“Hierarchical forecasting of causes of death with trend breaks in mortality modeling: Kenyan case”

AUTHORS	Nicholas Bett Juma Kasozi Daniel Ruturwa

ARTICLE INFO	Nicholas Bett, Juma Kasozi and Daniel Ruturwa (2025). Hierarchical forecasting of causes of death with trend breaks in mortality modeling: Kenyan case. *Insurance Markets and Companies*, *16*(1), 15-32. doi:10.21511/ins.16(1).2025.02
DOI	http://dx.doi.org/10.21511/ins.16(1).2025.02
RELEASED ON	Tuesday, 21 January 2025
RECEIVED ON	Sunday, 22 September 2024
ACCEPTED ON	Thursday, 07 November 2024

JOURNAL	"Insurance Markets and Companies"
ISSN PRINT	2616-3551
ISSN ONLINE	2522-9591
PUBLISHER	LLC “Consulting Publishing Company “Business Perspectives”
FOUNDER	LLC “Consulting Publishing Company “Business Perspectives”

NUMBER OF REFERENCES	NUMBER OF FIGURES	NUMBER OF TABLES
49	8	3

BUSINESS PERSPECTIVES

LLC "CPC "Business Perspectives"
Hryhorii Skovoroda lane, 10,
Sumy, 40022, Ukraine

www.businessperspectives.org

Nicholas Bett (Rwanda), **Juma Kasozi** (Rwanda, Uganda), **Daniel Ruturwa** (Rwanda)

HIERARCHICAL FORECASTING OF CAUSES OF DEATH WITH TREND BREAKS IN MORTALITY MODELING: KENYAN CASE

Abstract

Trends offer direction and momentum. However, trends in mortality are affected by trend breaks, which are a consequence of mortality shocks. Additionally, insufficient historical data challenge the credibility of the forecasted trends, which are useful for actuaries in pricing, reserving, and valuing life insurance products. To address these challenges, the study aims to determine and incorporate trend breaks among individual causes of death and coherently forecast them by applying the bottom-up hierarchical forecasting approach for life insurance models. The models used are categorized as base (linear model), auto-statistical (Arima, Exponential-Smoothing, and Prophet), and auto-machine learning. The data from the World Health Organization consisted of annualized mortality quantities by cause, gender, age, and period for Kenya. Results based on the mean absolute percentage error criteria across the causes of death showed that all the models apart from the base model showed significant improvement after accounting for the trend breaks with the best being the auto machine learning approach leading with seven causes of death. Updating forecasts based on the computed trend breakpoints that varied between 2007 to 2011 generally improved forecast accuracy. These results suggest that forecasting errors may be reduced after accounting for trend breaks and model specifications. Furthermore, this implies that insufficient data do not necessarily produce deficient forecasts. The study's contribution involved applying approaches that enhance the accuracy of forecasting models to prevent adverse effects of mortality shocks in actuarial modeling.

Keywords actuarial, machine learning, structural break, longevity, risk management, temporal, mortality shock, bottom-up

JEL Classification J11, G22

Received on: 22nd of September, 2024
Accepted on: 7th of November, 2024
Published on: 21st of January, 2025

Nicholas Bett, MSc., Junior Research Fellow, African Centre of Excellence in Data Science (ACEDS), College of Business and Economics, University of Rwanda, Rwanda. (Corresponding author)

Juma Kasozi, Professor, Department of Mathematics, College of Natural Sciences, Makerere University, Rwanda, Uganda.

Daniel Ruturwa, Senior Lecturer, Department of Applied Statistics, School of Economics, University of Rwanda, Rwanda.

Conflict of interest statement:
Author(s) reported no conflict of interest

INTRODUCTION

Long-term trends in mortality and longevity offer direction and momentum and are useful in mortality modeling, which is of value to actuaries in insurance, medical, and social security schemes. However, mortality trends are affected by structural trend breaks, which are also known as mortality shocks (Tang et al., 2022). For instance, the HIV/AIDS epidemic had the biggest impact on mortalities at the start of the century (Bett et al., 2023), specifically in the Kenyan case. Contextually, deaths due to HIV/AIDS peaked in 2000 for both males and females, negatively strained insurance liabilities, and contributed to the insolvency of insurance companies because of inaccurate estimations of mortality risk assumptions (Waweru, 2014). However, the eventual reduction of deaths due to HIV/AIDS accounted for the upward revision of the life expectancy in Kenya (UN, 2017). This scenario has further compounded uncertainty in mortality modeling not only due to HIV/AIDS but also other causes of death, hence the need to incorporate causes of deaths in modeling mortality.

Earlier mortality models were premised that mortality increases with age and were purely deterministic (Bengtsson & Keilman, 2019). However, patterns from mortality data have displayed uncertainty in the rates over time and led to the development of widely used models such as the Lee and Carter (LC) (1992) and the Cairns Blake Dowd (CBD) Model (Cairns et al., 2009). To date, these models form the backbone of mortality modeling in many jurisdictions, especially in developed countries with adequate historical data and long-term trends, which are necessary for linearity; however, this is not the case for developing countries like Kenya, where data are insufficient and unstable due to mortality shocks.

1. LITERATURE REVIEW

The studies of changepoint models are vast; they include non-parametric methods proposed by Pettitt (1979), as well as approaches linked to decision theory, as suggested by Dayanik et al. (2008). These studies have concluded that Bayesian approaches are the most appropriate in determining trend structural breaks. A drawback, however, is the increased difficulty in representing the posterior distribution, that is, in the detection of changes in a trend where new data are continually generated, necessitating the update of the posterior with each new piece of information. This iterative update can rapidly grow intricate, particularly in relation to dimensionality. An effective method eventually preferred is to compute the set of changepoints that are most probable, known as the Maximum A Posteriori (MAP) set (Fearnhead, 2005). However, the primary concern in applying this Bayesian approach is in the selection of priors.

The application of regression models in time series, especially in estimating structural breaks, was initially presented by Bai (1994) and then expanded to consist of several breaks (Bai & Perron, 1998) based on the algorithm exhibited in the study by Bai and Perron (2003) and further confirmed by Zeileis et al. (2003). The underlying algorithm is derived from dynamic programming as it calculates the appropriate breakpoints, given the specified number of breaks at the onset. The fundamental concept is that of the Bellman principle (Giuseppi & Pietrabissa, 2022) that is a solution to a complex problem can be established by solving smaller and individual subproblems optimally and thereafter combining them, which is advantageous because it enables systemization and computational efficiency. Coelho and Nunes (2011) and Tang et al. (2022) analyzed the topic of forecasting future mortality and life expectancy when there is a major shift in the underlying mortality pattern. They demonstrated, under the LC framework, the use of statistical testing to identify structural changes, resulting in a precise forecasting model for the overall mortality rate. Specifically, they conducted examinations to identify any changes in the mortality index's pattern. The three main categories of approaches proposed to include the impact of structural changes are the regime-switching model (Hamilton, 1989), the broken-trend stationary model as proposed by Perron (1989), and the difference stationary processes with breakpoints by van Berkum et al. (2013). Milidonis et al. (2011) observed that regime-switching models may illustrate different states of mortality. More precisely, they observed that distinct averages and variations characterize the error term of the mortality index in the LC model as it transitions between two different states. Additionally, they proposed possible departures from the assumption that the error term follows a normal distribution. A key critique of regime-flipping models is that it may be unrealistic to assume that influences on human mortality, such as medical breakthroughs, propel newer causes of death.

Several studies (Arnold & Glushko, 2021; Arnold & Sherris, 2015; Caselli et al., 2019; Robertson et al., 2013) have pushed for the inclusion of causes of death as a way to break down overall mortality. This has also been shown by the use of hierarchical forecasting in mortality modeling with causes of death (Li & Lu, 2018). In general, this approach allows for the organic breakdown of aggregated data into more specific subcategories, resulting in a hierarchical format (Athanasopoulos et al., 2009). The techniques for generating consistent predictions for both hierarchical time series encompass top-down, middle-out, combined, and bottom-up approaches (Hyndman et al., 2011). Research has demonstrated that the latter approach plays a crucial role in enabling accurate initial projections, thereby enhancing the efficiency of the reconciliation process and ensuring the consistency and precision of the final forecasts.

http://dx.doi.org/10.21511/ins.16(1).2025.02

The philosophy behind the disaggregation of the hierarchical forecasting approach is that the forecasts should be coherent, unbiased, and minimal. The base forecasts are the main focus of the bottom-up approach and encompass the causes of deaths that eventually aggregate into the total mortality forecast; this aligns with the key purpose of this study. Additionally, this approach optimizes the information at the base, as Zhang et al. (2023) assert that accurate initial predictions are crucial for producing consistent and unified predictions in hierarchical forecasting. This coherence ensures the proper alignment of predictions at different hierarchy levels with aggregation requirements. Enhancing the precision of the initial base forecasts improves the overall precision of the reconciled forecasts. In contrast, Gross and Sohl (1990) introduced the top-down approaches, which aim to predict the performance of individual goods in a product line by starting from the top series and progressively moving downward. The top-down approach leads to a loss of information, particularly from the lower levels, which impacts the overall reconciliation. On the other hand, the middle-out approach integrates both top and down methods, thereby averaging the results (Wickramasuriya et al., 2019). The bottom-up approach is advantageous because it minimizes information leakage and captures the dynamics of individual univariate series. Additionally, the advent of capable computer infrastructure reduces the cost of computing when modeling numerous individual series simultaneously.

