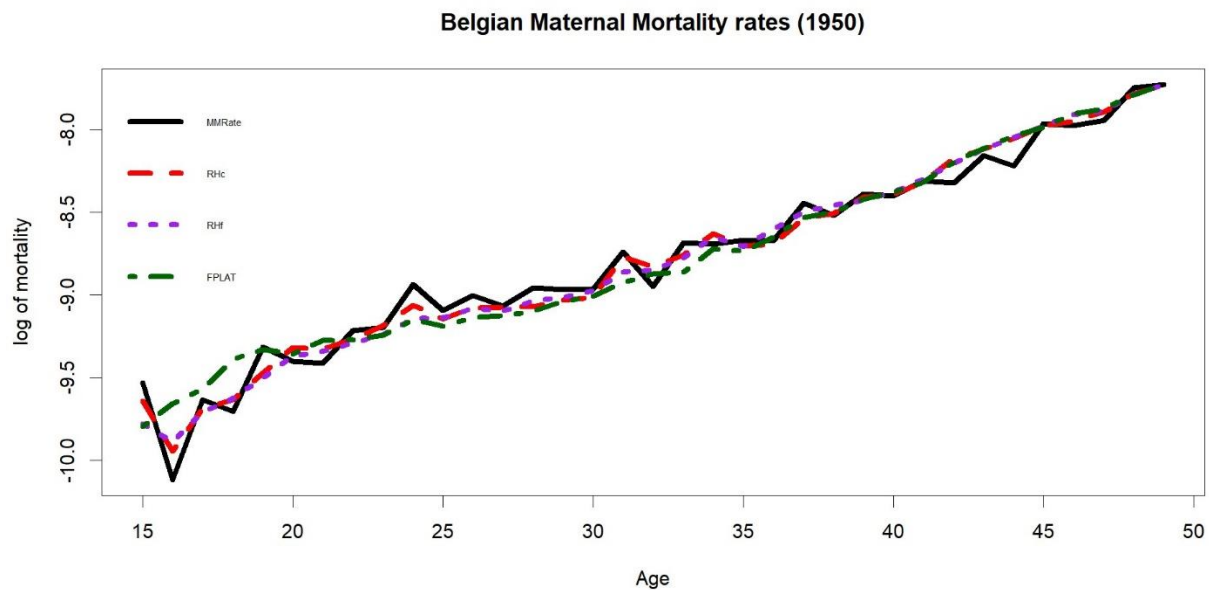


Figure 7 The RHc, RHf, and FPLAT parametric stochastic models fit of Belgian historical maternal mortality rates for the year 1950.

CHAPTER 5

EXAMINING AGE-GENDER EFFECT ON LEE-CARTER MODEL USING GNM FRAMEWORK.

This chapter delves into the study's fourth objective, which aims to propose a stochastic mortality model integrating gender-age effects. We examine the inclusion of gender as a factor in the stochastic mortality model, employing genders as dummy variables.

Different extensions to the Lee Carter model have been explored in the literature, such as Renshaw Haberman model, which consider cohort of the individual to account for the effect of the year of birth. However, these extensions typically overlook gender as a factor in mortality modelling and often fit separate models for each gender. In this chapter, we introduce gender as a factor using dummy variables. Initially, we stack the dataset for both genders, assigning zero to females and one to males. The model is basically fitted without considering gender for the stacked dataset, and subsequently, age-gender effects are incorporated into the model. We also set the age-modulating parameters to a unit in a separate case and compare all scenarios. Our observations indicate that the model with age-gender effects outperforms the other two scenarios.

To explore the impact of gender, we applied the LC model and conducted an analysis using Belgium's historical dataset. We utilized the Generalized Nonlinear Model (GNM) package to fit the proposed extended version of the LC stochastic model that incorporates gender as a factor. The choice of this package stems from its flexible

framework, allowing for the specification and estimation of models with nonlinear terms. This approach provides a comprehensive examination of gender effects in stochastic mortality modeling.

5.1 Introduction

Mortality rate is the measurement of the number of deaths in each population. Human mortality has recently experienced a significant and unequal reduction globally, compared to the historical records, despite international conflicts and other catastrophes [32], [119]. In recent years, there has been a notable surge in research on mortality change, driven by the collective efforts of economists, demographers, health policy makers, and other stakeholders. The increasing interest in this area stems from the quest to identify the mortality models that most accurately capture and explain mortality rates. Traditionally, the longevity of lives of individuals is understood as a good thing and a significant social advancement. However, this has posed a great worry to some organizations as the consequence may affect the distribution of resources in areas like welfare, health care, and pensions. Continuous increases in life expectancy in developed countries since 1970 have primarily been driven by a decrease in mortality rates among older people [33].

Extrapolating from past trends appears to be the most dependable approach to anticipate the future, given the intricacy and historical stability of these changes [33]. Contrarily, [32], [33], and [34] employed the LC model with Binomial, Poisson, and Negative binomial frameworks respectively to estimate the parameters using the maximum likelihood method, allowing for the evaluation of the model's parameter accuracy. [120] employs a first-principles analysis of the Lee-Carter model to

Australian population data, providing insights into potential future possibilities for actuarial analysis.

The subsequent sections of the paper are organized as follows: Section 2 provides a literature review of previous works. Section 3 highlights the methodology of the study. Section 4 discusses the results, and in section 5, the conclusion of the study is deliberated.

5.2 Literature Review

In the literature, various stochastic mortality models have been applied to historical datasets from different countries, encompassing both developing and developed nations. These models utilize either the generalized (non) linear package [59] or the Stochastics Mortality Model (StMoMo) package [27] in R software. These two packages are employed to fit mortality models for each gender independently. Notably, StMoMo is specifically designed to fit and forecast mortality model for a gender at a time.

The fundamental Lee-Carter method is elucidated in [2], containing its forecasts. The restrictions of the approach are deliberated, alongside its recent versions, uses, and methodological advancements. A concise overview of the method's incorporation into broader population predictions and forecasts of the US Social Security system's finances is also presented. The recent study of [35] demonstrates the methodology and conducts an ex-post evaluation of the original Lee-Carter projection for U.S. mortality, which was performed three decades ago. The authors discuss prominent extensions of the Lee-Carter method, addressing specific issues and providing an analysis of previous assessments and comparisons. They focused on evaluating the

effectiveness of different expansions. An extensive analysis of the link between the trend in economic growth as measured by the GDP and the latent trend in mortality dynamics was demonstrated by [39]. The Lee-Carter framework is then expanded by adding GDP as a second component next to the latent factor, which improves forecast interpretability and fit.

The original Lee-Carter model was extended and employed in the work by [38] to forecast death rates by using a new class of time series models called Generalized Autoregressive Score (GAS) or Dynamic Conditional Score (DCS) models. In essence, the study suggests five probability models—Poisson, binomial, negative binomial, Gaussian, and beta—that are based on the GAS framework and are used to calibrate the Lee-Carter parameters in a single, integrated step to dynamically forecast the mortality rates. The study concluded that, of the probability models, the LC model's negative binomial extension is the most effective at predicting death rates. The works of [121] and [9] emphasized the general characteristics and the potential to explain historical mortality patterns. These criteria encompass the robustness of parameter estimates concerning the data period, convenient to implement, simplicity, parsimony, inclusion of year-of-birth effects (at least for some countries), ability to generate meaningful age correlations, and fit quality. Furthermore, Plat (2009) stated an additional criterion that the model should be applicable to a complete age range.

In the realm of fitting and forecasting death, generalized linear models (GLMs) and generalized nonlinear models (GNMs), efficiently express stochastic mortality models. In a comprehensive study by [62], eight distinct mortality model structures

were compared. The R language was employed to delineate these structures, providing algebraic descriptions with the advantage of clarity and simplicity. Among the eight models under consideration, six were fitted using the glm approach, while the remaining two models such as the Lee–Carter and Renshaw–Haberman, employed gnm due to the presence of bilinear terms, that is, the product of two parameters. The gnm package[59] in R software permits the direct fitting of models with such specifications that is very similar to that of the glm function. The models employed Poisson structures with both log and logit link functions for the force of mortality, as well as binomial structures with logit and complementary log-log link functions for the rate of mortality within each structure. Furthermore, the study delved into the feasibility of forecasting using these models. Notably, the introduction of cohort effects was found to enhance overall model fit. However, it was acknowledged that incorporating cohort effects could pose challenges in forecasting with these models. The challenge of calibrating mortality and life expectancy in the event of a structural change is considered in the study conducted by [3]. The study displays how, within the framework of the Lee–Carter model, new improvements in statistical testing for structural changes may be used to derive adequately described models for the overall mortality rate. Specifically, the results of tests for a shift in the mortality index's trend and the existence of a unit root are utilized to determine the most suitable forecasting models. Additionally, almost all the countries examined had structurally different rates of fall in overall mortality, particularly in male populations, according to the study. Lastly, their research shows how to account for this kind of shift.

