Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A Dynamic Factor-Copula Framework for Cause-Specific Mortality Forecasting and Dependent Multiple Decrements: Evidence from U.S. Cause-of-Death Rates, 1979-2023

Anastasios-Tsampikos Statiou

statiou@gmail.com

University of the Aegean

Peter Hatzopoulos

University of Piraues

Research Article

Keywords: mortality forecasting, Lee-Carter, cause-specific mortality, multiple decrements, Gaussian copula, actuarial modeling

Posted Date: May 7th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9290989/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

A Dynamic Factor-Copula Framework for Cause-Specific Mortality Forecasting and Dependent Multiple Decrements: Evidence from U.S. Cause-of-Death Rates, 1979-2023

Abstract

This paper develops a dynamic factor-copula framework for coherent forecasting of all-cause and cause-specific mortality under a dependent multiple-decrement structure. The empirical application uses a U.S. cause-of-death panel covering 1979-2023, two sexes, 207 cause codes, and grouped ages from 0 to 100+. The model combines an all-cause Lee-Carter component, a cause-share factor layer estimated on logit-transformed decrement shares, and a Gaussian copula linking latent innovations across grouped causes. This architecture preserves aggregate coherence while allowing dependence across decrement-specific dynamics. After rescaling the source rates and regrouping causes into 13 macro-causes, forecasts were generated through 2070 using 500 stochastic trajectories. The results show numerically exact share and rate coherence, projected mortality declines for both sexes, and substantial long-horizon compositional change, with cardiovascular mortality losing relative weight while diabetes/metabolic, respiratory, and mental/neurological causes become more prominent. The copula layer reveals non-trivial latent cross-group dependence, especially involving respiratory mortality. Overall, the framework provides a tractable basis for coherent cause-specific forecasting, dependent decrement analysis, and future extensions toward richer dependence structures.

Keywords: mortality forecasting; Lee-Carter; cause-specific mortality; multiple decrements; Gaussian copula; actuarial modeling.

1. Introduction

Forecasting mortality remains central to actuarial science, demography, pension design, annuity valuation, public-health planning, and the measurement of longevity risk. The classical literature established a powerful baseline with parsimonious age-period structures, beginning from the Lee-Carter model and developing into broader single-population and multi-population frameworks (Lee & Carter, 1992; Booth & Tickle, 2008; Cairns et al., 2006; Hyndman & Ullah, 2007; Plat, 2009). Over time, the literature has also made clear that no single mortality model dominates across all countries, ages, and forecast horizons, which is why specification risk and model uncertainty remain central concerns (Booth & Tickle, 2008; Basellini et al., 2023; Diana et al., 2025).

Recent work has moved in several directions at once. One line of research seeks greater coherence, robustness, and control of forecast divergence across ages or populations (Chang & Shi, 2023; Richards, 2024; Scognamiglio, 2024). Another line introduces more flexible nonlinear structures

through kernel methods, regime-switching, mixed-frequency updates, boosting, and neural networks (Fu et al., 2023; Li, 2023; Wu et al., 2024; Li et al., 2025; Li et al., 2026). A third line emphasizes interpretability under richer settings, including multi-population and cause-specific mortality, rather than treating mortality data as generic machine-learning inputs (Debón et al., 2025; Euthum et al., 2024; De Mori et al., 2025; Zheng et al., 2025a).

The recent deep-learning literature is especially instructive because it highlights both promise and caution. Review work argues that architecture must remain aligned with actuarial structure, while empirical studies show improved forecasting when neural networks respect age-period smoothness, multi-task information sharing, and explainable latent representations (Euthum et al., 2024; De Mori et al., 2025; Zheng et al., 2025a, 2025b). For cause-of-death modeling specifically, interpretable neural-network and tensor approaches have been proposed precisely because simple independent forecasting of each cause is often too weak, whereas high-rank purely data-driven decompositions can lose transparency (Tanaka & Matsuyama, 2025).

Cause-specific mortality is difficult because the model must forecast multiple decrements while preserving consistency with the all-cause total and recognizing that causes co-move rather than evolve independently. Earlier solutions include cause-elimination and share-based structures, whereas recent work has added dependence-aware and machine-learning approaches (Alai et al., 2015; Bett et al., 2023; Lei et al., 2023; Tanaka & Matsuyama, 2025; Asplund et al., 2025). However, aggregate cause-specific mortality data do not reveal individual latent lifetimes under competing risks. Hence, one should not claim to identify true individual-level dependence from grouped mortality tables alone. What can be modeled, and what this paper models, is a latent dependence structure in aggregate cause-specific mortality shocks.

The purpose of the present paper is to develop a cautious but useful middle ground between overly restrictive independent decrement models and opaque black-box forecasting systems. The proposed framework keeps the all-cause Lee-Carter logic, models cause-specific shares through an interpretable factor layer, and introduces dependence by means of a Gaussian copula acting on latent decrement dynamics. In this way, the model remains coherent by construction, transparent in interpretation, and open to future generalization toward richer copulas or Bayesian estimation. Its contribution is not a new copula family or a new mortality estimator in isolation; rather, it is the coherent integration of aggregate mortality forecasting, compositional cause-share dynamics, and dependence-aware simulation within a single empirical architecture.

Empirically, the paper uses the supplied U.S. cause-of-death mortality panel for 1979-2023. After preprocessing and regrouping the raw causes into 13 macro-causes, the model generates forecasts through 2070 and derives all-cause and decrement-specific probabilities. The contribution is therefore twofold: first, to show that a relatively simple hierarchical architecture can deliver coherent all-cause and cause-specific forecasts; second, to show how a copula layer can summarize latent co-movement across decrement blocks in a way that is analytically tractable and empirically interpretable. The paper also adds a simple holdout check for the aggregate forecasting backbone so that coherence is complemented by a limited but explicit predictive validation step.

2. Methods

2.1 Data source, empirical setting, and preprocessing strategy

The empirical analysis was based on annual age-specific and cause-specific mortality data for the United States covering the period 1979–2023 and stratified by sex. The database contained central death rates by age, calendar year, sex, and cause of death. Let $x$denote age, $t$calendar year, $s$sex, and $j$the decrement or cause-of-death group. The fundamental quantity in the analysis was the central death rate

$$m^{(j,s)}(x,t),$$

which represents the age-specific mortality intensity for decrement j, sex s, at age $x$and year t. The corresponding all-cause mortality rate is denoted by

$$m^{(\mathrm{all},s)}(x,t).$$

The data preparation stage was not a trivial technical step but a necessary statistical component of the modelling framework. First, the raw mortality values were recorded on a per-100,000 scale. If such values were used directly in a forecasting and actuarial conversion framework, the resulting mortality probabilities would be severely overstated. Therefore, if $m_{\mathrm{raw}}(x,t)$denotes the observed mortality quantity in the original files, the working mortality rate used in all subsequent calculations was defined as

$$m(x,t) = \frac{m_{\mathrm{raw}}(x,t)}{100000}.$$

This rescaling ensured that all quantities were expressed on the natural mortality-rate scale. In practical terms, this correction was crucial because the later transformation from central death rates to interval mortality probabilities depends directly on the level of $m(x,t)$.