Developments in mortality modeling have also seen an increase in the application of algorithmic models that are an alternative to stochastic statistical models, as pointed out by Breiman (2001). These models are also known as machine learning models and are classified as supervised, unsupervised, and reinforcement learning (Richman, 2018). Mullainathan and Spiess (2017) point out that they constitute the general foundation of representation learning, where the machine receives raw data and subsequently detects patterns. Furthermore, the study asserts that machine learning has demonstrated its applicability in actuarial modeling, effectively addressing many regression model-based problems because manual implementation can introduce errors, and incorporating machine learning techniques may enhance model accuracy. However, one limitation of the above models is their reliance on sufficient data. That notwithstanding, using tuned deep neural learning methods, one can customize the neural architecture and capture the complexities of cause-of-death features with limited data. Additionally, automatic machine learning approaches (LeDell & Poirier, 2020) that train numerous models and iteratively tune vast hyperparameters to enhance accuracy justify these models. Thus, the incorporation of automated machine learning approaches to hierarchical forecasting of the cause of death is considered.

In the presence of mortality shocks, limited data, and less adapted forecasting models due to the nature of the data, there is a need to consider the incorporation of trend breaks and improved approaches in forecasting coherently. Changepoint models, particularly Bayesian approaches, are effective in identifying structural breaks in trends. However, they face challenges in updating posterior distributions with new and sequential data like time series, which can complicate the analysis as dimensionality increases, and therefore, regression models for time series utilize dynamic programming to determine breakpoints. Integrating hierarchical forecasting techniques like the bottom-up approach, which enhances mortality predictions by breaking down aggregated data into specific subcategories with machine learning models in mortality modeling, offers an alternative to traditional statistical methods, improving accuracy and efficiency, and can benefit from automated approaches for better forecasting of causes of death.

This study, therefore, aims to determine and incorporate trend breaks among individual causes of death and coherently forecast them by applying the bottom-up hierarchical forecasting approach for life insurance models.

2. METHOD

The dataset link is obtained from the World Health Organization, WHO database for Kenya (WHO, 2022). It is based on 131 causes of death (COD) across the years 2000–2019. The data are disaggregated as follows: the first level, 0, is composed

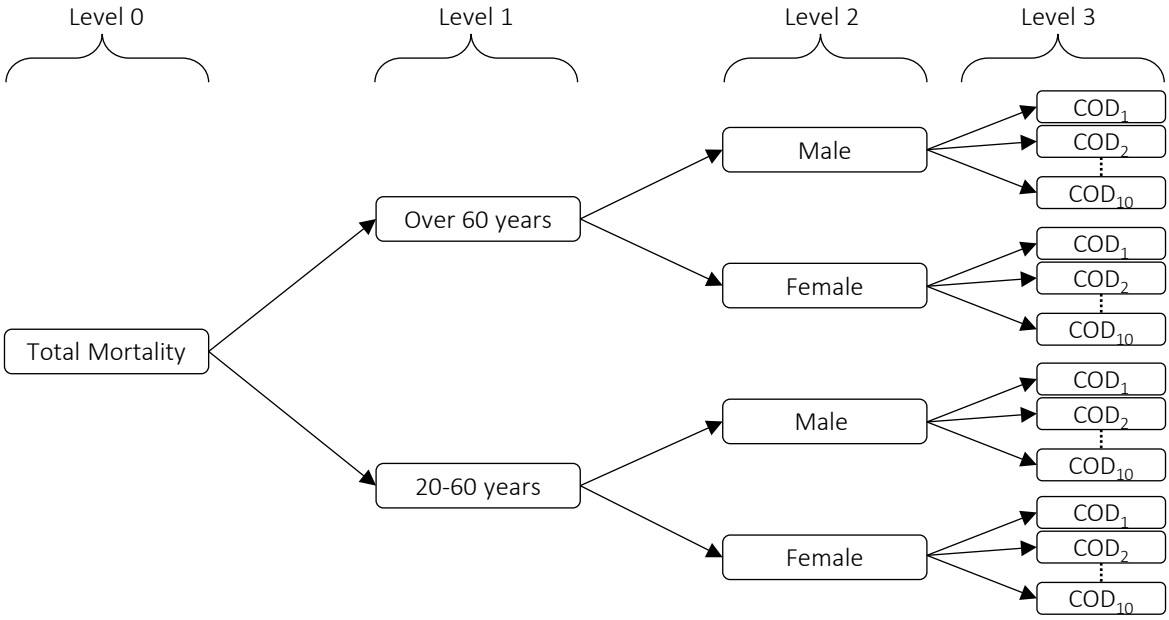


Figure 1. Three-level hierarchical data structure

of the aggregate mortality rate for the entire data set; here, the mortality rate is the average across age, gender, and causes of death. The second level, 1, consists of age-partitioned aggregate mortality. Here, only two age categories are considered: 20 to 60 years and over 60 years, averaged across gender and causes of death. The former age group caters to the aspects of mortality risks, legal eligibility for insurance purchases based on the data jurisdiction, and the working age group, while the latter caters for longevity risks or the post-retirement age group. The third level, 2, consists of the gender partition averaged across causes of deaths for specific age partitions, and the fourth level, 3, is based on the individual causes of deaths for specific age and gender partitions (Figure 1). The total mortality can be disaggregated by the age, gender, and cause of death variables.

2.1. Trend breaks detection

Breakpoints refer to the number of observations that indicate the conclusion of a segment. Suppose we have a set of time-series data, denoted as $y_{1:n} = (y_1,\ldots, y_n)$, and observation is univariate at each time t, y_t, then,

$$y_t = \alpha + \beta t + \sum_{k=1}^{m} (\alpha_k + \beta_k t) I\left(t \geq \tau_k\right) + \varepsilon_t, \quad (1)$$

where y_t is the observed value, α, β are intercepts of trend, τ_k time points where changepoints occur, and I $(t \geq \tau_k)$ is an indicator function. This approach aims to examine that the regression parameters remain constant against the possibility that at least one parameter changes with time such that:

$$H_0 : \beta_i = \beta_0 \qquad (i = 1,\ldots,n). \quad (2)$$

The model will have several changepoints, denoted as m, where, $\tau_{1:m} = (\tau_1, \ldots, \tau_m)$. The position of individual changepoints is a number ranging from 1 to n-1, including both endpoints. Let τ_0 be defined as 0 and τ_{m+1} as n. The changepoints are assumed to be arranged in ascending order, such that $\tau_i < \tau_j$ if and only if $i < j$. For this approach, the conventional linear regression model is examined (Killick et al., 2010).

2.2. Hierarchical forecasting approach

The projections at the bottom are referred to as the base forecasts. The items are arranged in the same sequence as the data. Subsequently, all methods of predicting future outcomes for hierarchical or clustered systems can be expressed as:

$$\tilde{y}_h = S\hat{y}_h. \quad (3)$$

The summing matrix S enables the combining structure to aggregate these forecasts coherently.

http://dx.doi.org/10.21511/ins.16(1).2025.02

This matrix is defined based on the implemented approach. The first three columns nullify the base forecasts, while the n-dimensional identity matrix selects only the base forecasts of the bottom level.

Consider the forecast notations for each of the lower-level series, predicting h-steps

Table 1. Forecast notations for lower-level 1

Cause of death data notation	Meaning
TM	Total aggregate mortality
20_60	Aged 20 to 60 years
60	Aged over 60 years
F_20_60	Female aged 20 to 60 years
F_60	Female aged over 60 years
M_20_60	Male aged 20 to 60 years
M_60	Male aged over 60 years
{COD_1_40}	Vector of 40 univariate series of individual causes of death (bottom series)

Such that,

$$\hat{y}_{TM,\,h},\hat{y}_{20\_60,\,h},\hat{y}_{60,\,h},\hat{y}_{F\_20\_60,\,h},\hat{y}_{F\_60,\,h},\quad \hat{y}_{M\_20\_60,\,h},\hat{y}_{M\_60,\,h},\left\{\hat{y}_{COD\_1\_40,\,h}\right\} \tag{4}$$

By summing them up we can get,

$$\hat{y}_{TM,h}=\hat{y}_{20\_60,h}+\hat{y}_{60,h}+\hat{y}_{F\_20\_60,h}+\hat{y}_{F\_60,h}+\hat{y}_{M\_20\_60,h}+\hat{y}_{M\_60,h}+\hat{y}_{COD\_1\_40,h} \tag{5}$$

By employing a more concise notation, the bottom-up technique becomes

$$\tilde{y}_h=S\hat{b}_h \tag{6}$$

where $\tilde{y}_h$ is an n-dimensional vector of coherent h-step ahead forecasts, and $\hat{b}_h$ is an m-dimensional vector of h-step ahead forecasts for the bottom level series. The vector $\tilde{y}_t$ represents coherent h-step ahead forecasts for an n-dimensional space, while

Table 2. Implemented models

Model Name	Formula	Parameters
1. Auto. Arima (Mélard & Pasteels, 2000)	$\nabla^d y_t=\alpha+\sum_{i=1}^{p}\varnothing_i\nabla^d y_{t-i}+\varepsilon_t+\sum_{j=1}^{q}\theta_j\nabla^d\varepsilon_{t-j}$	ARIMA(p, d, q)
2. ETS (Billah et al., 2006; Chen & Moraga, 2024)	$y_t=\left(l_{t-1}+b_{t-1}\right)s_{t-m}+\varepsilon_t$	l_t is the level; b_t is the trend; s_t is the seasonality and ε_t is the error term
3. Prophet (Taylor & Letham, 2018)	$y_t=g_t+s_t+h_t+\varepsilon_t$	h_t models holidays
4. Linear Model LM (Freedman, 2009)	$y=\beta_0+\sum_{i=1}^{p}\beta_i x_i+\varepsilon$	y is the dependent and x's are the independent variables. β_0 and β_i are coefficients and ε is the error
5. AutoML (LeDell & Poirier, 2020)		
a. Gradient Boosting Machine (GBM)	$F_m(x)=F_{m-1}(x)+v\cdot h_m(x)$	$F_m(x)$ is the model at iteration m and v is the learning rate
b. Distributed Random Forest (DRF)	$\hat{y}=\frac{1}{T}\sum_{t=1}^{T}h_t(x)$	$\hat{y}$ is the predicted output; T is the total number of trees; $h_t(x)$ is the prediction of the t-th tree
c. Generalized Linear Model (GLM)	$g\left(E\left(Y/X\right)\right)=\beta_0+\sum_{i=1}^{p}\beta_i X_i$	g is the link function; $E(Y/X)$ is the expected value of Y given X, β_0 and β_i are coefficients
d. Deep Learning (Neural Networks)	$\hat{y}=f(W^L\cdot f\left(W^{L-1}\cdot\ldots\cdot f\left(W^1\cdot x+b^1\right)+\ldots+b^{L-1}\right)+b^L)$	$\hat{y}$ is the predicted output; f is the activation function; W^i and b^i are weights and biases at layer i; L is the number of layers; x is the input
e. Extremely Randomized Trees (XRT)	$\hat{y}=\frac{1}{T}\sum_{t=1}^{T}h_t(x)$	Similar to DRF but splits in each tree are chosen at random
f. Stochastic Gradient Boosting (SGB)	$F_m(x)=F_{m-1}(x)+v\cdot h_m(x;\theta_m)$	Similar to GBM however it uses that uses subsampling at each iteration to reduce variance and improve generalization. $h_m(x;\theta_m)$ is the weak learner fitted on a random subset of the data
g. Stacked Ensembles	$\hat{y}=g\left(h_1(x),h_2(x),\ldots,h_n(x)\right)$	$\hat{y}$ is the final prediction; g is the meta-learner; are predictions of the i-th base learner; n is the number of the base learners