The work of [37] discusses on comparison of simulation methodologies for evaluating risk in death rate projections and associated values of life expectancy and annuity values in both period and cohort frameworks. The work of [40] also highlights various hidden characteristics of the LC death predicting method, demonstrating that it is a special case of a simpler and less frequently biased random walk with drift model. Furthermore, they establish that the age profile forecast from both approaches, whether predicting forward or backward in time, will eventually deviate from any given baseline and become less smooth and unrealistic after a certain point. Identifiability issues of the Lee-Carter model and its variants is discussed by [28], adopting a perspective referred to as the plug-in age-period or plug-in age-period-cohort model, to view the latent variable as a stochastic process from the outset. The study investigates the identifiability of expected values and covariances within the plug-in Lee-Carter models and defines a parameter vector that incorporates the underlying parameters of this process rather than its realizations. Similarly, the research of [122] employed mortality models to understand the relationship between death rates and population age using historical datasets.

In the study by [54], a new method for modeling subnational mortality is proposed. This method uses individual factors such as sex, geographical locations, and socioeconomic status, instead of focusing directly on subpopulations. The approach is considered parsimonious as it requires fewer parameters to explain mortality.

The research presents a broad Bayesian estimation method that is effective with many age-period-cohort mortality models suggested in existing literature. It utilizes Markov chain Monte Carlo methods by integrating deterministic filtering with

adaptive Metropolis steps, making it suitable for complex cases. Ultimately, the study shows that this approach, when paired with an appropriate mortality model, produces sensible, accurate, and consistent mortality predictions that align well with the data. In a study conducted by [55], a model proposed based on the mortality surface's breakdown into a polynomial foundation. The research demonstrates the application of cross-validation and regularization techniques to produce a parsimonious and logical predictive model for mortality forecasting. Furthermore, the study assesses the potential impact of COVID-19-like effects on predictions made using both classical and employed approaches. Essentially, predictions of mortality rates are generally more stable than those based on a cohort effect, and regularized models perform better in terms of prediction and stability than the so-called P-spline model. However, all the reviewed articles have fitted stochastic mortality models for either male or female gender or both genders separately. This article aims to fit the mortality rates for both genders collectively using our postulated extension of the original Lee-Carter methodology. Traditionally, the stochastic mortality models are fitted separately for male and female historical datasets, resulting in an equal number of parameters estimates for each gender. However, one of the key criteria for evaluating the quality of a model is parsimony, favoring accurate models with fewer estimated parameters. To address this, we propose a more parsimonious model that incorporates gender effect as a factor, compared to fitting the original Lee-Carter model separately for the two genders. Achieving accuracy with a reduced number of parameters is highly desirable.

5.3 Methodology

In this section, we will discuss the methodology used in this paper.

5.3.1 Assumptions

In line with the previous literature, we assumed throughout this study, that the mortality rate is the same within age groups and time but varies among different age groups and period. The death counts are independent for different individual features considered, such as age, gender, and year of birth. Similarly, we supposed the death counts D_{at} , a random variable with the fixed central exposure E_{at} follows Poisson distribution with a parameter proportional to exposures [45], [58], [68].

$$D_{at} \sim \text{Pois}(\theta_{at} E_{at}) \quad (5.1)$$

Additionally, we assumed independent and identically distributed random error terms with equal variances [40]; [105]. However, the assumption of constant variances of the error term is not always true [34].

Table 10 Description of variables/parameters

Variables/Parameters	Descriptions
D_{at}	deaths count of age group a at period t .
d_{at}	The realization of the random variable D_{at}
E_{at}	Fixed number of lives in age group a at period t that exposed to risk
θ_{at}	Observed mortality rate of age group a during calendar year t , with the estimated value, $\hat{\theta}_{at}$.
θ_{atg}	Observed mortality rate of age group a during calendar year t for gender g .
$\beta_a^{(1)}$	Average of log of mortality rates of age group a (fitted value)
$\beta_a^{(2)}$	Changes in age-specific mortality patterns

$\beta_a^{(3)}$	Sensitivity of mortality rates at age a to the observable factor, gender
γ_t	Random period parameter
μ_g	Gender parameter
ε_{at}	Random error term for the original Lee Carter
ε_{atg}	Random error term for the extended Lee Carter

5.3.2 Generalized non-linear models.

The generalized nonlinear models (gnm) are regression models in which the link-transformed mean is expressed as a sum of predictor terms, some of which may be non-linear in the unknown parameters. The gnm package in R provides facilities for fitting generalized nonlinear models. Central to the functionality offered by the gnm package is the model-fitting function `gnm`, which interprets a model formula and returns a model object. The user interface of `gnm` is patterned after `glm`, and indeed, `gnm` can be seen as a replacement for `glm` in specifying and fitting generalized nonlinear models. There is no reason to replace `glm` with `gnm` unless the model includes many incidental parameters treatable by `gnm`'s elimination mechanism. The primary purpose of the `gnm` package is to expand the features of models to incorporate nonlinear terms, providing a flexible framework for designing and estimating generalized models with nonlinear elements [62]. The `gnm` function is specified as:

```
gnm(Death ~ -1 + factor(Age)+ Mult(factor(Age), factor(Year)) + offset(logExp),
family = poisson(link = "log")),
```

where $\log\text{Exp} = \log(\text{Central Exposure})$ and function `Mult(·)` is used to specify nonlinearity in the model, that is the product of predictors is specified by the function.

Specifically, each of the features, age and year are expressed as a factor, using the function factor in R.

The feature in gnm for specifying nonlinear terms is made to work with the symbolic language that formula objects employ. In the model formula, nonlinear terms are basically defined as calls to functions of the class “nonlin.” The gnm package entails some nonlinear functions and allows users to define their own appropriate nonlinear functions as well. For further details on GNM, the reader should refer to [59].

5.3.3 Goodness of fit

The performance of the stochastic mortality model in fitting and predicting deaths will be examined using three metrics: Mean Absolute Percentage Error (MAPE), Root mean square error (RMSE), coefficient of determination (R^2) and mean absolute error (MAE). Lower values of Δ , RMSE, and MAE metrics signify a better fit of the model whereas the higher value of R^2 tells the model with best fit. Poisson Deviance, a usual metric in Generalized Nonlinear Model (GNM), is employed. This is because the form of many mortality models, such as the bilinear term $\beta_a^{(2)} \gamma_t$ in the Lee-Carter model, necessitates the use of gnm package. Therefore, the gnm command in the R-package gnm will be utilized for its compatibility with models featuring such structures.

MAPE assesses the absolute percentage error of the model by comparing actual maternal mortality rates with the values predicted by the model. Its formula is defined as:

$$MAPE = \frac{1}{n} \sum_{a,p} \left| \frac{d_{at} - \hat{d}_{at}}{d_{at}} \right| \times 100 \quad (5.2)$$

RMSE measures the average distance between the observed values and the predicted values of deaths counts. The model's overall performance is represented by a single flexible value of RMSE. It is defined as:

$$RMSE = \sqrt{\frac{\sum_{at} (d_{at} - \hat{d}_{at})^2}{n}}, \quad (5.3)$$

where n is the total number of observations.

The coefficient of determination, R^2 measures the proportion of the variation in the death counts that is explained by the features, such as age, gender, calendar year, and year of birth.

$$R^2 = 1 - \frac{\sum_{at} (d_{at} - \hat{d}_{at})^2}{\sum_{at} (d_{at} - \bar{d}_{at})^2}, \quad (5.4)$$

where $\bar{d}_{at}$ is the average of the observed death counts.

MAE describes the accuracy of the model by expressing the absolute difference between the fitted values and actual values of the deaths counts. Its mathematical expression is given as:

$$MAE = \frac{1}{n} \sum_a |d_{at} - \hat{d}_{at}| \quad (5.5)$$

5.3.4 Original Lee-Carter Model

The LC model, developed by [1], stands as a pioneering stochastic mortality model designed to extrapolate trends and age patterns in mortality. It is a widely

implemented model of its kind. The LC model offers a simple and empirical explanation for changes in mortality over time [2].