Second, the raw database contained overlapping age notations in the older ages, such as closed age intervals and plus-coded intervals referring to the same age boundary. If retained simultaneously, those variables induce duplication in the age-year surface and contaminate the decomposition of all-cause mortality into decrement-specific components. For this reason, the age structure was cleaned in a way that preserved the closed age intervals while retaining only the terminal open-ended category when required. In addition, because the open-ended highest age group tends to distort visual interpretation and sometimes destabilises long-horizon extrapolation, the substantive presentation of the results was restricted to ages up to 85. This choice did not alter the mathematical structure of the model; rather, it improved interpretability and prevented the open-ended terminal age interval from dominating the empirical discussion.

Third, the original detailed cause-of-death codes were aggregated into broader decrement groups. This aggregation was necessary because a highly disaggregated cause structure is statistically

unstable and tends to produce sparse, noisy, and weakly interpretable mortality surfaces. The grouped structure adopted in the final empirical implementation is summarised in Table 1.

Table 1. Cause-of-death groups used in the empirical decrement framework

Group	Interpretation
Cardiovascular diseases	Circulatory and vascular mortality causes
Neoplasms	Malignant and non-malignant neoplastic causes
Respiratory diseases	Chronic and acute respiratory causes
External causes	Injuries, accidents, violence, and related causes
Mental and neurological disorders	Dementia, neurological, and related mental causes
Infectious diseases	Infectious and parasitic causes
Digestive diseases	Diseases of the digestive system
Genitourinary diseases	Urinary and reproductive system causes
Endocrine and metabolic diseases	Diabetes and related endocrine/metabolic causes
Maternal and perinatal conditions	Maternal, neonatal, and perinatal causes
Congenital conditions	Congenital anomalies and related causes
Other non-communicable diseases	Remaining chronic disease categories
Other	Residual causes not otherwise classified

This grouping reflects a compromise between epidemiological meaning and statistical stability. The objective was not to reconstruct the entire fine-grained cause-of-death taxonomy, but to build a forecast system that is coherent, interpretable, and sufficiently robust for long-horizon projection.

2.2 Coherent decomposition of all-cause mortality

The core idea of the modelling strategy was that all-cause mortality should act as the aggregate anchor of the system, while decrement-specific mortality should be modelled as a coherent allocation of the all-cause total across cause groups. This was implemented through the identity

$$m^{(j,s)}(x,t) = m^{(\text{all},s)}(x,t)\, w^{(j,s)}(x,t),$$

where $w^{(j,s)}(x,t)$ denotes the share of all-cause mortality attributable to decrement group j. By definition,

$$0 \le w^{(j,s)}(x,t) \le 1, \sum_{j=1}^{J} w^{(j,s)}(x,t) = 1$$

for every age x, year t, and sex s. Hence the decomposition automatically implies

$$\sum_{j=1}^{J} m^{(j,s)}(x,t) = m^{(\text{all},s)}(x,t) \sum_{j=1}^{J} w^{(j,s)}(x,t) = m^{(\text{all},s)}(x,t).$$

This structure was deliberately chosen instead of independent cause-by-cause forecasting. If each cause were projected separately without an aggregation constraint, the resulting forecasts would generally fail to add up to all-cause mortality, thereby generating an incoherent mortality system. In contrast, the present framework imposes coherence by construction.

The decomposition can also be written in vector form. Define the decrement share vector

$$\mathbf{w}^{(s)}(x,t) = \left(w^{(1,s)}(x,t), \dots, w^{(J,s)}(x,t)\right)'.$$

Then $\mathbf{w}^{(s)}(x,t)$belongs to the simplex

$$\mathcal{S}^{J-1} = \left\{ \mathbf{w} \in \mathbb{R}^J : w_j \ge 0, \sum_{j=1}^{J} w_j = 1 \right\}.$$

Thus the empirical task was naturally divided into two coupled components: forecasting the all-cause mortality surface $m^{(\text{all},s)}(x,t)$, and forecasting the cause-share surface $\mathbf{w}^{(s)}(x,t)$subject to the simplex constraint.

2.3 All-cause mortality model

The first layer of the system modelled all-cause mortality separately for males and females using a Lee–Carter-type age-period factor specification. For each sex s,

$$\log\, m^{(\text{all},s)}(x,t) = a_x^{(s)} + b_x^{(s)} \kappa_t^{(s)} + \varepsilon^{(s)}(x,t).$$

In this formulation, $a_x^{(s)}$captures the average age profile of mortality, $b_x^{(s)}$measures how strongly age $x$responds to period variation, and $\kappa_t^{(s)}$is the latent period index summarising the time trend in mortality. The disturbance term $\varepsilon^{(s)}(x,t)$collects residual variation not captured by the rank-one factor structure.

The model is not identified without constraints because the factor loading and the period index can be rescaled jointly. To ensure uniqueness, the standard identification conditions were imposed:

$$\sum_x b_x^{(s)} = 1, \sum_t \kappa_t^{(s)} = 0.$$

The average age profile was estimated as

$$\hat{a}_x^{(s)} = \frac{1}{T}\sum_{t=1}^{T} \log\, m^{(\text{all},s)}(x,t),$$

where T is the number of years in the sample. The centred age-year matrix was then constructed as

$$Y_{x,t}^{(s)} = \log\, m^{(\text{all},s)}(x,t) - \hat{a}_x^{(s)}.$$

This matrix was approximated by a rank-one representation,

$$Y_{x,t}^{(s)} \approx b_x^{(s)} \kappa_t^{(s)},$$

through singular value decomposition. If

$$\mathbf{Y}^{(s)} = \mathbf{U}^{(s)}\mathbf{D}^{(s)}(\mathbf{V}^{(s)})',$$

then the first singular triplet was used to obtain

$$\hat{b}_x^{(s)} = u_{x1}^{(s)}, \hat{\kappa}_t^{(s)} = d_1^{(s)} v_{t1}^{(s)},$$

followed by the re-normalisation required to satisfy the identification constraints. More explicitly, if

$$\tilde{b}_x^{(s)} = u_{x1}^{(s)}, \tilde{\kappa}_t^{(s)} = d_1^{(s)} v_{t1}^{(s)},$$

then the identified quantities were

$$\hat{b}_x^{(s)} = \frac{\tilde{b}_x^{(s)}}{\sum_x \tilde{b}_x^{(s)}}, \hat{\kappa}_t^{(s)} = \left(\sum_x \tilde{b}_x^{(s)}\right)\tilde{\kappa}_t^{(s)} - \frac{1}{T}\sum_{u=1}^{T}\left(\sum_x \tilde{b}_x^{(s)}\right)\tilde{\kappa}_u^{(s)}.$$

The period index was then extrapolated by a random walk with drift,

$$\kappa_t^{(s)} = \kappa_{t-1}^{(s)} + d^{(s)} + \eta_t^{(s)}, \eta_t^{(s)} \sim \mathcal{N}(0, \sigma_{\kappa,s}^2),$$

where the drift term was estimated by

$$\hat{d}^{(s)} = \frac{\hat{\kappa}_T^{(s)} - \hat{\kappa}_1^{(s)}}{T-1}.$$

For a forecast horizon h, the point forecast of the period index is therefore

$$\hat{\kappa}_{T+h}^{(s)} = \hat{\kappa}_T^{(s)} + h\hat{d}^{(s)}.$$

Substituting into the Lee–Carter equation gives the all-cause mortality forecast

$$\hat{m}^{(\text{all},s)}(x, T+h) = \exp\left(\hat{a}_x^{(s)} + \hat{b}_x^{(s)}\hat{\kappa}_{T+h}^{(s)}\right).$$

The forecast horizon in the empirical application extended from 2024 to 2070.