the vector $\hat{b}_h$ represents h-step-ahead forecasts for each individual series in an m-dimensional space. The R package "hts" will be loaded in order to generate forecasts for the models shown in Table 2.

2.3. Forecasting measurement

A forecast error refers to the disparity between an observed value and its predicted value, representing the unpredictable component of the observation (Smolensky et al., 2013). The expression can be formulated as follows: the training data consist of $\{y_1, \ldots, y_T\}$, while the test data consist of $\{y_{T+1}, y_{T+2}, \ldots\}$. Forecast errors are computed from the test set.

$$e_{T+h} = y_{T+h} - \hat{y}_{T+h|T}. \tag{7}$$

In this scenario, the forecast errors are calculated for multi-step forecasts with a horizon of h = 3. Mean Absolute Percentage Error (MAPE) metric would be considered. The reason of choosing this error metric is due to its independence from scaling and sensitivity to errors. MAPE is a comparative metric that is not influenced by the magnitude of the data. This enables the comparison of different models on the same dataset without being affected by the magnitude of the real values and susceptibility to errors. Additionally, MAPE ensures equal treatment of both overestimations and underestimations by utilizing the absolute values of errors. This mitigates the problem of positive and negative errors nullifying each other, a phenomenon that can occur with alternative metrics such as Mean Error.

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right| \cdot 100, \tag{8}$$

where n is the number of observations, A_t is the observed value at time t, while F_t is the forecasted value at time t. R programming language (Team, 2020) would be the main software tool to be used.

3. RESULTS

3.1. Exploratory analysis

The temporal top nine causes of death for each gender and age partition over the years 2000 to 2019 are presented in Figures 2-5.

Figure 6 represents the tenth cause, averaged over the remaining causes of death and labeled as others across the age and gender groups. The lowest value for this indicator is observed across years, which means that, on average, causes other than the nine main ones had a smaller impact in the study, taking into account time, age, and gender.

The trend change points were determined by structural trend change models averaged over the