The LC model is characterized by two time-invariant age parameters, denotes as $\beta_a^{(1)}$ and $\beta_a^{(2)}$, along with a random period parameter γ_t . Furthermore, $\beta_a^{(2)}$ is assumed to be fixed over time for all age group, a (i.e., it does not vary over time) and γ_t is also assumed to be constant over age groups for all periods, denoted as t [40]. This assumption is referred to as the independent assumption for the logarithm of the mortality rate.

In addition, the model incorporates a random homoscedasticity error term ε_{at} , assuming a mean of 0 and a variance of σ_ε^2 [40]; [105]; [39]. It is defined as:

$$\theta_{at} = \log^{-1}(\beta_a^{(1)} + \beta_a^{(2)} \gamma_t + \varepsilon_{at}), \quad (5.6)$$

For identifiability issue, LC model attracts the following parameter constraints [3],

$$\sum_a \beta_a^{(2)} = 1 \text{ and } \sum_t \gamma_t = 0. \quad (5.7)$$

From equation (1) above, the maximum likelihood estimate of θ_{at} is obtained as:

$$\hat{\theta}_{at} = \frac{\text{Death counts of age group } a \text{ during period } t.}{\text{exposure of age group } a \text{ during period } t.} = \frac{d_{at}}{E_{at}}, \quad (5.8)$$

where d_{at} is the observed values of the random variables D_{at} . The prediction of γ_t and estimated parameters $\hat{\beta}_a^{(1)}$ and $\hat{\beta}_a^{(2)}$ play a prominent role in forecasting mortality.

The LC model is flexible and parsimonious. Several versions of the Lee Carter model have been discussed in the literature; for example, the cohort extension of Lee Carter

called Renshaw-Haberman stochastic model [52]. Nevertheless, most of these editions of Lee-Carter model attract more parameters estimation.

5.3.5 Extended Lee- Carter Model

In this section, our analysis is based on the historical mortality dataset of Belgium, a country with a complete record of death counts for the year and age group considered. Belgium's mortality dynamics are often assessed in the research. The original LC model fits the two genders combinedly, using the function `gnm` in the `gnm` package. All the goodness of fit in Eqs. (2), (3), (4), and (5) are measured, and the results are displayed in Table 2.0.

Following the proposed generalized form of the LC model by [39], we formulate our extended Lee-Carter model as:

$$\ln\theta_{atg} = \beta_a^{(1)} + \beta_a^{(2)}\gamma_t + \beta_a^{(3)}\mu_g + \varepsilon_{atg}, \quad (5.9)$$

where γ_t is time varying latent variable, $\beta_a^{(1)}$, $\beta_a^{(2)}$, and $\beta_a^{(3)}$ are the invariant age parameters with the same interpretation as in original Lee-Carter model. μ_g is gender parameter. Parameter γ_t has similar constraints as in original Lee-Carter and parameters $\beta_a^{(2)}$, and $\beta_a^{(3)}$ share the same constraints like $\beta_a^{(2)}$ in original Lee-Carter model. The error term ε_{atg} has no correlation with γ_t and μ_g . From equation (9), we can obtain the original Lee-Carter model by letting $\beta_a^{(3)} = 0$.

Similarly, for female, that is, if the gender is zero, eq.(9) reduced to original Lee Carter model and for the male gender, that is when gender is one, the model is obtained as:

$$\ln\theta_{atg} = \beta_a^* + \beta_a^{(2)}\gamma_t + \varepsilon_{atg}, \quad (5.10)$$

where $\beta_a^* = \beta_a^{(1)} + \beta_a^{(3)}$.

5.4 Historical Dataset

The historical mortality data, sourced from the Human Mortality Database [111], is organized in rectangular arrays containing death counts, D_{at} and central exposure E_{at} for male, female, and total separately, for each country. The arrangement is suitable for the StMoMo package [27], where rows represent age groups a_i ($i = 1, 2, \dots, k$), and a_k is the maximum age, an individual can attain). The columns of the rectangular array represent calendar years p_h ($h = 1, 2, \dots, H$), and p_H denotes the maximum calendar year considered). For our study, we restructured the Belgian historical mortality dataset by creating different columns for each feature or variable, including age, gender, calendar year, and year of birth or cohort. Subsequently, we stacked the dataset for both male and female genders. This reformation facilitates compatibility with the gnm package[26], which is a key tool in our analysis. The year 1950 to 2020 and ages ranging from 0 to 100 inclusively are considered. The total number of observations, 14342, is obtained by the product of one hundred and one age groups, the seventy-one different calendar year and the two genders considered.

5.5 Results and discussion

In alignment with the previous studies, we initiated our analysis by applying the original Lee Carter model separately to male and female genders, utilizing the

Belgian historical dataset, followed by a comparative assessment. The parameters of the original Lee-Carter model were estimated under a Poisson setting [71] with a log link function. The estimations were performed using the StMoMo package in R software [27].

Table 11 Performance metrics for original Lee Carter fitted to each gender of Belgian historical dataset.

Metrics	Male	Female	Male + Female
Number of Parameters	273	273	546
MAPE	0.1118	0.1166	0.2284
RMSE	61.97534	36.39974	98.37508
R^2	0.9904	0.9972	-
MAE	32.20389	20.97022	53.17411

Table 11 above presents the outcomes of the original Lee Carter model applied to both genders separately. The results exhibit an equal number of parameters estimates for both genders. However, the male gender displays smaller values for MAPE whereas RMSE and MAE for the female give smaller values compared to values for male gender, indicating a superior model fit compared to the male gender. The analysis encompassed the age range of 0 to 100 and the calendar years from 1950 to 2020, totaling 71years for both genders. Notably, the model explained 99.72% of variations in female mortality rates and 99.04% of variations in male death rates. Estimating the two genders separately appears time-consuming, redundant, and uneconomical.

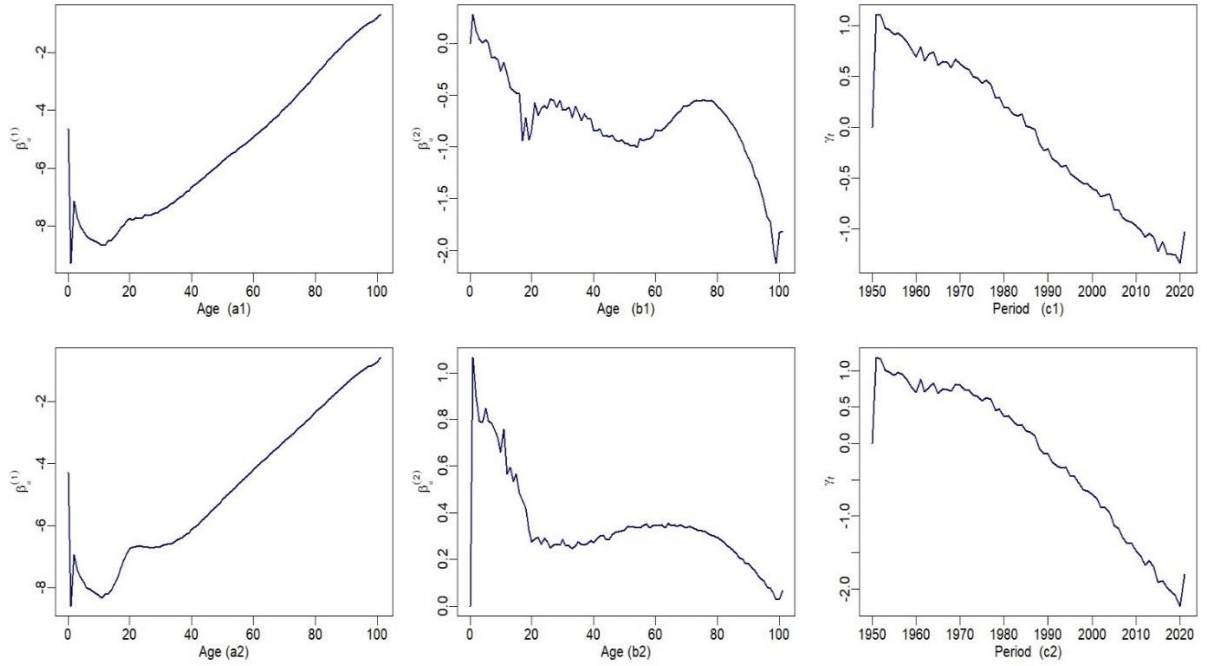


Figure 8 Original LC model parameter values plotted for datasets of both genders of Belgian historical mortality data.