Table 2. Summary of the all-cause mortality forecasting layer

Component	Mathematical role
$a_x^{(s)}$	Baseline age mortality pattern
$b_x^{(s)}$	Age-specific sensitivity to period change
$\kappa_t^{(s)}$	Common period mortality index
$\varepsilon^{(s)}(x,t)$	Residual age-year variation
Forecast equation	$\hat{m}^{(\mathrm{all},s)}(x,t) = \exp\left(\hat{a}_x^{(s)} + \hat{b}_x^{(s)}\hat{\kappa}_t^{(s)}\right)$

The purpose of this all-cause layer was not to exhaust every demographic nuance, but to create a stable, interpretable, and coherent aggregate backbone for the decrement system.

2.4 Construction of decrement shares

The second layer allocated the all-cause mortality rate across decrement groups. For each age x, year t, sex s, and decrement group j, the observed cause-share was defined as

$$w^{(j,s)}(x,t) = \frac{m^{(j,s)}(x,t)}{m^{(\mathrm{all},s)}(x,t)}.$$

Because shares are bounded and compositional, they cannot be modelled reliably with an unconstrained linear Gaussian specification on the raw scale. A transformation was therefore required. The implementation used a logit transform applied group by group:

$$z^{(j,s)}(x,t) = \log\left(\frac{w^{(j,s)}(x,t)}{1 - w^{(j,s)}(x,t)}\right).$$

This transformation maps the open unit interval to the real line and stabilises the modelling of the share surface. For numerical stability, the shares were bounded away from exactly zero and one before transformation, i.e.

$$w_\epsilon^{(j,s)}(x,t) = \min\left\{\max\left[w^{(j,s)}(x,t), \epsilon\right], 1-\epsilon\right\},$$

with a very small positive constant ϵ. The transformed value was then

$$z^{(j,s)}(x,t) = \log\left(\frac{w_\epsilon^{(j,s)}(x,t)}{1 - w_\epsilon^{(j,s)}(x,t)}\right).$$

This step was essential because extremely small shares at certain ages and causes would otherwise create undefined logarithms.

2.5 Factor model for transformed decrement shares

For each decrement group $j$and sex s, the transformed share surface was modelled by an age-period factor structure analogous to the all-cause layer:

$$z^{(j,s)}(x,t) = \alpha_x^{(j,s)} + \beta_x^{(j,s)}\gamma_t^{(j,s)} + \xi^{(j,s)}(x,t).$$

Here, $\alpha_x^{(j,s)}$is the average age profile of the transformed share, $\beta_x^{(j,s)}$is the age-specific sensitivity to time variation in the share, $\gamma_t^{(j,s)}$is the latent period factor for decrement j, and $\xi^{(j,s)}(x,t)$is the residual term.

The average age profile of the transformed share was estimated as

$$\hat{\alpha}_x^{(j,s)} = \frac{1}{T}\sum_{t=1}^{T} z^{(j,s)}(x,t).$$

The centred matrix

$$Z_{x,t}^{(j,s)} = z^{(j,s)}(x,t) - \hat{\alpha}_x^{(j,s)}$$

was again approximated by rank one,

$$Z_{x,t}^{(j,s)} \approx \beta_x^{(j,s)}\gamma_t^{(j,s)}.$$

Applying singular value decomposition,

$$\mathbf{Z}^{(j,s)} = \mathbf{U}^{(j,s)}\mathbf{D}^{(j,s)}(\mathbf{V}^{(j,s)})',$$

and retaining the first singular component yielded preliminary estimates

$$\tilde{\beta}_x^{(j,s)} = u_{x1}^{(j,s)}, \tilde{\gamma}_t^{(j,s)} = d_1^{(j,s)}v_{t1}^{(j,s)}.$$

These were identified by imposing

$$\sum_x \beta_x^{(j,s)} = 1, \sum_t \gamma_t^{(j,s)} = 0,$$

which leads to

$$\hat{\beta}_x^{(j,s)} = \frac{\tilde{\beta}_x^{(j,s)}}{\sum_x \tilde{\beta}_x^{(j,s)}}, \hat{\gamma}_t^{(j,s)} = \left(\sum_x \tilde{\beta}_x^{(j,s)}\right)\tilde{\gamma}_t^{(j,s)} - \frac{1}{T}\sum_{u=1}^{T}\left(\sum_x \tilde{\beta}_x^{(j,s)}\right)\tilde{\gamma}_u^{(j,s)}.$$

The time factor $\gamma_t^{(j,s)}$was projected by a random walk with drift,

$$\gamma_t^{(j,s)} = \gamma_{t-1}^{(j,s)} + c^{(j,s)} + \nu_t^{(j,s)}, \nu_t^{(j,s)} \sim \mathcal{N}(0, \sigma_{\gamma,j,s}^2),$$

with drift estimate

$$\hat{c}^{(j,s)} = \frac{\hat{\gamma}_T^{(j,s)} - \hat{\gamma}_1^{(j,s)}}{T-1}.$$

Hence, for horizon h,

$$\hat{\gamma}_{T+h}^{(j,s)} = \hat{\gamma}_T^{(j,s)} + h\hat{c}^{(j,s)}.$$

The transformed-share forecast then becomes

$$\hat{z}^{(j,s)}(x, T+h) = \hat{\alpha}_x^{(j,s)} + \hat{\beta}_x^{(j,s)}\hat{\gamma}_{T+h}^{(j,s)}.$$

2.6 Inverse transformation and renormalisation of shares

The transformed-share forecasts were mapped back to the share scale by the inverse logit function:

$$\tilde{w}^{(j,s)}(x,t) = \frac{\exp\left[\hat{z}^{(j,s)}(x,t)\right]}{1 + \exp\left[\hat{z}^{(j,s)}(x,t)\right]}.$$

However, because the inverse logit was applied separately to each decrement group, the resulting provisional shares do not necessarily satisfy

$$\sum_{j=1}^{J} \tilde{w}^{(j,s)}(x,t) = 1.$$

Therefore, a renormalisation step was introduced:

$$\hat{w}^{(j,s)}(x,t) = \frac{\tilde{w}^{(j,s)}(x,t)}{\sum_{k=1}^{J} \tilde{w}^{(k,s)}(x,t)}.$$

By construction,

$$\sum_{j=1}^{J} \hat{w}^{(j,s)}(x,t) = 1.$$

This renormalisation is not a cosmetic correction. It is the mathematical step that restores compositional coherence and ensures that the forecast decrement shares remain interpretable as a partition of all-cause mortality.