Figure 2. Top nine causes of death for males over 60 years of age

http://dx.doi.org/10.21511/ins.16(1).2025.02

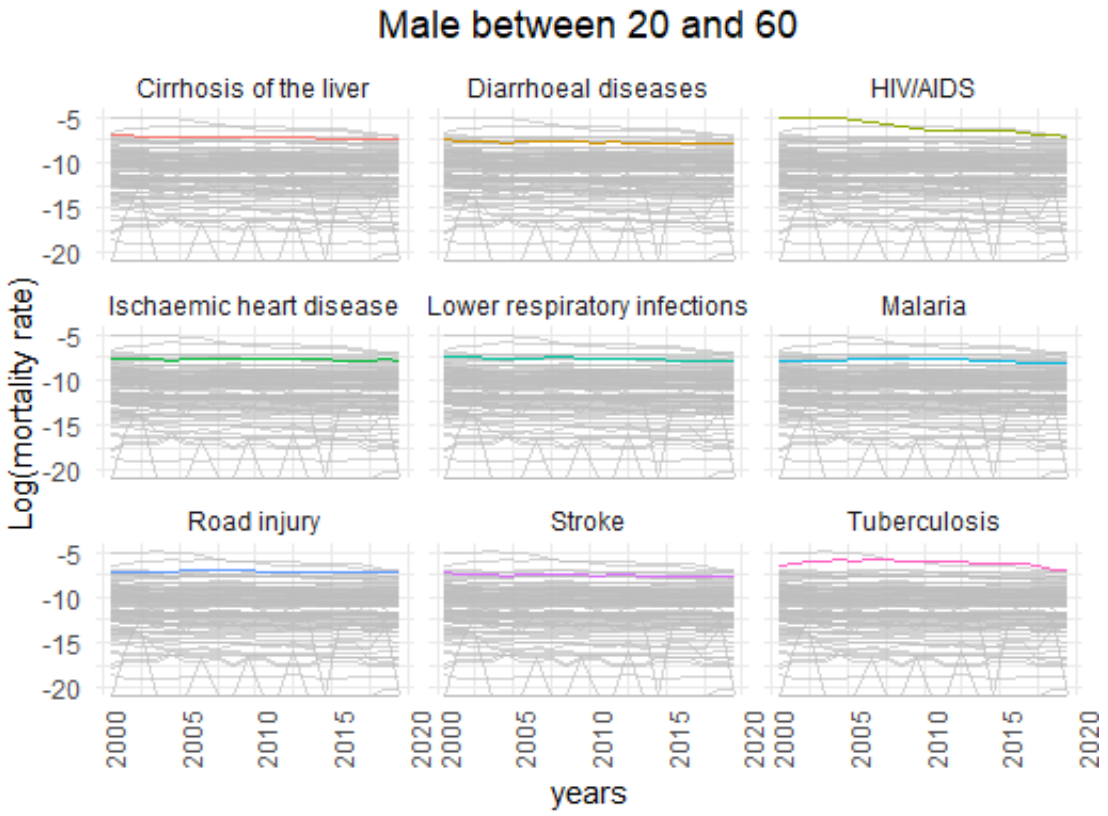


Figure 3. Top nine causes of death for males aged 20 to 60 years

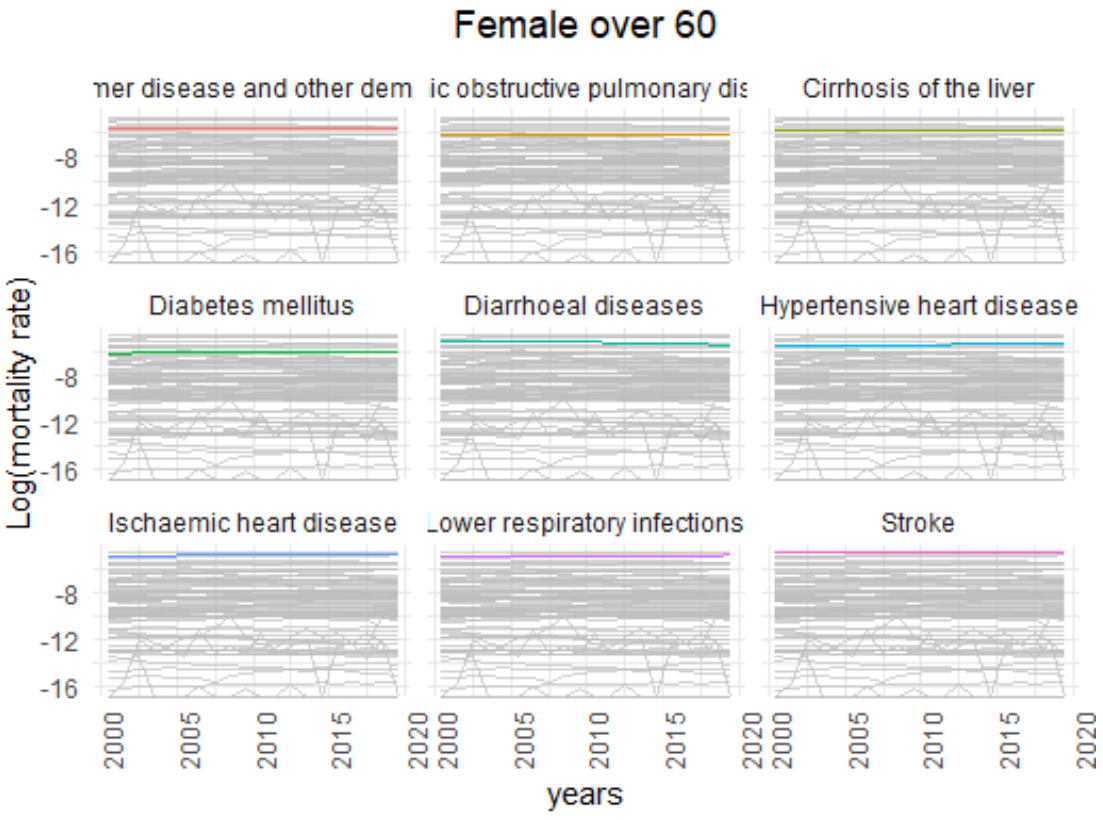


Figure 4. Top nine causes of death for females over 60 years of age

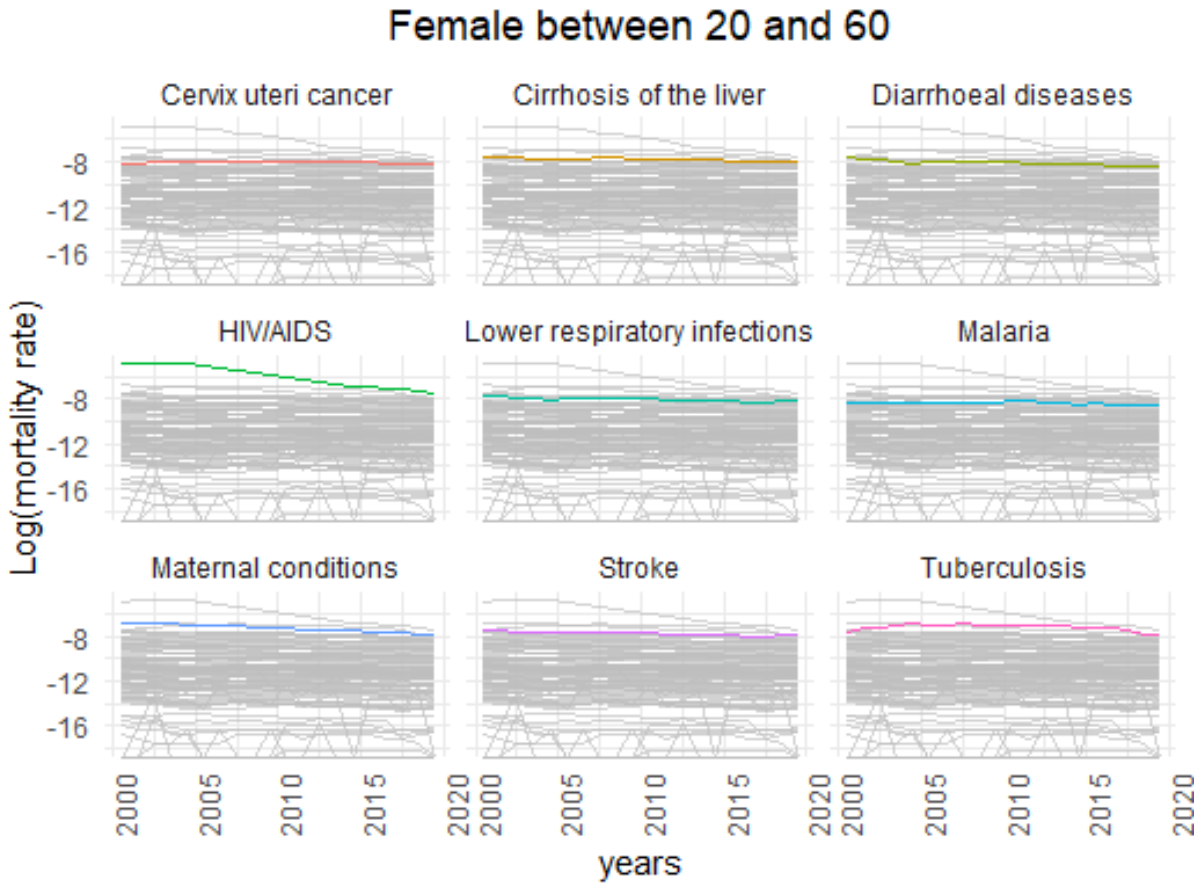


Figure 5. Top nine causes of death for females aged 20 to 60 years

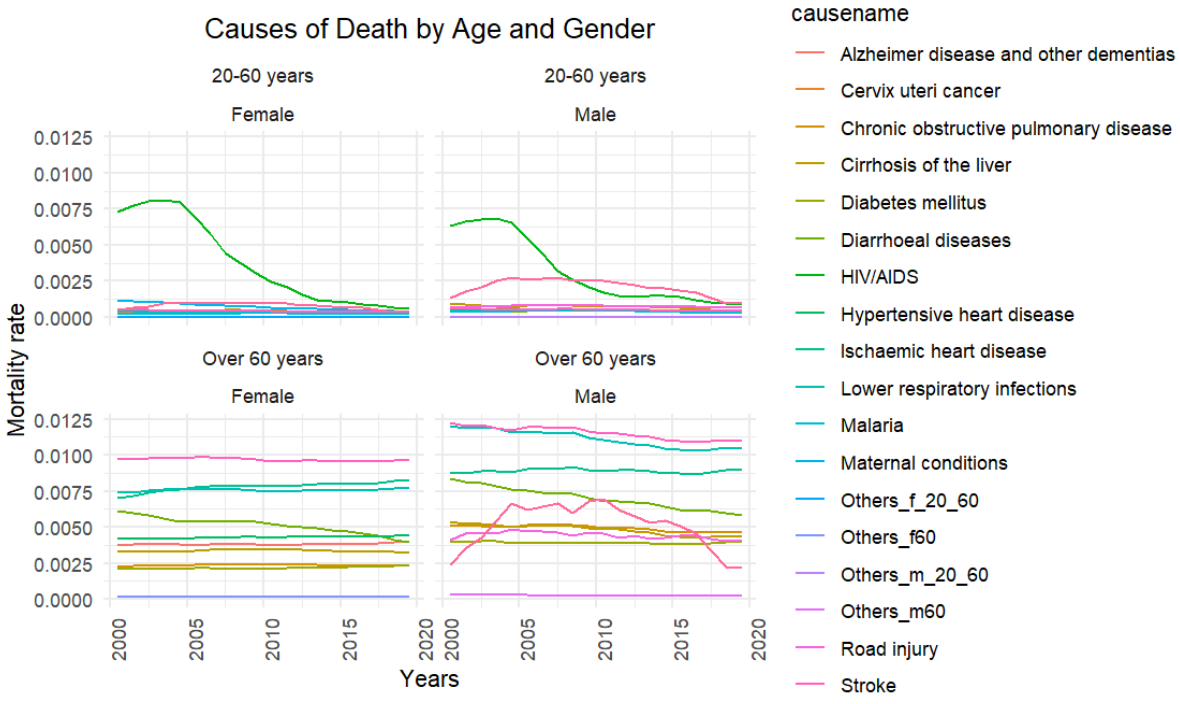


Figure 6. Other causes of death for males and females by age group

http://dx.doi.org/10.21511/ins.16(1).2025.02

segment breaks. Individual trend breaks were determined for all causes of death, and the breaks ranged between 2007 and 2011. Appendix A visually represents each trend break point for each cause of death.

Subsequently, the hierarchical forecasts for the test periods 2017, 2018, and 2019 based on the bottom-up approach were presented in Table 3 and Figure 7, that is, the distribution of mean absolute percentage error (MAPE) scores across the years pre

Table 3. Model results based on MAPE grouped by the causes of death

Univariate	BP	Linear Model (Base)		Auto. Arima		ETS		Prophet		Auto Machine Learning	
Series		Before BP	After BP	Before BP	After BP	Before BP	After BP	Before BP	After BP	Before BP	After BP
hiv_aids_female_20_60_years	2008	51.26	291.52	14.1	8.59	109.82	51.26	289.85	35.14	47.87	23.39
maternal_conditions_female_20_60_years	2008	17.37	15.5	7.22	15.04	13.16	17.37	16.81	5.65	13.24	33.2
tuberculosis_female_20_60_years	2011	12.24	85.11	97.62	15.89	97.61	12.24	89.31	25.58	76.42	25.23
stroke_female_20_60_years	2009	7.05	3.84	2.23	7.05	4.3	7.05	5.02	8.82	10.49	6.12
cirrhosis_of_the_liver_female_20_60_years	2009	2.32	6.14	5.26	2.32	6.83	2.32	5.98	5.15	11.41	2.17
diarrhoeal_diseases_female_20_60_years	2009	6.43	11.16	16.02	6.43	3.67	6.43	10.02	4.52	32.52	8.15
lower_respiratory_infections_female_20_60_years	2009	5.21	4.17	4.41	5.21	5.21	5.21	5.61	6.26	11.65	13.21
cervix_uteri_cancer_female_20_60_years	2010	3.32	9.03	10.15	3.28	6.04	3.32	7.37	1.35	7.95	8.62
malaria_female_20_60_years	2009	7.33	13.64	19.07	7.33	4.5	7.33	14.55	4.99	5.68	6.45
other_causes_female_20_60_years	2009	1.46	3.17	1.4	1.46	4.73	1.46	3.45	2.95	5.96	3
stroke_female_over_60_years	2007	1.87	0.79	0.83	1.87	0.83	1.87	0.92	0.98	0.7	1.24
ischaemic_heart_disease_female_over_60_years	2009	3.94	0.69	2.41	3.94	2.67	3.94	0.45	0.47	8.6	1.49
lower_respiratory_infections_female_over_60_years	2010	0.49	1.78	1.9	0.49	1.9	0.49	1.79	5.25	0.58	3.9
diarrhoeal_diseases_female_over_60_years	2010	5.53	9.08	7.38	5.53	8.17	5.53	9	0.84	9.77	0.42
hypertensive_heart_disease_female_over_60_years	2008	0.14	0.61	0.19	0.13	0.61	0.14	0.77	1.95	0.23	0.17
alzheimer_disease_and_other_dementias_female_over_60_years	2009	1.09	3.88	1.13	1.18	3.62	1.09	2.62	1.66	3.02	1.13
cirrhosis_of_the_liver_female_over_60_years	2010	0.66	4.39	3.13	0.66	0.62	0.66	1.61	1.48	0.77	0.97
chronic_obstructive_pulmonary_disease_female_over_60_years	2009	0.19	3.37	1.71	0.49	0.55	0.19	0.7	3.09	0.3	0.37
diabetes_mellitus_female_over_60_years	2010	0.79	2.64	2.55	0.79	2.64	0.79	0.81	5.07	2.65	1.7
other_causes_female_over_60_years	2009	2.35	1.99	2.9	2.35	2.1	2.35	2.14	1.04	7.1	0.61
other_causes_male_over_60_years	2008	2.42	2.8	2.01	2.42	0.45	2.42	2.16	0.88	0.33	2.9
diabetes_mellitus_male_over_60_years	2009	1.68	3.44	2.9	1.68	3.44	1.68	3.65	1.65	3.43	0.57
road_injury_male_over_60_years	2010	2.26	4.38	9.08	2.26	6.97	2.26	5.23	2.04	10.62	2.23
chronic_obstructive_pulmonary_disease_male_over_60_years	2009	2.21	1.42	3.47	2.2	0.75	2.21	3.9	3.86	0.59	1.79
cirrhosis_of_the_liver_male_over_60_years	2009	3.17	1.11	4.49	3.17	1.64	3.17	7.89	2.88	1.22	4.05
tuberculosis_male_over_60_years	2010	22.81	136.44	129.35	22.81	118.6	22.81	144.47	7.82	98.03	7.78
diarrhoeal_diseases_male_over_60_years	2009	10.48	1.12	1.83	10.48	1.24	10.48	1.44	3.8	0.87	11.47
ischaemic_heart_disease_male_over_60_years	2009	1.53	1.28	1.55	1.56	2.67	1.53	3.77	3.88	2.58	1.5
lower_respiratory_infections_male_over_60_years	2008	4.64	3.09	3.87	4.41	3.55	4.64	5.45	2.11	1.51	3.24
stroke_male_over_60_years	2008	3.66	1.06	1.14	3.66	0.89	3.66	3.59	2.59	0.81	2.88
other_causes_male_20_60_years	2009	4.63	1.82	2.27	4.65	1.63	4.63	1.9	7.31	5.93	18.05
malaria_male_20_60_years	2009	1.7	27.5	33.75	1.7	4.57	1.7	28.19	4.05	5.98	1.61
ischaemic_heart_disease_male_20_60_years	2009	4.03	3.94	7.21	4.03	4.11	4.03	3.76	9.85	4.07	5.27
diarrhoeal_diseases_male_20_60_years	2009	9.33	1.57	3.01	9.33	5.38	9.33	3.12	9.56	25.34	8.03
lower_respiratory_infections_male_20_60_years	2009	7.09	3.26	4.3	7.1	1.93	7.09	4.71	11.07	6.06	6.69
stroke_male_20_60_years	2009	7.11	2.39	2.97	7.11	1.85	7.11	4	11.69	6.69	4.21
road_injury_male_20_60_years	2009	4.23	9.79	14.03	4.23	6.02	4.23	10.47	2.53	7.95	3
cirrhosis_of_the_liver_male_20_60_years	2009	5.49	2.93	4.4	5.49	1.82	5.49	4.75	9.22	19.44	8.87
tuberculosis_male_20_60_years	2011	13.29	93.84	114.93	16.7	114.92	13.29	99.41	26.36	86.69	12.05
hiv_aids_male_20_60_years	2007	39.13	183.73	21.39	28.57	23.8	39.13	184.79	26.13	71.06	23.52