Figure 8 illustrates the parameter values of the original LC model for the male and female Belgian historical datasets. The top three graphs (a1, b1, and c1) represent the three parameters of LC model estimated for the female gender. Conversely, the other three graphs (a2, b3, and c2) depict the three parameters of the LC model for the male gender. The average level of mortality at each age group, $\beta_a^{(1)}$ which denotes the general shape of the logarithm of the age specific death rates, exhibits similar trends for both genders. Although the hump around the age of 20 for males is sharper compared to the one around the same age for females.

The changes in age-specific death patterns over the years, $\beta_a^{(2)}$ vary between male and female genders. The sensitivity of ages to the change in mortality over time is initially higher in males compared to females. Male mortality rates show a quick

reduction, while female mortality remain relatively stable from around age 20 to age 80, after which a significant decline occurs. Lastly, the random period parameter for both genders also demonstrate a downward trend in mortality after a sudden increase from 1950 to around 1993, followed by a continuous decline. However, in males, there are fluctuations observed from around 1960 to 1970.

The second analysis of the study focused on fitting stochastic mortality model to a historical dataset for both genders combined. We then compared the previous approach with three different cases of the proposed method. We stacked male and female mortality datasets from Belgium and used them to calibrate the parameters of the extended LC model for three scenarios: $\beta_a^{(3)} = 0$, $\beta_a^{(3)} = 1$, and $0 \neq \beta_a^{(3)} \neq 1$.

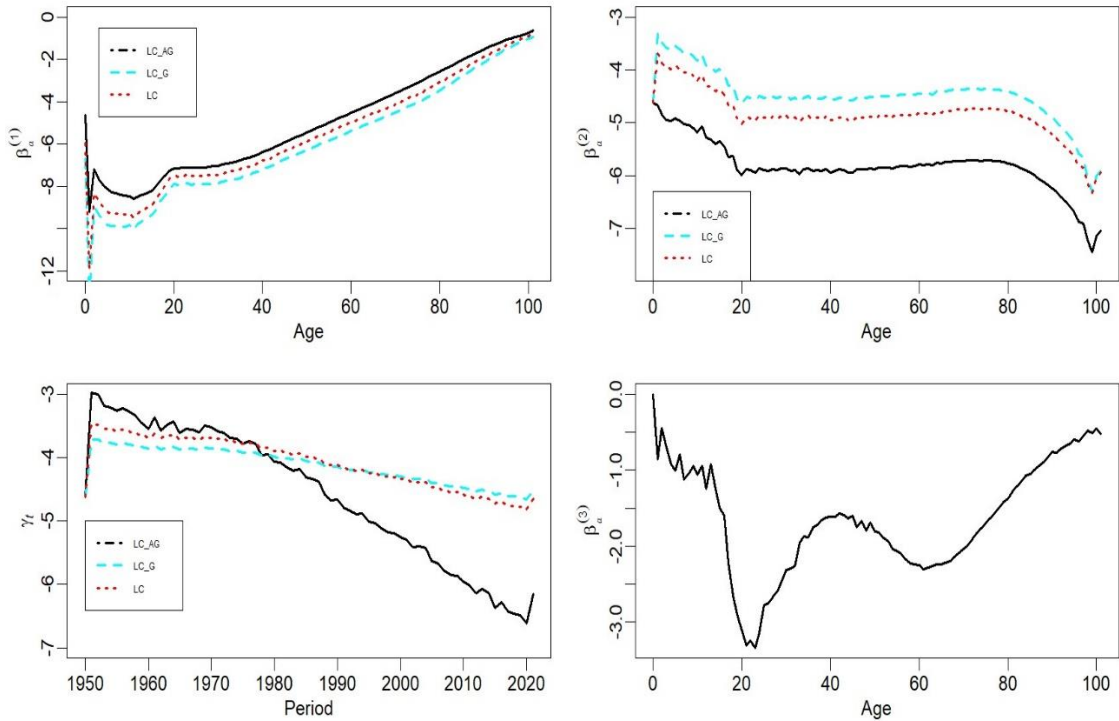


Figure 9 Parameter values for the LC model with age-gender effect, gender effect only, and no gender effect, applied to the combined male and female datasets from Belgian historical records.

The results for the combined genders are displayed in Figure 9 below for the three cases. All the three parameters in each case demonstrate the same patterns except for the period effect parameter. In Table 12, the performance metric for the extended Lee Carter model are reported. The fourth figure for the age effect across the gender, $\beta_a^{(3)}$, is obtained only for the case of age-gender effect. A very sharp hump is observed around the age of 20 and another at around the age of 60, followed by a drastic fall in mortality. Unlike in separate fit of mortality models for genders, the study observed a similar pattern in parameter $\beta_a^{(2)}$, representing the responsiveness of ages to the variations in mortality over years, for the three scenarios.

Table 12 Performance metrics for extended Lee Carter fitted to combined genders of Belgian historical mortality dataset.

Models	No of Par.	Metrics			
		MAPE	RMSE	R^2	MAE
LC_AG	376	0.1242	66.19643	0.9899	35.5024
LC_G	274	0.1700	87.08432	0.9826	51.4514
LC	273	0.3142	198.9053	0.9092	124.1841

In the last column of Table 2, the number of estimated parameters and the metric values for both males and females are summed together. The three cases studied are more economical in terms of the number of parameters required for estimation compared to fitting the models separately for each gender. Specifically, the Lee-Carter Age-Gender effect (LC_AG) model and Lee-Carter Gender effect (LC_G) outperform separate fits, based on lower values of MAPE, RMSE, and MAE. The ordinary Lee-Carter model fitted for stacked datasets of both male and female genders performed poorly among the three cases and compared to separate fits for each gender. Notably, the extended Lee-Carter (LC_AG) model demonstrates the

highest performance among the three cases as well as when compared to the addition of separate fits for genders. The LC_AG model explains 98.99% of the variation in male and female mortality. The LC_G model follows, explaining 98.26% of the variation in mortality.

CHAPTER 6

CONCLUSION

This chapter focuses on the discussion of summary and conclusion of the whole study. The study aimed to conduct analysis that compared death models enhanced by machine learning techniques. It sought to fit maternal mortality datasets using the identified best stochastic mortality model and propose a model incorporating gender as a factor. This involved examining ten prevalent stochastic mortality models, including LC, RHf, RHc, APC, CBD (and its variants M6, M7, M8), and PLAT (both FPLAT and RPLAT), using historical datasets from four developed countries: Switzerland, Italy, Belgium, and Japan.

In chapter three, the study established a comparative study of the models, considering age groups from 0 to 100 years and the period from 1950 to 2020. Among these models, RHc, RHf, and FPLAT demonstrated exceptional performance across all countries and genders. Furthermore, in the same chapter three, the selected models were enhanced by applying machine learning techniques. The improvement in the performance of the three models was observed through reductions in RMSE and MAE values and an increase in the value of R^2 . Initially, the RHc model outperformed others for Belgian female mortality data. However, after incorporating machine learning, the FPLAT model emerged as the overall best model for Belgium's historical mortality data, showcasing the importance of machine learning in identifying model strengths and weaknesses and aiding in model selection.

In chapter four, the three exceptional models selected from chapter three were applied to fitting maternal mortality of Belgium, with the RHc model maintaining the highest performance.

Subsequently, in chapter five, a study examining the incorporation of gender as a factor in the LC model was established. This involved three cases: Lee-Carter-Age-Gender (LC-AG), Lee-Carter-Gender (LC-G), and the ordinary Lee-Carter model for stacked data. The LC model fitted for the stacked datasets of male and female genders, but ignoring gender effects, exhibited the least performance among the three cases. Conversely, the LC-AG model, which incorporates gender-age effects, showed the highest performance with the lowest values of MAPE, RMSE, and MAE. The study demonstrated the importance of incorporating gender as a factor when fitting mortality models.