Table 3. Summary of the decrement-share modelling layer

Component	Mathematical role

Component	Mathematical role
$w^{(j,s)}(x,t)$	Observed decrement share
$z^{(j,s)}(x,t)$	Logit-transformed share
$\alpha_x^{(j,s)}$	Mean age profile of transformed share
$\beta_x^{(j,s)}$	Age loading for transformed share dynamics
$\gamma_t^{(j,s)}$	Period factor for decrement share
$\hat{w}^{(j,s)}(x,t)$	Forecast share after inverse logit and renormalisation

2.7 Reconstruction of decrement-specific mortality forecasts

Having obtained forecasts for the all-cause mortality surface and the cause-share surface, the decrement-specific mortality rates were reconstructed as

$$\hat{m}^{(j,s)}(x,t) = \hat{m}^{(\text{all},s)}(x,t)\,\hat{w}^{(j,s)}(x,t).$$

This equation is the key synthesis step of the framework. It yields a full age-year-sex-cause mortality forecast while preserving coherence. The aggregate identity follows directly:

$$\sum_{j=1}^{J} \hat{m}^{(j,s)}(x,t) = \hat{m}^{(\text{all},s)}(x,t) \sum_{j=1}^{J} \hat{w}^{(j,s)}(x,t) = \hat{m}^{(\text{all},s)}(x,t).$$

Thus the model is coherent both historically and in forecast.

A useful alternative representation is to define the vector of decrement-specific forecasts

$$\hat{\mathbf{m}}^{(s)}(x,t) = \left(\hat{m}^{(1,s)}(x,t), \ldots, \hat{m}^{(J,s)}(x,t)\right)',$$

so that

$$\hat{\mathbf{m}}^{(s)}(x,t) = \hat{m}^{(\text{all},s)}(x,t)\,\hat{\mathbf{w}}^{(s)}(x,t).$$

This highlights that the entire cause-of-death system is generated by the scalar all-cause forecast and the simplex-valued share forecast.

2.8 Latent dependence across decrement innovations

An important objective of the study was to move beyond purely independent decrement forecasting and allow dependence across the cause-specific forecast innovations. However, the available data

were aggregate age-year-cause mortality rates rather than individual competing-risk histories. Therefore, individual-level dependence among decrements is not directly identifiable from the data. For this reason, dependence was introduced at the level of latent shocks in the decrement-share dynamics.

Let the innovation vector for sex $s$at time $t$be defined as

$$\mathbf{u}_t^{(s)} = \left(u_t^{(1,s)}, u_t^{(2,s)}, \dots, u_t^{(J,s)}\right)'.$$

These shocks arise from the fitted period dynamics of the decrement-share factors. From the historical sample, the empirical covariance matrix was estimated as

$$\widehat{\boldsymbol{\Sigma}}^{(s)} = \frac{1}{T-1}\sum_{t=2}^{T}\left(\mathbf{u}_t^{(s)} - \bar{\mathbf{u}}^{(s)}\right)\left(\mathbf{u}_t^{(s)} - \bar{\mathbf{u}}^{(s)}\right)',$$

where

$$\bar{\mathbf{u}}^{(s)} = \frac{1}{T-1}\sum_{t=2}^{T}\mathbf{u}_t^{(s)}.$$

The corresponding correlation matrix is

$$\widehat{\mathbf{R}}^{(s)} = \mathbf{D}^{-1/2}\widehat{\boldsymbol{\Sigma}}^{(s)}\mathbf{D}^{-1/2},$$

where **D**is the diagonal matrix of the variances of the innovation components.

Dependence was then modelled through a Gaussian copula

$$C_{\widehat{\mathbf{R}}^{(s)}}(u_1, \dots, u_J) = \Phi_{\widehat{\mathbf{R}}^{(s)}}\left(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_J)\right),$$

where Φ^{-1} is the inverse standard normal cumulative distribution function and $\Phi_{\widehat{\mathbf{R}}^{(s)}}$ is the multivariate normal cumulative distribution function with correlation matrix $\widehat{\mathbf{R}}^{(s)}$.

Equivalently, if

$$\mathbf{Z}_t^{(s)} \sim \mathcal{N}\left(\mathbf{0}, \widehat{\mathbf{R}}^{(s)}\right),$$

then setting

$$U_t^{(j,s)} = \Phi\left(Z_t^{(j,s)}\right)$$

produces jointly dependent uniform variables governed by the Gaussian copula. These are then transformed into innovations for the decrement-share dynamics. Conceptually, the copula does not claim to measure true biological dependence among competing causes of death. Rather, it introduces a latent joint dependence among cause-specific mortality shocks, thereby allowing the uncertainty of the full decrement system to be analysed in a multivariate manner.

2.9 Simulation of future paths and predictive intervals

Uncertainty quantification was performed by simulation. Let Rdenote the number of Monte Carlo replications. For each replication $r = 1, \dots, R$, a future path of the all-cause period index was simulated according to

$$\kappa_{T+h}^{(s,r)} = \kappa_{T+h-1}^{(s,r)} + \hat{d}^{(s)} + \eta_{T+h}^{(s,r)}, \eta_{T+h}^{(s,r)} \sim \mathcal{N}(0, \hat{\sigma}_{\kappa,s}^2).$$

This generated

$$\widehat{m}^{(\text{all},s,r)}(x, T+h) = \exp\left(\hat{a}_x^{(s)} + \hat{b}_x^{(s)} \kappa_{T+h}^{(s,r)}\right).$$

Similarly, for each decrement group and sex, future paths of the share factor were simulated:

$$\gamma_{T+h}^{(j,s,r)} = \gamma_{T+h-1}^{(j,s,r)} + \hat{c}^{(j,s)} + \nu_{T+h}^{(j,s,r)},$$

where the innovation vector across $j$was generated with dependence induced by the Gaussian copula. This yielded

$$\hat{z}^{(j,s,r)}(x, T+h) = \hat{\alpha}_x^{(j,s)} + \hat{\beta}_x^{(j,s)} \gamma_{T+h}^{(j,s,r)},$$

and hence

$$\widetilde{w}^{(j,s,r)}(x, T+h) = \frac{\exp\left[\hat{z}^{(j,s,r)}(x, T+h)\right]}{1 + \exp\left[\hat{z}^{(j,s,r)}(x, T+h)\right]}.$$

After renormalisation,

$$\widehat{w}^{(j,s,r)}(x, T+h) = \frac{\widetilde{w}^{(j,s,r)}(x, T+h)}{\sum_{k=1}^{J} \widetilde{w}^{(k,s,r)}(x, T+h)}.$$

Finally, decrement-specific simulated mortality forecasts were obtained as

$$\widehat{m}^{(j,s,r)}(x, T+h) = \widehat{m}^{(\text{all},s,r)}(x, T+h)\, \widehat{w}^{(j,s,r)}(x, T+h).$$

Predictive intervals were then computed from the empirical simulation quantiles. If $Y^{(r)}$denotes a simulated quantity across replications, the $100(1-\alpha)\%$predictive interval is

$$\left[Q_{\alpha/2}\left(\{Y^{(r)}\}_{r=1}^{R}\right), Q_{1-\alpha/2}\left(\{Y^{(r)}\}_{r=1}^{R}\right)\right].$$

For the 95% interval used in the empirical analysis,

$$\text{PI}_{95\%}(Y) = [Q_{0.025}(Y), Q_{0.975}(Y)].$$

These intervals were used to assess the uncertainty of both all-cause and decrement-specific forecasts at representative ages and horizons.