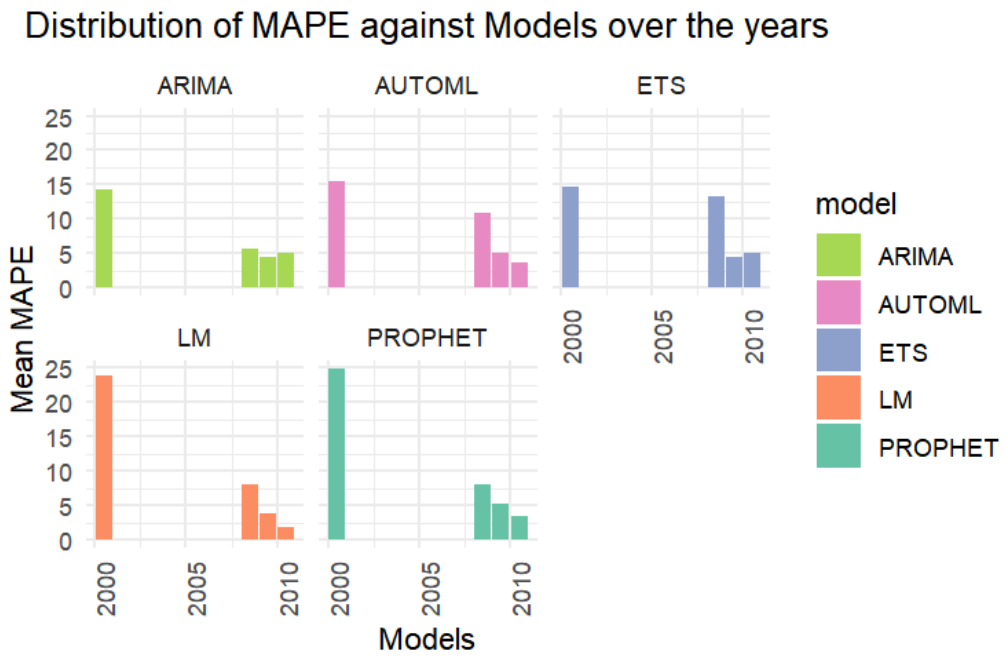


Figure 7. Distribution of MAPE against models over the years 2000–2019

(from 2000) and post (based on updated points from 2007 to 2011) and grouped by model. The year 2000 had the highest level of errors, possibly due to the presence of structural breaks in the entire dataset. However, 2008, 2009, and 2010 have reduced errors, possibly due to the update of the forecasts starting from the trend change points.

3.2. Forecasting accounting for trend breaks

Auto ML performed best in seven univariate series after accounting for trend breaks; they include cirrhosis of the liver for females aged 20 to 60 years, diarrhoeal diseases for females over 60 years, diabetes mellitus for males over 60 years, tuberculosis for males over 60 years, malaria for males 20 to 60 years, tuberculosis for males over 60 years, and other causes for females over 60 years, as shown in Table 3. Prophet performed best in four univariates after accounting for trend breaks; they include cervix for uteri cancer for females 20 to 60 years, road injury for males over 60 years, lower respiratory infections for males over 60 years, and road injury for males 20 to 60 years, as shown in Table 3. LM performed best in four univariates after accounting for trend breaks; they include lower respiratory infections for females aged 20 to 60 years, cirrhosis of the liver for males aged over 60 years, ischaemic heart disease for males aged over 60 years, and diarrheal diseases for males aged 20 to 60 years, as shown in Table 3. Auto Arima performed best in two univariates after accounting for trend breaks; they include HIV/AIDS for females aged 20 to 60 years old and hypertensive heart disease for females over 60 years old, as shown in Table 3. ETS performed poorly, shy of being the worst as compared to the base model after accounting for trend break, as shown in Table 3.

3.3. Forecasting without accounting for trend breaks

Exponential Smoothing (ETS) performed best in seven univariate analyses without accounting for trend breaks. They include diarrheal diseases for females aged 20 to 60 years, malaria for females aged 20 to 60 years, cirrhosis of the liver for females over 60 years, cirrhosis of the liver for males aged 20 to 60 years, lower respiratory infections for males aged 20 to 60 years, stroke for males aged 20 to 60 years, and other causes for males aged 20 to 60 years, as shown in Table 3. Auto ML performed best in five univariate series without accounting for trend breaks; they include stroke for females over 60 years, other causes for males over 60 years, chronic obstructive pulmonary disease for males over 60 years, diarrheal diseases for males over 60 years, and stroke for males over 60 years (see

http://dx.doi.org/10.21511/ins.16(1).2025.02

Table 3). Auto Arima performed best in four univariates without accounting for trend breaks; they include maternal conditions for females 20 to 60 years, stroke for females 20 to 60 years, others for females 20 to 60 years, and HIV/AIDS for females 20 to 60 years, as shown in Table 3. Prophet performed best in two univariates without accounting for trend breaks; they include ischaemic heart disease for females over 60 years and ischaemic heart disease for males aged 20 to 60 years, as shown in Table 3. LM performed best in one univariate without accounting for trend breaks; they include tuberculosis for females aged 20 to 60 years, as shown in Table 3.

However, other results were also obtained: auto-arima, ETS, and LM led with similar lowest scores for MAPE in four univariate series; they include lower respiratory infections for females over 60 years, Alzheimer's disease and other dementias for females over 60 years, chronic obstructive pulmonary disease for females over 60 years, and diabetes mellitus for females over 60 years (see Table 3).

4. DISCUSSIONS

The change in the cause of death trend confirms that each univariate series has a point where the accuracy score improves. At this realization point, the ultimate model update performs better. The auto-ML model performed the best after accounting for trend changes based on the number of series it was leading. Generally, the remaining models also performed well when applied to the entire dataset. However, the ETS model had no leading univariate forecast after the trend break; in fact, it performed the best together with LM over the entire dataset from 2000. It is noteworthy to suggest that ETS requires a longer historical dataset. Similar studies (Hyndman & Athanasopoulos, 2018; Petropoulos et al., 2013) have reached the same conclusion. Therefore, insufficient data do not necessarily imply deficient forecasts.

According to Bett et al. (2022), mortality shocks such as pandemics and famines introduce trend breaks and distort the temporal series either upwards or downwards. Such events are uncertain, inevitable, and ultimately affect total aggregate mortality. These results demonstrate that such scenarios would necessitate a customized approach to dealing with the causes of death that form the base forecasts. The bottom-up hierarchical forecast reconciliation approach is appropriate because it focuses on the base forecasts and coherently aggregates them into the total. As a result, actuaries should consider this approach as confirmed by (Van Berkum et al., 2016), especially while dealing with mortality shocks.

The forecasting models were implemented using three types of models categorized as base (LM), auto-statistical (Auto. Arima, ETS, and Prophet), and auto-machine learning (GBM, DRF, GLM, Neural Networks, XRT, SGB, and ensembles), which also contained deep learning approaches. Each cause of death yielded a unique set of optimal models across 40 univariate series separately. This is because the fitting and forecast were achieved using a bottom-up model, which is advantageous as it models from bottom-up based on individual causes and endeavors to apply all the available resources, which is consistent with the data and studies (Zhang et al., 2023). From the results among the models, the auto-machine learning models gave the best outcome, being the second when all the dataset was used and being the top after allowance for structural breaks were incorporated. Studies by Mancuso et al. (2021) and Abolghasemi et al. (2019) establish this conclusion. Other models implemented ranked highly also across the different years, meaning that none were redundant, however, this result shows that updating trend breaks is vital as it improves the overall result.