References

- [1] R. D. Lee and L. R. Carter, “Modeling and Forecasting U. S. Mortality,” *J Am Stat Assoc*, vol. 87, no. 419, p. 659, 1992, doi: 10.2307/2290201.
- [2] R. Lee, “The lee-carter method for forecasting mortality, with various extensions and applications,” *North American Actuarial Journal*, vol. 4, no. 1, pp. 80–91, 2000, doi: 10.1080/10920277.2000.10595882.
- [3] E. Coelho and L. C. Nunes, “Forecasting mortality in the event of a structural change,” 2011.
- [4] A. E. Renshaw and S. Haberman, “A cohort-based extension to the Lee-Carter model for mortality reduction factors,” *Insur Math Econ*, vol. 38, no. 3, pp. 556–570, 2006, doi: 10.1016/j.insmatheco.2005.12.001.
- [5] I. D. Currie, M. Durban, and P. H. Eilers, “Smoothing and forecasting mortality rates,” *Stat Modelling*, vol. 4, no. 4, pp. 279–298, Dec. 2004, doi: 10.1191/1471082X04st080oa.
- [6] A. Hunt and A. M. Villegas, “Robustness and convergence in the Lee-Carter model with cohort effects,” *Insur Math Econ*, vol. 64, no. December 2017, pp. 186–202, 2015, doi: 10.1016/j.insmatheco.2015.05.004.
- [7] A. J. G. Cairns, D. Blake, K. Dowd, G. D. Coughlan, D. Epstein, and M. Khalaf-Allah, “Mortality density forecasts: An analysis of six stochastic mortality models,” *Insur Math Econ*, vol. 48, no. 3, pp. 355–367, 2011, doi: 10.1016/j.insmatheco.2010.12.005.
- [8] A. J. G. Cairns, D. P. Blake, K. Dowd, G. Coughlan, and D. Epstein, “A Quantitative Comparison of Stochastic Mortality Models Using Data from England & Wales and the United States,” *SSRN Electronic Journal*, Jul. 2012, doi: 10.2139/ssrn.1340389.
- [9] R. Plat, “On stochastic mortality modeling,” *Insur Math Econ*, vol. 45, no. 3, pp. 393–404, Dec. 2009, doi: 10.1016/j.insmatheco.2009.08.006.
- [10] M. Bansal, A. Goyal, and A. Choudhary, “A comparative analysis of K-Nearest Neighbor, Genetic, Support Vector Machine, Decision Tree, and Long Short-Term Memory algorithms in machine learning,” *Decision Analytics Journal*, vol. 3, no. May, p. 100071, 2022, doi: 10.1016/j.dajour.2022.100071.

- [11] S. Levantesi and V. Pizzorusso, “Application of machine learning to mortality modeling and forecasting,” *Risks*, vol. 7, no. 1, Mar. 2019, doi: 10.3390/risks7010026.
- [12] G. James, Daniela Witten, T. Hastie, and R. Tibshirani, *Springer Texts in Statistics An Introduction to Statistical Learning with application in R*. 2013.
- [13] B. Gupta, A. Rawat, A. Jain, A. Arora, and N. Dhimi, “Analysis of Various Decision Tree Algorithms for Classification in Data Mining,” *Int J Comput Appl*, vol. 163, no. 8, pp. 15–19, 2017, doi: 10.5120/ijca2017913660.
- [14] G. Meyfroidt, F. Güiza, J. Ramon, and M. Bruynooghe, “Machine learning techniques to examine large patient databases,” *Best Pract Res Clin Anaesthesiol*, vol. 23, no. 1, pp. 127–143, 2009, doi: 10.1016/j.bpa.2008.09.003.
- [15] M. Amr E., “Comparative Study of Four Supervised Machine Learning Techniques for Classification,” *Int J Appl Sci Technol*, vol. 7, no. 2, pp. 5–18, 2017.
- [16] H. H. Patel and P. Prajapati, “Study and Analysis of Decision Tree Based Classification Algorithms,” *International Journal of Computer Sciences and Engineering*, vol. 6, no. 10, pp. 74–78, 2018, doi: 10.26438/ijcse/v6i10.7478.
- [17] R. A. D. Gordon, “Review Reviewed Work (s): Classification and Regression Trees . by L . Breiman , J . H . Friedman , R . A . Olshen and C . J . Stone Published by : International Biometric Society Stable URL : <https://www.jstor.org/stable/2530946>,” vol. 40, no. 3, pp. 3–4, 1984.
- [18] G. Biau and E. Scornet, “A random forest guided tour,” *Test*, vol. 25, no. 2, pp. 197–227, 2016, doi: 10.1007/s11749-016-0481-7.
- [19] W. H. Hong, J. H. Yap, G. Selvachandran, P. H. Thong, and L. H. Son, “Forecasting mortality rates using hybrid Lee–Carter model, artificial neural network and random forest,” *Complex and Intelligent Systems*, vol. 7, no. 1, pp. 163–189, 2021, doi: 10.1007/s40747-020-00185-w.
- [20] E. D. Attanasi and T. C. Coburn, “Random Forest,” pp. 1182–1185, 2023, doi: 10.1007/978-3-030-85040-1_265.
- [21] P. Florek and A. Zagdański, “Benchmarking state-of-the-art gradient boosting algorithms for classification,” 2023, [Online]. Available: <http://arxiv.org/abs/2305.17094>

- [22] T. Chen and C. Guestrin, “XGBoost: A scalable tree boosting system,” *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, vol. 13-17-Aug, pp. 785–794, 2016, doi: 10.1145/2939672.2939785.
- [23] J. Fan *et al.*, “Comparison of Support Vector Machine and Extreme Gradient Boosting for predicting daily global solar radiation using temperature and precipitation in humid subtropical climates: A case study in China,” *Energy Convers Manag*, vol. 164, no. February, pp. 102–111, 2018, doi: 10.1016/j.enconman.2018.02.087.
- [24] A. Natekin and A. Knoll, “Gradient boosting machines, a tutorial,” *Front Neurorobot*, vol. 7, no. DEC, 2013, doi: 10.3389/fnbot.2013.00021.
- [25] C. O’Hare and Y. Li, “Models of Mortality - Analysing the Residuals,” *SSRN Electronic Journal*, pp. 1–34, 2014, doi: 10.2139/ssrn.2512573.
- [26] H. Turner and D. Firth, “Generalized nonlinear models in R : An overview of the gnm package Generalized linear models Preamble,” *Soc Res (New York)*, 2008.
- [27] A. A. Villegas, P. M. Pietromillo, S. V. K. Vladimir, and M. A. Villegas, “Package ‘ StMoMo ,’” 2022, doi: 10.2307/2290201>.
- [28] E. Beutner, S. Reese, and J. P. Urbain, “Identifiability issues of age–period and age–period–cohort models of the Lee–Carter type,” *Insur Math Econ*, vol. 75, pp. 117–125, 2017, doi: 10.1016/j.insmatheco.2017.04.006.
- [29] S. Haberman and A. Renshaw, “A comparative study of parametric mortality projection models,” *Insur Math Econ*, vol. 48, no. 1, pp. 35–55, Jan. 2011, doi: 10.1016/j.insmatheco.2010.09.003.
- [30] G. A. Spedicato and G. P. Clemente, “Mortality projection with demography and lifecontingencies packages,” no. 2013, pp. 1–7, 2013, [Online]. Available: https://cran.r-project.org/web/packages/lifecontingencies/vignettes/mortality_projection.pdf
- [31] J. Felix-Cardoso, H. Vasconcelos, P. Rodrigues, and R. Cruz-Correia, “Excess mortality during COVID-19 in five European countries and a critique of mortality analysis data,” *medRxiv*, p. 2020.04.28.20083147, 2020, doi: 10.1101/2020.04.28.20083147.

- [32] D. Wang and P. Lu, “Modelling and forecasting mortality distributions in England and Wales using the Lee-Carter model,” *J Appl Stat*, vol. 32, no. 9, pp. 873–885, 2005, doi: 10.1080/02664760500163441.
- [33] N. Brouhns, M. Denuit, and J. K. Vermunt, “A Poisson log-bilinear regression approach to the construction of projected lifetables,” *Insur Math Econ*, vol. 31, no. 3, pp. 373–393, 2002, doi: 10.1016/S0167-6687(02)00185-3.
- [34] W. K. Li and G. Li, “Negative binomial version of the Lee–Carter model for mortality forecasting,” *Appl Stoch Models Bus Ind*, vol. 23, no. January, pp. 237–239, 2009, doi: 10.1002/asmb.
- [35] U. Basellini, C. G. Camarda, and H. Booth, “Thirty years on: A review of the Lee–Carter method for forecasting mortality,” *Int J Forecast*, vol. 39, no. 3, pp. 1033–1049, 2023, doi: 10.1016/j.ijforecast.2022.11.002.
- [36] A. Debón, F. Montes, and F. Puig, “Modelling and forecasting mortality in Spain,” *Eur J Oper Res*, vol. 189, no. 3, pp. 624–637, 2008, doi: 10.1016/j.ejor.2006.07.050.
- [37] A. E. Renshaw and S. Haberman, “On simulation-based approaches to risk measurement in mortality with specific reference to Poisson Lee-Carter modelling,” *Insur Math Econ*, vol. 42, no. 2, pp. 797–816, 2008, doi: 10.1016/j.insmatheco.2007.08.009.
- [38] C. Neves, C. Fernandes, and H. Hoeltgebaum, “Five different distributions for the Lee–Carter model of mortality forecasting: A comparison using GAS models,” *Insur Math Econ*, vol. 75, pp. 48–57, 2017, doi: 10.1016/j.insmatheco.2017.04.004.
- [39] G. Niu and B. Melenberg, “Trends in Mortality Decrease and Economic Growth,” *Demography*, vol. 51, no. 5, pp. 1755–1773, 2014, doi: 10.1007/s13524-014-0328-3.
- [40] F. Girosi and G. King, “Understanding the Lee-Carter Mortality Forecasting Method,” no. 617, p. 33, 2007.
- [41] P. Barrieu *et al.*, “Understanding , Modeling and Managing Longevity Risk : Key Issues and Main Challenges To cite this version : Understanding , Modeling and Managing Longevity Risk : Key Issues and Main Challenges *,” 2010.