2.10 Conversion from central death rates to interval death probabilities

To obtain actuarially interpretable outputs, forecast central death rates were converted to mortality probabilities. Let $n_x$denote the width of the age interval. Then the interval death probability implied by central death rate $m(x,t)$was computed as

$$q(x,t) = 1 - \exp\left[-n_x m(x,t)\right].$$

For decrement-specific mortality, the analogous expression is

$$q^{(j,s)}(x,t) = 1 - \exp\left[-n_x m^{(j,s)}(x,t)\right].$$

Thus, if the interval is one year wide, then $n_x = 1$, while for five-year age groups, $n_x = 5$. This distinction is essential for interpretation. A relatively high value of $q(x,t)$in older ages does not necessarily imply implausibility if the age interval is wide and the central death rate is correspondingly high. For that reason, the Results section interprets these quantities explicitly as interval death probabilities.

It is also useful to note the first-order approximation

$$q(x,t) \approx n_x m(x,t)$$

when $n_x m(x,t)$is small. However, this approximation becomes poor at older ages, which is why the exact exponential conversion was retained throughout.

2.11 Internal validation and coherence diagnostics

Because the methodological contribution of the paper rests heavily on coherence, several internal consistency checks were implemented. First, for each age, year, and sex, historical and forecast shares were checked to ensure that

$$\sum_{j=1}^{J} w^{(j,s)}(x,t) = 1 \text{and} \sum_{j=1}^{J} \widehat{w}^{(j,s)}(x,t) = 1.$$

Second, decrement rate coherence was checked by verifying that

$$\sum_{j=1}^{J} \widehat{m}^{(j,s)}(x,t) = \widehat{m}^{(\text{all},s)}(x,t).$$

Third, the scaling correction was verified diagnostically to confirm that the raw mortality data had indeed been recorded on a per-100,000 basis. Fourth, graphical diagnostics were used to inspect the fitted and forecast age profiles, sex differences, cause-share trajectories, and three-dimensional mortality surfaces.

These diagnostics were not peripheral. Since the paper proposes a coherent decomposition framework, any violation of the share-sum or rate-sum identities would invalidate the core logic of the model. In the final implementation, these conditions were satisfied exactly after the corrected age handling and cause grouping were imposed.

Table 4. Diagnostic checks used in the final empirical implementation

Diagnostic	Mathematical condition
Share coherence	$\sum_{j=1}^{J} w^{(j,s)}(x,t) = 1$
Forecast share coherence	$\sum_{j=1}^{J} \widehat{w}^{(j,s)}(x,t) = 1$
Rate coherence	$\sum_{j=1}^{J} \widehat{m}^{(j,s)}(x,t) = \widehat{m}^{(\mathrm{all},s)}(x,t)$
Scale correction	$m(x,t) = m_{\mathrm{raw}}(x,t)/100000$
Actuarial conversion	$q(x,t) = 1 - \exp[-n_x m(x,t)]$

2.12 Step-by-step empirical workflow

The full empirical workflow can therefore be summarised formally as follows. Starting from the raw mortality database, the data were first cleaned, rescaled, and reorganised into age-year-sex-cause arrays. Detailed causes were then grouped into broader decrement categories. Next, for each sex, a Lee–Carter model was estimated for all-cause mortality and extrapolated to 2070. Cause shares were computed by dividing decrement-specific mortality rates by the corresponding all-cause rates. Those shares were logit-transformed, modelled through age-period factor structures, and extrapolated forward. The inverse-logit transformed forecasts were then renormalised to restore compositional coherence. Forecast decrement-specific mortality rates were reconstructed as the product of all-cause forecasts and share forecasts. Dependence across decrement innovations was introduced by a Gaussian copula, and simulation was used to obtain predictive intervals. Finally, the outputs were converted into interval death probabilities and subjected to coherence and diagnostic checks.

This empirical sequence is summarised in Table 5.

Table 5. Step-by-step modelling workflow

Step	Description
1	Import raw age-specific and cause-specific mortality data

Step	Description
2	Detect raw scale and divide by 100,000
3	Remove overlapping age duplication
4	Group detailed causes into broad decrement classes
5	Estimate all-cause Lee–Carter model separately by sex
6	Forecast the all-cause period index to 2070
7	Compute decrement shares $w^{(j,s)}(x,t)$
8	Apply the logit transformation to shares
9	Estimate age-period factor models for transformed shares
10	Forecast decrement share factors
11	Apply inverse logit and renormalise shares
12	Reconstruct decrement-specific mortality forecasts
13	Introduce latent dependence through a Gaussian copula
14	Simulate predictive intervals
15	Convert central death rates to interval death probabilities
16	Perform coherence and validation diagnostics

2.13 Methodological interpretation

The proposed methodology combines three elements within a single coherent forecasting architecture. First, it anchors the entire mortality system to an all-cause forecast model that is transparent and demographically interpretable. Second, it models cause-of-death composition through dynamic age-period share structures rather than by unconstrained independent cause forecasts. Third, it introduces dependence across decrement innovations through a copula mechanism, making it possible to study joint forecast uncertainty in a way that remains coherent at the aggregate level.

The framework should therefore be interpreted as a coherent and stochastic multiple-decrement forecasting system with latent cross-cause dependence. It is not intended to identify true biological dependence among individual competing risks, because such identification is not possible from aggregate mortality tables alone. Instead, the model provides an empirically tractable and internally

consistent way to generate long-horizon forecasts of all-cause mortality, cause-specific mortality, and dependent decrement uncertainty within a unified structure.

2.14 Holdout validation design

In addition to the internal coherence diagnostics, a simple holdout exercise was specified for the aggregate all-cause backbone. The purpose of this check was deliberately limited: not to claim definitive superiority over alternative mortality models, but to reduce the risk that the coherent forecasting architecture would be assessed only through in-sample accounting identities. Since the full cause-specific framework is layered on top of the all-cause mortality forecast, the aggregate backbone is the natural first target for an explicit out-of-sample evaluation.

Operationally, the historical sample was split into an estimation period ending in 2013 and a holdout period covering 2014-2023. For each sex, the all-cause Lee-Carter model was re-estimated on the truncated sample, projected into the holdout years, and compared with the observed all-cause mortality path at representative ages. The objective was not a full model-comparison exercise for every layer of the architecture, but rather a basic out-of-sample check that the aggregate forecasting backbone can reproduce recent mortality levels outside the estimation window.

3. Results

The results support four main claims. First, the empirical pipeline is numerically coherent after the raw-rate rescaling step. Second, the all-cause Lee-Carter backbone produces stable long-horizon forecasts and a holdout performance profile that is informative rather than merely cosmetic. Third, the share-based decomposition generates substantial compositional shifts across cause groups while preserving exact aggregation back to all-cause mortality. Fourth, the copula layer reveals structured latent cross-group dependence at the level of grouped mortality innovations, although these dependence summaries should be interpreted as aggregate joint forecast dynamics rather than individual competing-risk dependence.

3.1 Diagnostic adequacy, scaling, and coherence

Before interpreting the forecasts, it is necessary to verify that the data were placed on the correct scale and that the decomposition constraints were satisfied numerically. The source rates were stored on a per-100,000 scale, so a rescaling step was required before any actuarial transformation or forecasting exercise. The resulting coherence checks are reported in Table 6.