The linear model LM has been set as a base model to specifically track the trend component, which does not take into account any change points. These results confirm that the specific models are suitable for non-linear and linear data structures, and as seen, the total aggregate mortality is based on disaggregate causes of deaths; therefore, individual models would be viable to be modelled individually and progress collectively. Additionally, this study suggests that for specific causes of death, the length of the data is not necessarily a measure of quality data. The forecast accuracy of the 40-time series models demonstrates this. According to the findings, updating forecasts

based on the change point generally improves forecast accuracy. These results suggest that, to reduce fitting and forecasting errors, each series should undergo trend breaks and model specification prior to forecasting.

Due to the insufficiency of historical data, the forecast horizon can only be applied to shorter time intervals, preferably one to three years. For instance, it can match the annual actuarial valuation of life and pension schemes or even periodic life tables that span three years of experience in computation.

It is important to note that the years 2007 and 2011 have been omitted for the purpose of this study because they contain only one and two causes, respectively, which is inadequate for comparative studies. Therefore, future studies may consider investigating such a scenario.

CONCLUSION

The aim of this study is to determine and incorporate trend breaks among individual causes of death and coherently forecast them by applying the bottom-up hierarchical forecasting approach for life insurance models. The study employed a two-pronged approach: Firstly, it aimed to identify the individual series' structure and pinpoint the trend change point, achieved through a linear structure change. Secondly, it employed a bottom-up hierarchical forecast approach that maximized the base series. This approach forecasted the causes of death using appropriate models, including a base, auto-statistical, and auto-machine learning approach, all based on the mean absolute percentage error score. The comparative results were presented as pre- and post-trend change points, and generally, a reduction in mean absolute percentage error was observed across the models because each series showed different results. These results suggest that incorporating trend change generally improves forecast accuracy. However, it is important to note that each model performs differently, regardless of the length of data and the presence of mortality shocks. These studies contribute to the forecast reconciliation approach by incorporating trend breaks. Further research could consider taking into account uncertainties associated with trend breaks, such as bootstrapping, which measures levels of accuracy in cause-of-death prediction estimates.

AUTHOR CONTRIBUTIONS

Conceptualization: Nicholas Bett, Juma Kasozi, Daniel Ruturwa.
Data curation: Nicholas Bett, Juma Kasozi, Daniel Ruturwa.
Formal analysis: Nicholas Bett, Juma Kasozi, Daniel Ruturwa.
Funding acquisition: Juma Kasozi, Daniel Ruturwa.
Investigation: Nicholas Bett, Juma Kasozi, Daniel Ruturwa.
Methodology: Nicholas Bett, Juma Kasozi, Daniel Ruturwa.
Project administration: Juma Kasozi, Daniel Ruturwa.
Resources: Nicholas Bett, Juma Kasozi, Daniel Ruturwa.
Software: Nicholas Bett, Juma Kasozi, Daniel Ruturwa.
Supervision: Juma Kasozi, Daniel Ruturwa.
Validation: Nicholas Bett, Juma Kasozi, Daniel Ruturwa.
Visualization: Nicholas Bett.
Writing – original draft: Nicholas Bett.
Writing – review & editing: Nicholas Bett, Juma Kasozi, Daniel Ruturwa.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

http://dx.doi.org/10.21511/ins.16(1).2025.02

REFERENCES

1. Abolghasemi, M., Hyndman, R. J., Tarr, G., & Bergmeir, C. (2019). *Machine learning applications in time series hierarchical forecasting* (arXiv:1912.00370). https://doi.org/10.48550/arXiv.1912.00370
2. Arnold, S., & Glushko, V. (2021). Cause-specific mortality rates: Common trends and differences. *Insurance: Mathematics and Economics, 99,* 294-308. https://doi.org/10.1016/j.insmatheco.2021.03.027
3. Arnold, S., & Sherris, M. (2015). Causes-of-death mortality: What do we know on their dependence? *North American Actuarial Journal, 19*(2), 116-128. https://doi.org/10.1080/10920277.2015.1011279
4. Athanasopoulos, G., Ahmed, R. A., & Hyndman, R. J. (2009). Hierarchical forecasts for Australian domestic tourism. *International Journal of Forecasting, 25*(1), 146-166. https://doi.org/10.1016/j.ijforecast.2008.07.004
5. Bai, J. (1994). Least Squares Estimation of a Shift in Linear Processes. *Journal of Time Series Analysis, 15*(5), 453-472. https://doi.org/10.1111/j.1467-9892.1994.tb00204.x
6. Bai, J., & Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica, 66*(1), 47-78. https://doi.org/10.2307/2998540
7. Bai, J., & Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics, 18*(1), 1-22. https://doi.org/10.1002/jae.659
8. Bengtsson, T., & Keilman, N. (2019). *Old and new perspectives on mortality forecasting*. Springer Nature.
9. Bett, N., Kasozi, J., & Ruturwa, D. (2022). Temporal Clustering of the Causes of Death for Mortality Modelling. *Risks*, *10*(5), 99. https://doi.org/10.3390/risks10050099
10. Bett, N., Kasozi, J., & Ruturwa, D. (2023). Dependency Modeling Approach of Cause-Related Mortality and Longevity Risks: HIV/AIDS. *Risks, 11*(2), 38. https://doi.org/10.3390/risks11020038
11. Billah, B., King, M. L., Snyder, R. D., & Koehler, A. B. (2006). Exponential smoothing model selection for forecasting. *International Journal of Forecasting, 22*(2), 239-247. https://doi.org/10.1016/j.ijforecast.2005.08.002
12. Breiman, L. (2001). Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical Science, 16*(3), 199-231. https://doi.org/10.1214/ss/1009213726
13. Cairns, A. J., Blake, D., Dowd, K., Coughlan, G. D., Epstein, D., Ong, A., & Balevich, I. (2009). A quantitative comparison of stochastic mortality models using data from England and Wales and the United States. *North American Actuarial Journal, 13*(1), 1-35. https://doi.org/10.1080/10920277.2009.10597538
14. Caselli, G., Vallin, J., & Marsili, M. (2019). How useful are the causes of death when extrapolating mortality trends. An update. In Bengtsson, T., & Keilman, N. (Eds.), *Old and New Perspectives on Mortality Forecasting* (pp. 237-259). Demographic Research Monographs. Cham: Springer. https://doi.org/10.1007/978-3-030-05075-7_18
15. Chen, X., & Moraga, P. (2024). Assessing dengue forecasting methods: A comparative study of statistical models and machine learning techniques in Rio de Janeiro, Brazil. *medRxiv,* 2024-06. https://doi.org/10.1101/2024.06.12.24308827
16. Coelho, E., & Nunes, L. C. (2011). Forecasting mortality in the event of a structural change. *Journal of the Royal Statistical Society Series A: Statistics in Society, 174*(3), 713-736. https://doi.org/10.1111/j.1467-985X.2010.00687.x
17. Dayanik, S., Goulding, C., & Poor, H. V. (2008). Bayesian Sequential Change Diagnosis. *Mathematics of Operations Research, 33*(2), 475-496. https://doi.org/10.1287/moor.1070.0307
18. Fearnhead, P. (2005). Exact Bayesian curve fitting and signal segmentation. *IEEE Transactions on Signal Processing, 53*(6), 2160-2166. https://doi.org/10.1109/TSP.2005.847844
19. Freedman, D. A. (2009). *Statistical models: Theory and practice*. Cambridge University Press.
20. Giuseppi, A., & Pietrabissa, A. (2022). Bellman's principle of optimality and deep reinforcement learning for time-varying tasks. *International Journal of Control, 95*(9), 2448-2459. https://doi.org/10.1080/00207179.2021.1913516
21. Gross, C. W., & Sohl, J. E. (1990). Disaggregation methods to expedite product line forecasting. *Journal of Forecasting, 9*(3), 233-254. https://doi.org/10.1002/for.3980090304
22. Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica: Journal of the Econometric Society, 57*(2), 357-384. https://doi.org/10.2307/1912559
23. Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and practice*. OTexts. Retrieved from https://otexts.com/fpp3/
24. Hyndman, R. J., Ahmed, R. A., Athanasopoulos, G., & Shang, H. L. (2011). Optimal combination forecasts for hierarchical time series. *Computational Statistics & Data Analysis, 55*(9), 2579-2589. https://doi.org/10.1016/j.csda.2011.03.006
25. Killick, R., Eckley, I. A., Ewans, K., & Jonathan, P. (2010). Detection of changes in variance of oceanographic time-series using changepoint analysis. *Ocean Engineering*, *37*(13), 1120-1126. https://doi.org/10.1016/j.oceaneng.2010.04.009
26. LeDell, E., & Poirier, S. (2020). H2o autoML: Scalable automatic machine learning. *Proceedings of the AutoML Workshop at ICML, 2020*. Retrieved from https://www.automl.org/wp-content/uploads/2020/07/AutoML_2020_paper_61.pdf

27. Lee, R. D., & Carter, L. R. (1992). Modeling and Forecasting U.S. Mortality. *Journal of the American Statistical Association, 87*(419), 659-671. https://doi.org/10.1080/01621459.1992.10475265

28. Li, H., & Lu, Y. (2018). Modeling cause-of-death mortality using hierarchical Archimedean copula. *Scand. Actuarial Journal, 2019*(3), 247-272. https://doi.org/10.1080/03461238.2018.1546224

29. Mancuso, P., Piccialli, V., & Sudoso, A. M. (2021). A machine learning approach for forecasting hierarchical time series. *Expert Systems with Applications, 182,* 115102. https://doi.org/10.1016/j.eswa.2021.115102