- [42] V. D. Amato, S. Haberman, G. Piscopo, and M. Russolillo, “Insurance : Mathematics and Economics Modelling dependent data for longevity projections ☆,” vol. 51, pp. 694–701, 2012, doi: 10.1016/j.insmatheco.2012.09.008.
- [43] A. Hunt and D. P. Blake, “Identifiability in Age/Period/Cohort Mortality Models,” *SSRN Electronic Journal*, no. September 2015, 2020, doi: 10.2139/ssrn.3552213.
- [44] G. Giordano, S. Haberman, and M. Russolillo, “Coherent modeling of mortality patterns for age-specific subgroups,” *Decisions in Economics and Finance*, vol. 42, no. 1, pp. 189–204, 2019, doi: 10.1007/s10203-019-00245-y.
- [45] A. Hunt and A. M. Villegas, “Robustness and convergence in the Lee-Carter model with cohort effects,” *Insur Math Econ*, vol. 64, pp. 186–202, 2015, doi: 10.1016/j.insmatheco.2015.05.004.
- [46] J. S. T. Wong, J. J. Forster, and P. W. F. Smith, “Bayesian model comparison for mortality forecasting,” *J R Stat Soc Ser C Appl Stat*, vol. 72, no. 3, pp. 566–586, 2023, doi: 10.1093/jrssc/qlad021.
- [47] S. J. Richards, I. D. Currie, T. Kleinow, and G. P. Ritchie, “A stochastic implementation of the APCI model for mortality projections,” vol. 24, pp. 1–26, 2019, doi: 10.1017/S1357321718000260.
- [48] M. C. Fung, G. W. Peters, and P. V. Shevchenko, “Cohort effects in mortality modelling: A Bayesian state-space approach,” *Annals of Actuarial Science*, vol. 13, no. 1, pp. 109–144, 2019, doi: 10.1017/S1748499518000131.
- [49] A. Renshaw and S. Haberman, “Lee-carter mortality forecasting: A parallel generalized linear modelling approach for England and Wales mortality projections,” *J R Stat Soc Ser C Appl Stat*, vol. 52, no. 1, pp. 119–137, 2003, doi: 10.1111/1467-9876.00393.
- [50] A. E. Renshaw and S. Haberman, “Lee-Carter mortality forecasting with age-specific enhancement,” *Insur Math Econ*, vol. 33, no. 2, pp. 255–272, 2003, doi: 10.1016/S0167-6687(03)00138-0.
- [51] A. Roshani, M. Izadi, and B. Khaledi, “Modeling and Forecasting Mortality Rate with Cohort Effect and 1 Introduction,” no. October, 2022.

- [52] A. E. Renshaw and S. Haberman, “A cohort-based extension to the Lee-Carter model for mortality reduction factors,” *Insur Math Econ*, vol. 38, no. 3, pp. 556–570, Jun. 2006, doi: 10.1016/j.insmatheco.2005.12.001.
- [53] S. Haberman and A. Renshaw, “On age-period-cohort parametric mortality rate projections,” *Insur Math Econ*, vol. 45, no. 2, pp. 255–270, 2009, doi: 10.1016/j.insmatheco.2009.07.006.
- [54] J. F. Bégin, B. Sanders, and X. Xu, “Modeling and Forecasting Subnational Mortality in the Presence of Aggregated Data,” *North American Actuarial Journal*, vol. 0, no. 0, pp. 1–27, 2023, doi: 10.1080/10920277.2023.2231996.
- [55] K. Barigou, S. Loisel, and Y. Salhi, “Parsimonious predictive mortality modeling by regularization and cross-validation with and without covid-type effect,” *Risks*, vol. 9, no. 1, pp. 1–18, 2021, doi: 10.3390/risks9010005.
- [56] H. Aro and T. Pennanen, “Stochastic modelling of mortality and financial markets,” *Scand Actuar J*, vol. 1238, no. 6, pp. 483–509, 2014, doi: 10.1080/03461238.2012.724442.
- [57] A. J. G. Cairns and K. Dowd, “A TWO-FACTOR MODEL FOR STOCHASTIC MORTALITY WITH PARAMETER UNCERTAINTY :,” vol. 73, no. 4, pp. 687–718, 2006.
- [58] A. Debón, S. Haberman, F. Montes, and E. Otranto, “Do different models induce changes in mortality indicators? That is a key question for extending the lee-carter model,” *Int J Environ Res Public Health*, vol. 18, no. 4, pp. 1–16, 2021, doi: 10.3390/ijerph18042204.
- [59] H. Turner, D. Firth, H. Turner, and D. Firth, “Generalized nonlinear models in R : An overview of the gnm package ESRC National Centre for Research Methods Generalized nonlinear models in R : An overview of the gnm package”.
- [60] H. Booth, R. J. Hyndman, L. Tickle, and P. de Jong, “Lee-Carter mortality forecasting: A multi-country comparison of variants and extensions,” *Demogr Res*, vol. 15, 2006, doi: 10.4054/demres.2006.15.9.
- [61] S. J. Richards, I. D. Currie, T. Kleinow, and G. P. Ritchie, “A stochastic implementation of the APCI model for mortality projections,” *British Actuarial Journal*, vol. 24, pp. 1–26, 2021, doi: 10.1017/S1357321718000260.

- [62] I. D. Currie, “On fitting generalized linear and non-linear models of mortality,” *Scand Actuar J*, vol. 2016, no. 4, pp. 356–383, 2016, doi: 10.1080/03461238.2014.928230.
- [63] C. Maccheroni and S. Nocito, “Backtesting the Lee–Carter and the Cairns–Blake–Dowd stochastic mortality models on italian death rates,” *Risks*, 2017, doi: 10.3390/risks5030034.
- [64] K. Dowd, A. J. G. Cairns, D. Blake, G. D. Coughlan, D. Epstein, and M. Khalaf-Allah, “Evaluating the goodness of fit of stochastic mortality models,” *Insur Math Econ*, vol. 47, no. 3, pp. 255–265, 2010, doi: 10.1016/j.insmatheco.2010.06.006.
- [65] P. M. Barrieu and L. A. M. Veraart, “Pricing q-forward contracts: an evaluation of estimation window and pricing method under different mortality models,” *Scand Actuar J*, vol. 2016, no. 2, pp. 146–166, 2016, doi: 10.1080/03461238.2014.916228.
- [66] G. Kesavaraj and S. Sukumaran, “A study on classification techniques in data mining,” *2013 4th International Conference on Computing, Communications and Networking Technologies, ICCCNT 2013*, pp. 1–7, 2013, doi: 10.1109/ICCCNT.2013.6726842.
- [67] W. Y. Loh, “Fifty years of classification and regression trees,” *International Statistical Review*, vol. 82, no. 3, pp. 329–348, 2014, doi: 10.1111/insr.12016.
- [68] P. Deprez, P. V. Shevchenko, and M. V. Wüthrich, “Machine Learning Techniques for Mortality Modeling,” May 2017, [Online]. Available: <http://arxiv.org/abs/1705.03396>
- [69] S. Scognamiglio, “Calibrating the Lee-Carter and the Poisson Lee-Carter Models Via Neural Networks,” *ASTIN Bulletin*, vol. 52, no. 2, pp. 519–561, 2022, doi: 10.1017/asb.2022.5.
- [70] A. Kaya, “Conflicted Principals, Uncertain Agency: The International Monetary Fund and the Great Recession,” *Glob Policy*, vol. 3, no. 1, pp. 24–34, 2012, doi: 10.1111/j.1758-5899.2011.00096.x.
- [71] S. A. Rakhmawan, M. H. Omar, M. Riaz, and N. Abbas, “Hotelling T2 Control Chart for Detecting Changes in Mortality Models Based on Machine-Learning Decision Tree,” *Mathematics*, vol. 11, no. 3, Feb. 2023, doi: 10.3390/math11030566.