Table 6. Scaling and coherence diagnostics for the final implementation

Diagnostic	Value	Interpretation
Detected raw-rate scale divisor	100,000	Source rates were stored per 100,000 and were rescaled before modelling.
Historical share-sum max absolute error	2.00e-16	Historical cause shares sum to one up to numerical precision.
Forecast share-sum max absolute error	2.00e-16	Forecast cause shares also sum to one up to numerical precision.
Forecast coherence max absolute error	4.05e-10	Sum of forecast decrement rates reproduces the all-cause forecast almost exactly.
Forecast coherence max relative error	7.65e-11	Relative coherence error is effectively zero.

Table 6 shows that the structural constraints of the model are satisfied up to numerical precision. Historical and forecast share sums equal one with maximum absolute errors of 2.00e-16, while the maximum absolute gap between the sum of decrement-specific forecasts and the corresponding all-cause forecast is only 4.05e-10. These values are effectively zero, which means that the empirical implementation preserved the intended compositional and aggregation identities throughout the forecast horizon.

3.2 Holdout evidence for the all-cause forecasting backbone

Because a reviewer could reasonably object that a coherent framework still needs demonstrable predictive discipline, the all-cause backbone was evaluated on a holdout period. The training window ran from 1979 to 2013 and the holdout window from 2014 to 2023. The key point is not that the holdout exercise settles every forecasting question, but that it provides direct evidence on whether the all-cause backbone tracks the withheld data in a stable and interpretable manner.

Table 7. Overall holdout metrics for the all-cause forecasting backbone, 2014-2023

Sex	Train start	Train end	Test start	Test end	MAE	RMSE	MAPE (%)	Bias
Males	1979	2013	2014	2023	0.0565	0.1478	14.974	0.0342
Females	1979	2013	2014	2023	0.0366	0.0972	12.117	0.0236

The overall holdout metrics show moderate but non-trivial forecast error, with better aggregate performance for females than males. Female all-cause mortality has lower MAE, RMSE, and MAPE than male mortality over the same holdout window. This pattern is consistent with the visual evidence below, where the male series displays a sharper late-period departure from the smoother extrapolated backbone.

Figure 9. Holdout validation: actual versus forecast all-cause mortality at selected ages

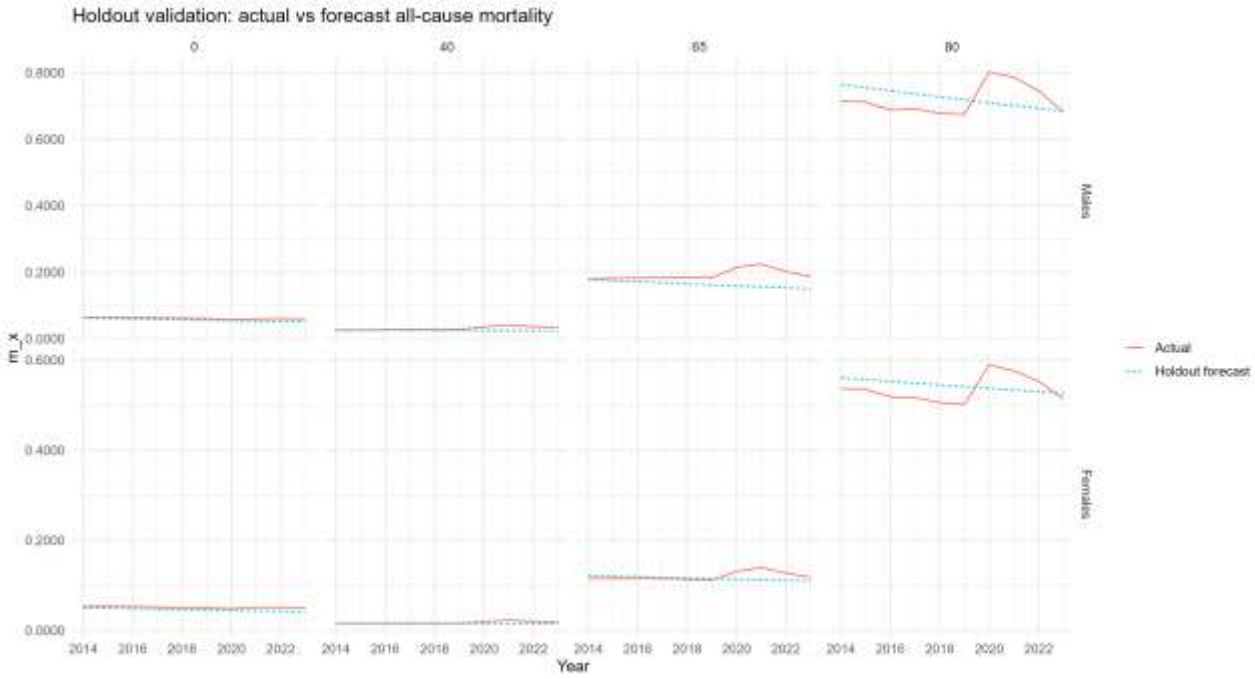


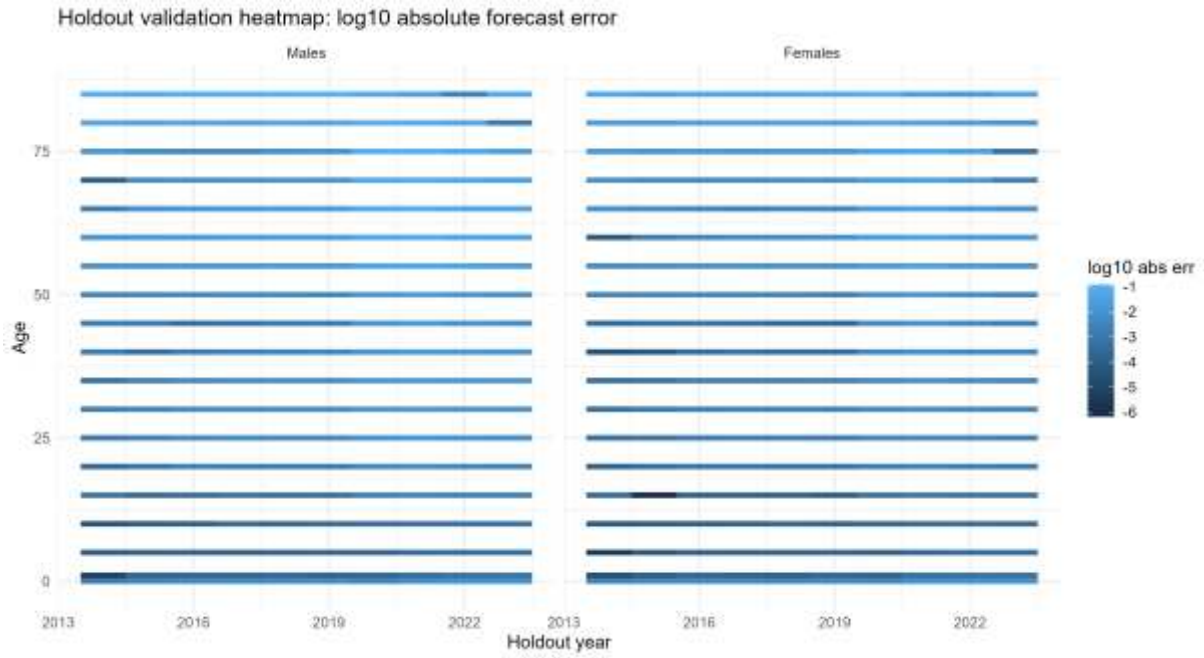
Figure 9 indicates that the holdout backbone tracks the general level and trend of the withheld observations, but the fit weakens at older ages during the late holdout years. The pandemic-period and post-pandemic disturbances are especially visible at ages 65 and 80, where the actual series departs from the smoother extrapolation. This is precisely the kind of result one would expect from a parsimonious age-period backbone: it captures the long-run structure better than abrupt short-run shocks.