30. Mélard, G., & Pasteels, J.-M. (2000). Automatic ARIMA modeling including interventions, using time series expert software. *International Journal of Forecasting, 16*(4), 497-508. https://doi.org/10.1016/S0169-2070(00)00067-4

31. Milidonis, A., Lin, Y., & Cox, S. H. (2011). Mortality Regimes and Pricing. *North American Actuarial Journal, 15*(2), 266-289. https://doi.org/10.1080/10920277.2011.10597621

32. Mullainathan, S., & Spiess, J. (2017). Machine learning: An applied econometric approach. *Journal of Economic Perspectives*, *31*(2), 87-106. https://doi.org/10.1257/jep.31.2.87

33. Perron, P. (1989). The great crash, the oil price shock, and the unit root hypothesis. *Econometrica: Journal of the Econometric Society, 57*(6), 1361-1401. https://doi.org/10.2307/1913712

34. Petropoulos, F., Nikolopoulos, K., Spithourakis, G. P., & Assimakopoulos, V. (2013). Empirical heuristics for improving intermittent demand forecasting. *Industrial Management & Data Systems*, *113*(5), 683-696. https://doi.org/10.1108/02635571311324142

35. Pettitt, A. N. (1979). A Non-Parametric Approach to the Change-Point Problem. *Applied Statistics*, *28*(2), 126-135. https://doi.org/10.2307/2346729

36. Richman, R. (2018). *AI in actuarial science*. Actuarial Science of South Africa. Retrieved from https://www.actuarialsociety.org.za/wp-content/uploads/2018/11/AI-IN-ACTUARIAL-SCIENCE-Ronald-Richman.pdf

37. Robertson, T., Batty, G. D., Der, G., Fenton, C., Shiels, P. G., & Benzeval, M. (2013). Is socioeconomic status associated with biological aging as measured by telomere length? *Epidemiologic Reviews, 35*(1), 98-111. https://doi.org/10.1093/epirev/mxs001

38. Smolensky, P., Mozer, M. C., & Rumelhart, D. E. (2013). *Mathematical perspectives on neural networks*. Psychology Press. https://doi.org/10.4324/9780203772966

39. Tang, S., Li, J., & Tickle, L. (2022). A New Fourier Approach under the Lee-Carter Model for Incorporating Time-Varying Age Patterns of Structural Changes. *Risks, 10*(8), 147. https://doi.org/10.3390/risks10080147

40. Taylor, S. J., & Letham, B. (2018). Forecasting at Scale. *The American Statistician*, *72*(1), 37-45. https://doi.org/10.1080/00031305.2017.1380080

41. Team, R. C. (2020). R: A language and environment for statistical computing, R Foundation for Statistical. *Computing*. Retrieved from https://cir.nii.ac.jp/crid/1370298755636824325

42. United Nations (UN). (2017). *World Population Prospects: The 2017 Revision Data Booklet (ST/ESA/SER. A/401)*. Retrieved from https://www.un.org/development/desa/pd/content/world-population-prospects-2017-revision-data-booklet

43. Van Berkum, F., Antonio, K., & Vellekoop, M. (2016). The impact of multiple structural changes on mortality predictions. *Scandinavian Actuarial Journal, 2016*(7), 581-603. https://doi.org/10.1080/03461238.2014.987807

44. Van Berkum, F., Antonio, K., Vellekoop, M., & Leuven, K. (2013). *Structural changes in mortality rates with an application to Dutch and Belgian data*. Retrieved from https://feb.kuleuven.be/drc/AFI/research/AFIInsuranceFolder/InsurancePapers/afi-1379.pdf

45. Waweru, M. N. (2014). *Determinants of insolvency in selected insurance companies in Kenya* (Ph.D. Thesis). University of Nairobi. Retrieved from http://erepository.uonbi.ac.ke/bitstream/handle/11295/95501/Waweru_Determinants%20of%20insolvency%20in%20selected%20insurance%20companies%20in%20Kenya.pd

46. Wickramasuriya, S. L., Athanasopoulos, G., & Hyndman, R. J. (2019). Optimal Forecast Reconciliation for Hierarchical and Grouped Time Series Through Trace Minimization. *Journal of the American Statistical Association, 114*(526), 804-819. https://doi.org/10.1080/01621459.2018.1448825

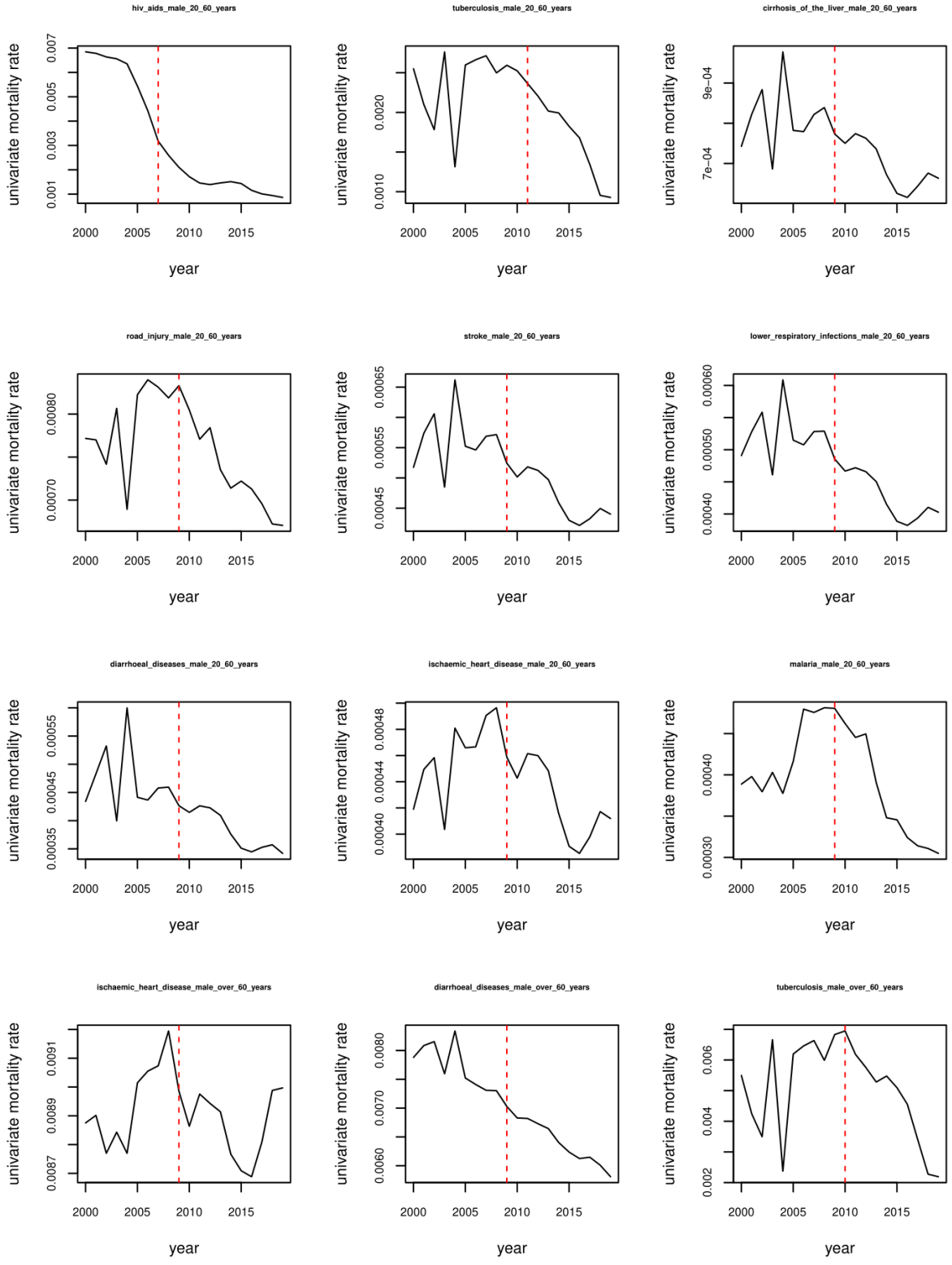
47. World Health Organization (WHO). (2022). *Global health estimates: Leading causes of death.* Retrieved from https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates/ghe-leading-causes-of-death

48. Zeileis, A., Kleiber, C., Krämer, W., & Hornik, K. (2003). Testing and dating of structural changes in practice. *Computational Statistics & Data Analysis, 44*(1-2), 109-123. https://doi.org/10.1016/S0167-9473(03)00030-6