- [72] Patrick Zöhbauer, “Data Science in Non-Life Insurance Pricing,” no. July, 2017, [Online]. Available: <https://ethz.ch/content/dam/ethz/special-interest/math/risklab-dam/documents/walter-saxer-preis/ma-zoechbauer.pdf>
- [73] W. H. Hong, J. H. Yap, G. Selvachandran, P. H. Thong, and L. H. Son, “Forecasting mortality rates using hybrid Lee–Carter model, artificial neural network and random forest,” *Complex and Intelligent Systems*, vol. 7, no. 1, pp. 163–189, 2021, doi: 10.1007/s40747-020-00185-w.
- [74] A. Natekin and A. Knoll, “Gradient boosting machines, a tutorial,” *Front Neurorobot*, vol. 7, no. DEC, 2013, doi: 10.3389/fnbot.2013.00021.
- [75] M. Nalluri, M. Pentela, and N. R. Eluri, “A Scalable Tree Boosting System: XG Boost,” *Int. J. Res. Stud. Sci. Eng. Technol*, vol. 7, no. 12, pp. 36–51, 2020.
- [76] S. Fafalios, P. Charonyktakis, and I. Tsamardinos, “Gradient Boosting Trees,” *Stefanos Fafalios, Pavlos Charonyktakis, Ioannis Tsamardinos Gnosis Data Analysis PC*, no. April, pp. 1–13, 2020, [Online]. Available: https://www.gnosisda.gr/wp-content/uploads/2020/07/Gradient_Boosting_Implementation.pdf
- [77] T. Chen and C. Guestrin, “LearningSys_2015_paper_32,” pp. 1–6, 2015.
- [78] C. Bentéjac, A. Csörgö, and G. Martínez-Muñoz, *A comparative analysis of gradient boosting algorithms*, vol. 54, no. 3. Springer Netherlands, 2021. doi: 10.1007/s10462-020-09896-5.
- [79] M. Denil, D. Matheson, and N. De Freitas, “Narrowing the Gap: Random Forests In TheDenil, M., Matheson, D., & De Freitas, N. (2014). Narrowing the Gap: Random Forests In Theory and In Practice. Proceedings of The 31st International Conference on Machine Learning, (1998), 665–673. Retrieved from ht,” *Proceedings of The 31st International Conference on Machine Learning*, no. 1998, pp. 665–673, 2014, [Online]. Available: <http://jmlr.org/proceedings/papers/v32/denil14.html>
- [80] A. Ziegler and I. R. König, “Mining data with random forests: Current options for real-world applications,” *Wiley Interdiscip Rev Data Min Knowl Discov*, vol. 4, no. 1, pp. 55–63, 2014, doi: 10.1002/widm.1114.
- [81] Z. Jin, J. Shang, Q. Zhu, C. Ling, W. Xie, and B. Qiang, “RFRSF: Employee Turnover Prediction Based on Random Forests and Survival Analysis,” *Lecture Notes in Computer Science (including subseries Lecture*

- Notes in Artificial Intelligence and Lecture Notes in Bioinformatics*), vol. 12343 LNCS, pp. 503–515, 2020, doi: 10.1007/978-3-030-62008-0_35.
- [82] J. M. Klusowski, “Complete Analysis of a Random Forest Model,” *ArXiv*, vol. 13, pp. 1063–1095, 2018.
 - [83] R. Genuer, J.-M. Poggi, and C. Tuleau, “Random Forests: some methodological insights,” 2008, [Online]. Available: <http://arxiv.org/abs/0811.3619>
 - [84] P. Probst, M. N. Wright, and A. L. Boulesteix, “Hyperparameters and tuning strategies for random forest,” *Wiley Interdiscip Rev Data Min Knowl Discov*, vol. 9, no. 3, pp. 1–15, 2019, doi: 10.1002/widm.1301.
 - [85] J. Lee, J. Cai, F. Li, and Z. A. Vesoulis, “Predicting mortality risk for preterm infants using random forest,” *Sci Rep*, vol. 11, no. 1, pp. 1–9, 2021, doi: 10.1038/s41598-021-86748-4.
 - [86] S. Levantesi and A. Nigri, “A random forest algorithm to improve the Lee–Carter mortality forecasting: impact on q-forward,” *Soft comput*, vol. 24, no. 12, pp. 8553–8567, 2020, doi: 10.1007/s00500-019-04427-z.
 - [87] B. E. McNellis, A. M. S. Smith, A. T. Hudak, and E. K. Strand, “Tree mortality in western U.S. forests forecasted using forest inventory and Random Forest classification,” *Ecosphere*, vol. 12, no. 3, 2021, doi: 10.1002/ecs2.3419.
 - [88] J. Pirneskoski *et al.*, “Random Forest machine learning method outperforms prehospital National Early Warning Score for predicting one-day mortality: A retrospective study,” *Resusc Plus*, vol. 4, no. May, p. 100046, 2020, doi: 10.1016/j.resplu.2020.100046.
 - [89] J. L. Speiser, M. E. Miller, J. Tooze, and E. Ip, “A comparison of random forest variable selection methods for classification prediction modeling,” *Expert Syst Appl*, vol. 134, pp. 93–101, 2019, doi: 10.1016/j.eswa.2019.05.028.
 - [90] J. Abellán, C. J. Mantas, and J. G. Castellano, “A Random Forest approach using imprecise probabilities,” *Knowl Based Syst*, vol. 134, pp. 72–84, 2017, doi: 10.1016/j.knosys.2017.07.019.
 - [91] R. Das, S. Biswas, and A. Mukherjee, “Maternal Mortality at a Teaching Hospital of Rural India: A Retrospective Study,” *International Journal of Biomedical and Advance Research*, 2014, doi: 10.7439/ijbar.v5i2.578.

- [92] S. Hertog, “R Scripts for computing adult and maternal mortality from DHS sibling survival histories,” 2019, [Online]. Available: <https://www.un.org/en/development/desa/population/publications/pdf/technical/TP2019-2.pdf>
- [93] J. Toohill, Y. Chadha, and S. Nowlan, “An interactive decision-making framework (i-DMF) to scale up maternity continuity of carer models,” *Journal of Research in Nursing*, vol. 25, no. 6–7, pp. 561–576, 2020, doi: 10.1177/1744987119887424.
- [94] P. Idoko, M. O. Anyanwu, and S. Bass, “A retrospective analysis of trends in maternal mortality in a Gambian tertiary health centre,” *BMC Res Notes*, vol. 10, no. 1, pp. 1–7, 2017, doi: 10.1186/s13104-017-2817-0.
- [95] S. Harrison, J. Henderson, F. Alderdice, and M. A. Quigley, “Methods to increase response rates to a population-based maternity survey: A comparison of two pilot studies,” *BMC Med Res Methodol*, vol. 19, no. 1, pp. 1–8, 2019, doi: 10.1186/s12874-019-0702-3.
- [96] Z. Amarin, Y. Khader, A. Okour, H. Jaddou, and R. Al-Qutob, “National maternal mortality ratio for Jordan, 2007–2008,” *International Journal of Gynecology and Obstetrics*, vol. 111, no. 2, pp. 152–156, 2010, doi: 10.1016/j.ijgo.2010.05.016.
- [97] S. a Sarpong, “Modeling and Forecasting Maternal Mortality ; an Application of ARIMA Models,” *Int J Appl Sci Technol*, vol. 3, no. 1, pp. 19–28, 2013.
- [98] S. Ahmed and K. Hill, “Estimation de la mortalité maternelle au niveau sous-national: Une méthode à base de modèles appliquée au Bangladesh,” *Bull World Health Organ*, vol. 89, no. 1, pp. 12–21, 2011, doi: 10.2471/BLT.10.076851.
- [99] L. Onambele *et al.*, “Trends, Projections, and Regional Disparities of Maternal Mortality in Africa (1990–2030): An ARIMA Forecasting Approach,” *Epidemiologia*, vol. 4, no. 3, pp. 322–351, 2023, doi: 10.3390/epidemiologia4030032.
- [100] D. Kaur, V. Kaur, and V. I. Yuel, “Alarming high maternal mortality in 21st century,” *JK Science*, vol. 9, no. 3, pp. 123–126, 2007.
- [101] R. Pittrof, O. M. R. Campbell, and V. G. A. Filippi, “What is quality in maternity care? An international perspective,” *Acta Obstet Gynecol Scand*, vol. 81, no. 4, pp. 277–283, 2002, doi: 10.1034/j.1600-0412.2002.810401.x.