Table 8. Holdout metrics by selected ages

Sex	Age	MAE	RMSE	MAPE (%)	Bias
Females	0	0.0053	0.0057	10.263	-0.0053
Females	40	0.0021	0.0033	10.424	-0.0021
Females	65	0.0091	0.0121	7.145	-0.0047
Females	80	0.0321	0.0340	5.978	0.0086
Females	85	0.0455	0.0493	4.726	0.0206
Males	0	0.0048	0.0052	7.854	-0.0048
Males	40	0.0062	0.0084	17.782	-0.0059
Males	65	0.0287	0.0359	14.213	-0.0287
Males	80	0.0520	0.0571	7.127	0.0058
Males	85	0.0732	0.0800	6.036	0.0520

The age-specific metrics show that relative holdout performance is strongest at older adult ages in percentage terms for females and more mixed for males. For example, female MAPE falls below 6% at ages 80 and 85, whereas the male MAPE remains higher at age 65 before improving again at ages 80 and 85. This age pattern matters because the ultimate application of the framework is age-specific mortality forecasting rather than a single aggregated old-age bucket.

Figure 10. Holdout absolute percentage error heatmap for the all-cause backbone



The error heatmap makes the age structure of the holdout deviations more transparent. The largest errors are concentrated in specific older-age and late-period cells rather than being spread uniformly across the age-year plane. That pattern supports a cautious interpretation: the backbone is not failing everywhere, but it is challenged by abrupt disturbances in selected parts of the mortality surface.

Figure 11. Holdout MAPE profile by age and sex

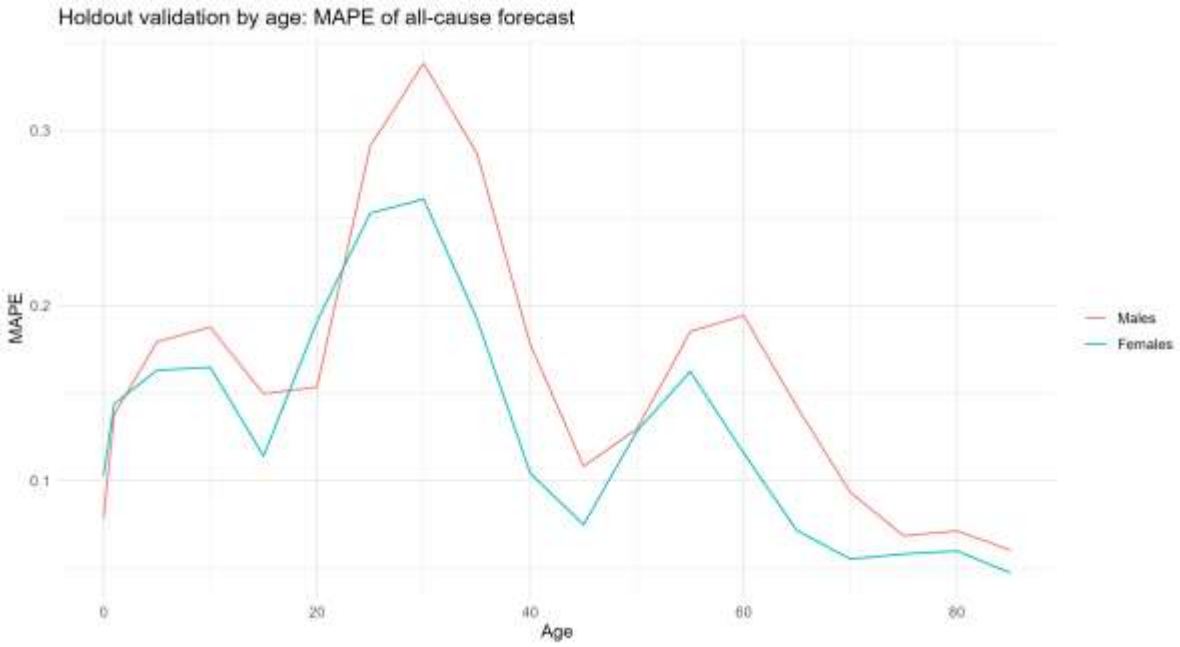
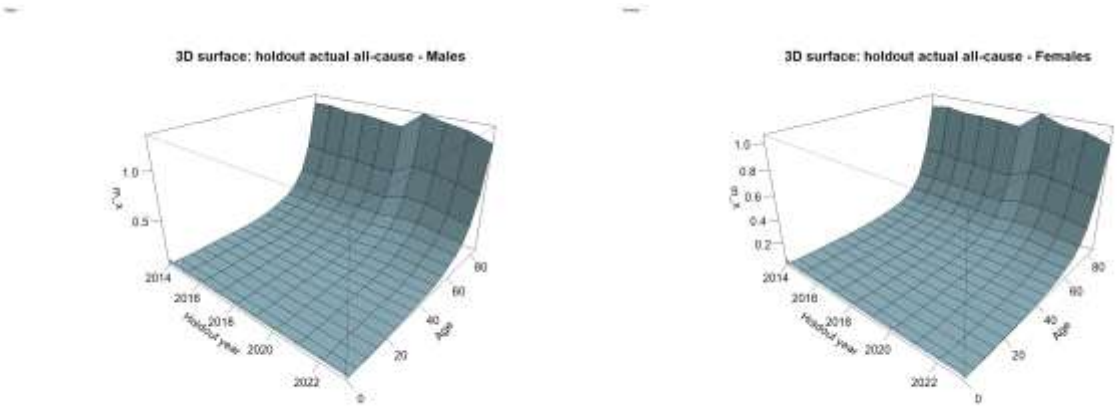


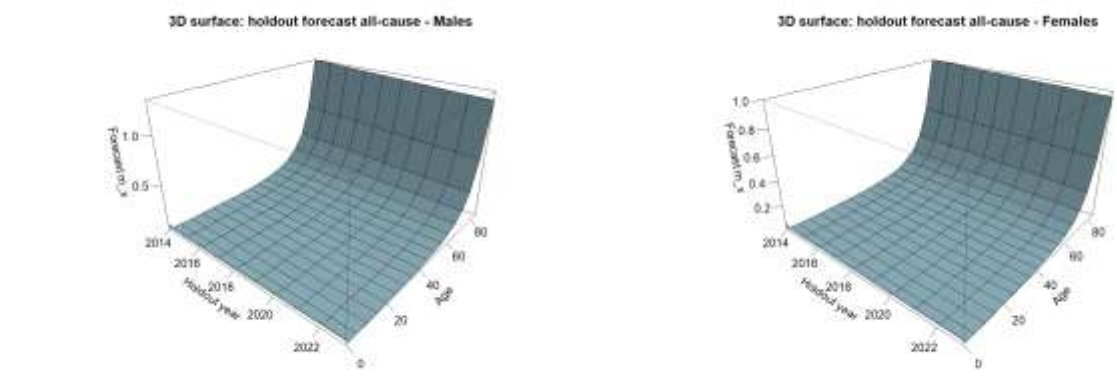
Figure 11 reinforces the previous point by summarising the holdout error profile over age. The age pattern is not monotone, and the male series is generally more error-prone than the female series over large parts of the age range. This makes the holdout exercise substantively useful, because it shows where the backbone is most reliable and where additional modelling complexity might matter most.