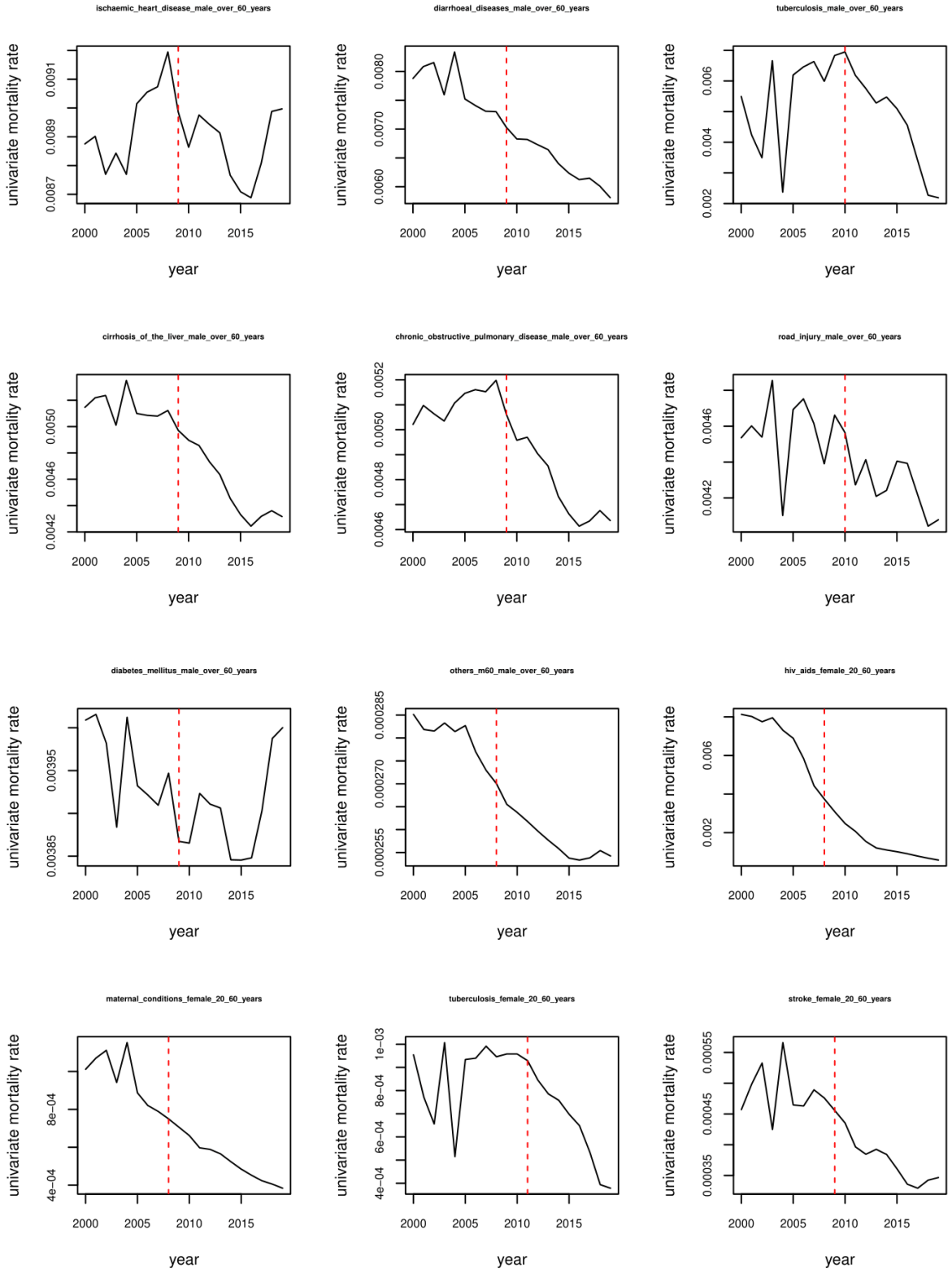
49. Zhang, B., Kang, Y., Panagiotelis, A., & Li, F. (2023). Optimal reconciliation with immutable forecasts. *European Journal of Operational Research, 308*(2), 650-660. https://doi.org/10.1016/j.ejor.2022.11.035

http://dx.doi.org/10.21511/ins.16(1).2025.02

APPENDIX A. Causes of death trend breaks based on 2000–2019 dataset

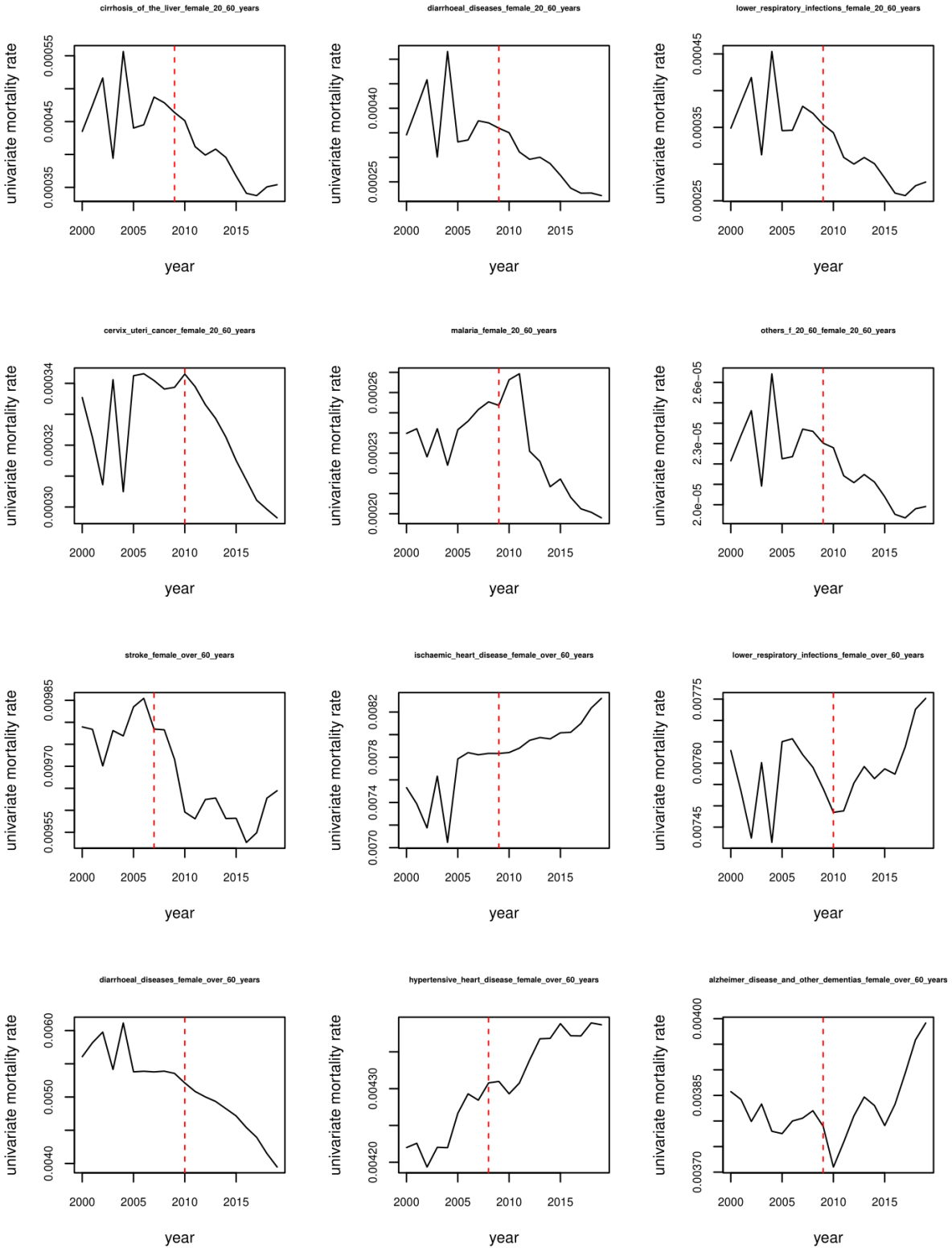


Figures A1. Plot of causes of death trend breaks based on 2000–2019 dataset

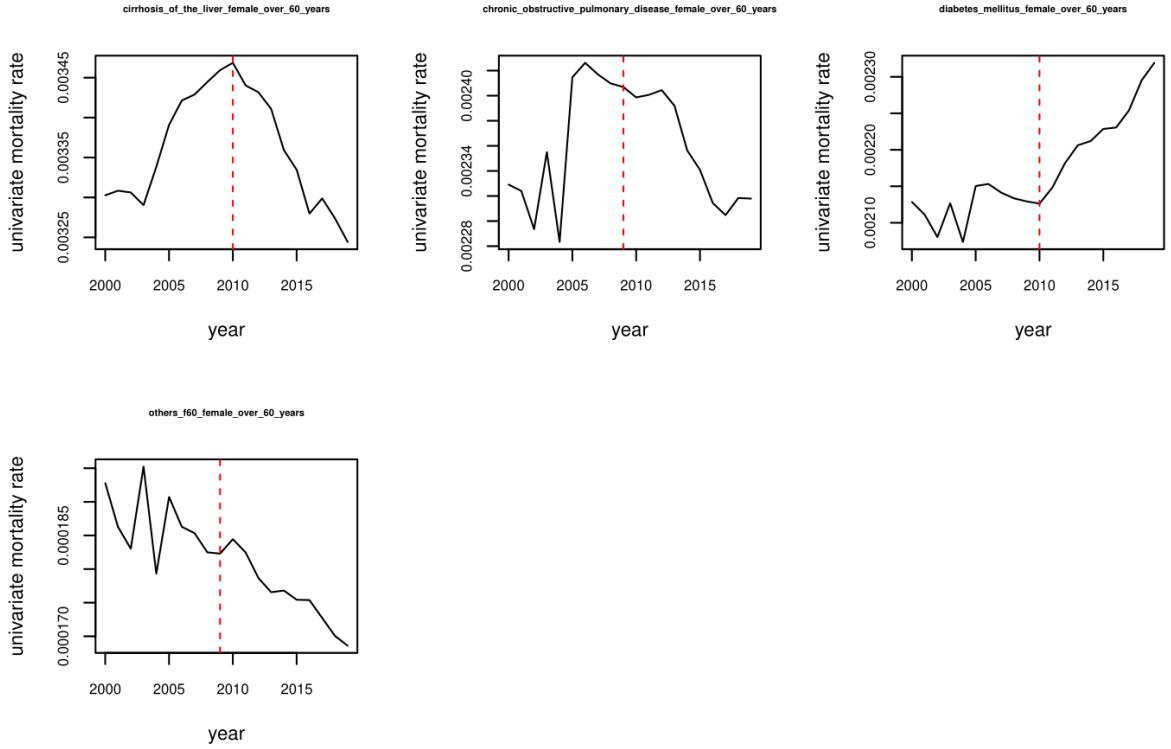


Figures A1 (cont.). Plot of causes of death trend breaks based on 2000–2019 dataset

http://dx.doi.org/10.21511/ins.16(1).2025.02



Figures A1 (cont.). Plot of causes of death trend breaks based on 2000–2019 dataset



Figures A1 (cont.). Plot of causes of death trend breaks based on 2000–2019 dataset

http://dx.doi.org/10.21511/ins.16(1).2025.02