- [102] R. Das, S. Biswas, and A. Mukherjee, “Maternal Mortality at a Teaching Hospital of Rural India: A Retrospective Study,” *International Journal of Biomedical and Advance Research*, vol. 5, no. 2, p. 114, 2014, doi: 10.7439/ijbar.v5i2.578.
- [103] Y. M. Adamu, H. M. Salihu, N. Sathiakumar, and G. R. Alexander, “Maternal mortality in Northern Nigeria: A population-based study,” *European Journal of Obstetrics and Gynecology and Reproductive Biology*, vol. 109, no. 2, pp. 153–159, 2003, doi: 10.1016/S0301-2115(03)00009-5.
- [104] H. Booth, L. Tickle, and L. Smith, “Evaluation of the Variants of the Lee-Carter Method of Forecasting Mortality: A Multi-Country Comparison,” *New Zealand Population Review, Special Issue*, vol. 31, no. 1, pp. 13–34, 2005.
- [105] N. Brouhns, M. Denuit, and J. K. Vermunt, “A Poisson log-bilinear regression approach to the construction of projected lifetables,” *Insur Math Econ*, vol. 31, no. 3, pp. 373–393, 2002, doi: 10.1016/S0167-6687(02)00185-3.
- [106] H. Booth, J. Maindonald, and L. Smith, “AGE-TIME INTERACTIONS IN MORTALITY PROJECTION : Working Papers in Demography Revised 2 August 2002,” *North American Actuarial Journal*, no. August, 2002.
- [107] H. Booth, J. Maindonald, and L. Smith, “Applying Lee-Carter under conditions of variable mortality decline,” *Popul Stud (NY)*, vol. 56, no. 3, pp. 325–336, 2002, doi: 10.1080/00324720215935.
- [108] R. J. Hyndman and M. Shahid Ullah, “Robust forecasting of mortality and fertility rates: A functional data approach,” *Comput Stat Data Anal*, vol. 51, no. 10, pp. 4942–4956, 2007, doi: 10.1016/j.csda.2006.07.028.
- [109] E. Debonneuil, “A simple model of mortality trends aiming at universality: Lee Carter + Cohort,” *Quantitative Finance Papers 1003*, p. 5, 2010, [Online]. Available: <http://arxiv.org/abs/1003.1802>
- [110] A. J. G. Cairns, D. Blake, K. Dowd, G. D. Coughlan, D. Epstein, and M. Khalaf-Allah, “Mortality density forecasts: An analysis of six stochastic mortality models,” *Insur Math Econ*, vol. 48, no. 3, pp. 355–367, 2011, doi: 10.1016/j.insmatheco.2010.12.005.
- [111] Human Mortality Database, “University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany),” Available at www.mortality.org, 2020.

- [112] T. M. Therneau and E. J. Atkinson, “An introduction to recursive partitioning using the RPART routines. Mayo foundation, available from: <http://www.mayo.edu/hsr/techrpt/61.pdf>,” p. 52, 1997.
- [113] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification And Regression Trees*. Routledge, 2017. doi: 10.1201/9781315139470.
- [114] A. Liaw and M. Wiener, “Classification and Regression by randomForest,” *R News*, vol. 2, no. 3, pp. 18–22, 2002.
- [115] G. Ridgeway, “Generalized Boosted Models: A guide to the gbm package.package ‘gbm’, version 1.5-7.,” *Update*, vol. 1, no. 1, pp. 1–15, 2007.
- [116] J. R. Wilmoth *et al.*, “A New Method for Deriving Global Estimates of Maternal Mortality,” *Stat Politics Policy*, vol. 3, no. 2, Jan. 2012, doi: 10.1515/2151-7509.1038.
- [117] WHO, “WHO guidance for measuring maternal mortality from a census,” *Report*, 2013.
- [118] WHO, *Trends in maternal mortality 2000 to 2020: estimates*. 2021. [Online]. Available: <https://www.who.int/reproductivehealth/publications/maternal-mortality-2000-2017/en/>
- [119] D. Cutler, A. Deaton, and A. Lleras-Muney, “The determinants of mortality,” *Journal of Economic Perspectives*, vol. 20, no. 3, pp. 97–120, 2006, doi: 10.1257/jep.20.3.97.
- [120] R. S. John, “Lee-Carter Model in a Generalised Linear Modelling Framework 2 . Credibility Models and Generalised Linear,” no. 1996, pp. 1–24, 2019.
- [121] A. J. G. Cairns, D. Blake, and K. Dowd, “Modelling and management of mortality risk: A review,” *Scand Actuar J*, vol. 1238, no. 2–3, pp. 79–113, 2008, doi: 10.1080/03461230802173608.
- [122] G. Díaz-Rojo, A. Debón, and J. Mosquera, “Multivariate control chart and Lee–Carter models to study mortality changes,” *Mathematics*, vol. 8, no. 11, pp. 1–17, Nov. 2020, doi: 10.3390/math8112093.

Vitae

Name :Sulyman Iyanda

Nationality :Nigerian

Date of Birth :10/5/1992

Email :sulymaniyanda@gmail.com

Address :147, Alimi Road Alore, Ilorin Kwara, Nigeria

Postal code 240272

Academic Background

MSc. in Applied Statistics 2024
King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia.

MSc. in Mathematical Sciences 2018/2019
African Institute for Mathematical Sciences (AIMS Cameroon), Cameroon.

BSc. in Statistics 2012-2016
University of Ilorin, Ilorin, Kwara State, Nigeria.

West African Senior School Certificate Examination, 2010
Community Secondary School, Banni, Ilorin, Kwara State, Nigeria

RESEARCH INTEREST

- Statistical modelling and Machine learning
- Applied Statistics in Public Health - Biostatistics
- Computational Bayesian Statistics

- Statistical Process control
- Demographic and Survival Analysis

TEACHING EXPERIENCE

Graduate Teaching Assistant 2022-till date

Department of Mathematics, KFUPM, Dhahran, Saudi Arabia.

High School Principal 2018, 2020-2021

Marokiz-l-Hidayah College, Kwara State, Nigeria.

High School Maths Teacher (National Youth Service Corps). 2017

Government Day Secondary School, Niger State, Nigeria.

Mathematics Tutor 2014-2016

Adex Educational consult and Computer academy, Kwara State, Nigeria.

Undergraduate Tutor 2013-2014

Muslim Students Society of Nigeria (MSSN), University of Ilorin chapter

Mathematics Teacher (Community Service) 2010-2012

Alore Descendants Community Development Association (ADCDA),

MANUSCRIPTS/ PUBLICATION

Implementing Age-Gender Effects on Stochastic Mortality Models Using the GNM Framework: A Machine Learning Enhancement. *(Submitted for publication)*

Comparison of stochastic mortality model enhanced by machine learning *(Drafted)*

WORKSHOP/ CONFERENCE

First KFUPM Interdisciplinary Graduate Students Conference May 2024

By Mr. Sulyman Iyanda

Examining age-gender effect on Lee-Carter model Using GNM framework.
 3rd IRC-ISS Student Research Symposium at KFUPM May 2024
By Mr. Sulyman Iyanda

Exploring Gender Effects in Stochastic Mortality Models Using
 Machine Learning Techniques
 King Fahd University May. 2023
 Women in Data Science
 King Fahd University Oct. 2022
By Prof. Muhammad Riaz.

Poisson Adaptive EWMA control chart for monitoring Range of shifts.
 AIMS-Cameroon Nov. 2018
By Prof. Roger Stern, University of Reading.

Climatic data and small-scale farmers: What's the role of a Statistician?
 AIMS-Cameroon Jan. 2019
By Prof. Jane Hutton, Warwick University, UK.

"Statistics and the law"- Statistical evidence in civil case.
 AIMS-Cameroon Feb. 2019
*By Prof. Kristan A. Schneider, University of Applied Sciences,
 Mittweida, Germany.*

Mathematical models of the spread of drug resistance in malaria.