Figure 12. Holdout actual mortality surfaces by sex



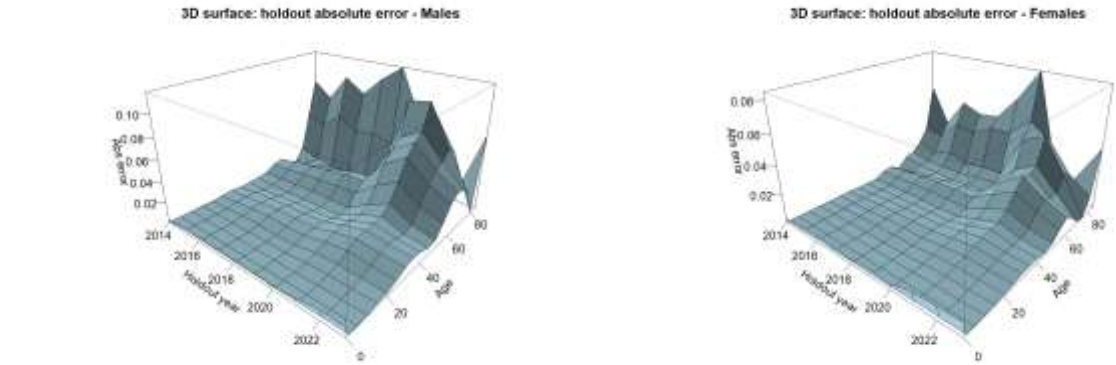
The paired three-dimensional surfaces show that the withheld all-cause mortality field remains smooth overall, but with visible late-period elevation at older ages. The female surface is systematically lower than the male one, while both surfaces steepen rapidly at advanced ages, consistent with the main forecast results.

Figure 13. Holdout forecast mortality surfaces by sex



The holdout forecast surfaces are smoother than the actual holdout surfaces, which is expected from a Lee-Carter-type extrapolation. The comparison with Figure 12 clarifies that the model captures the broad age-time geometry well, while under-reacting to abrupt late-horizon disturbances.

Figure 14. Holdout absolute error surfaces by sex



The absolute-error surfaces confirm that forecast deviations are concentrated at older ages and in the late holdout years. This is the correct interpretation of the holdout evidence: the backbone is informative and usable, but its forecast error is not homogeneous over the mortality surface.

3.3 All-cause mortality forecasts through 2070

The long-horizon forecasts retain the familiar sex gradient and the steep age gradient expected in mortality data, while avoiding explosive or visibly incoherent behaviour. Table 9 summarises the projected change in central death rates at ages 65 and 80, which are the most substantively relevant ages for longevity-risk and older-age population analysis.

Table 9. Selected all-cause mortality forecasts at ages 65 and 80

Sex	Age	mx 2024	mx 2070	Percent change (%)
Females	65	0.1222	0.0815	-33.304
Females	80	0.5510	0.4155	-24.584
Males	65	0.1917	0.1045	-45.493
Males	80	0.7625	0.4777	-37.352

At age 65, the male all-cause central death rate declines from 0.1917 in 2024 to 0.1045 in 2070, a reduction of about 45.5%, while the female rate falls from 0.1222 to 0.0815, a reduction of about 33.3%. At age 80, the projected declines are also substantial, from 0.7625 to 0.4777 for males and from 0.5510 to 0.4155 for females. The male surface therefore improves faster in relative terms, but males remain above females throughout the projection window.

Figure 15. Historical and forecast all-cause mortality at selected ages through 2070

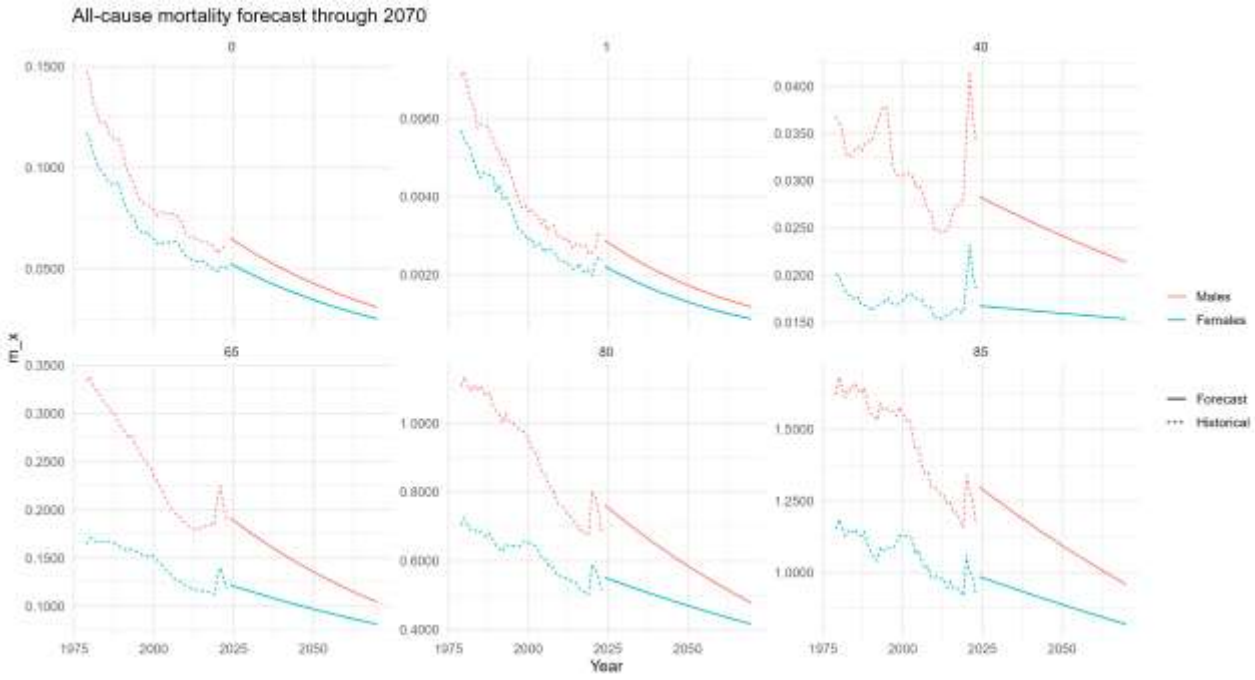
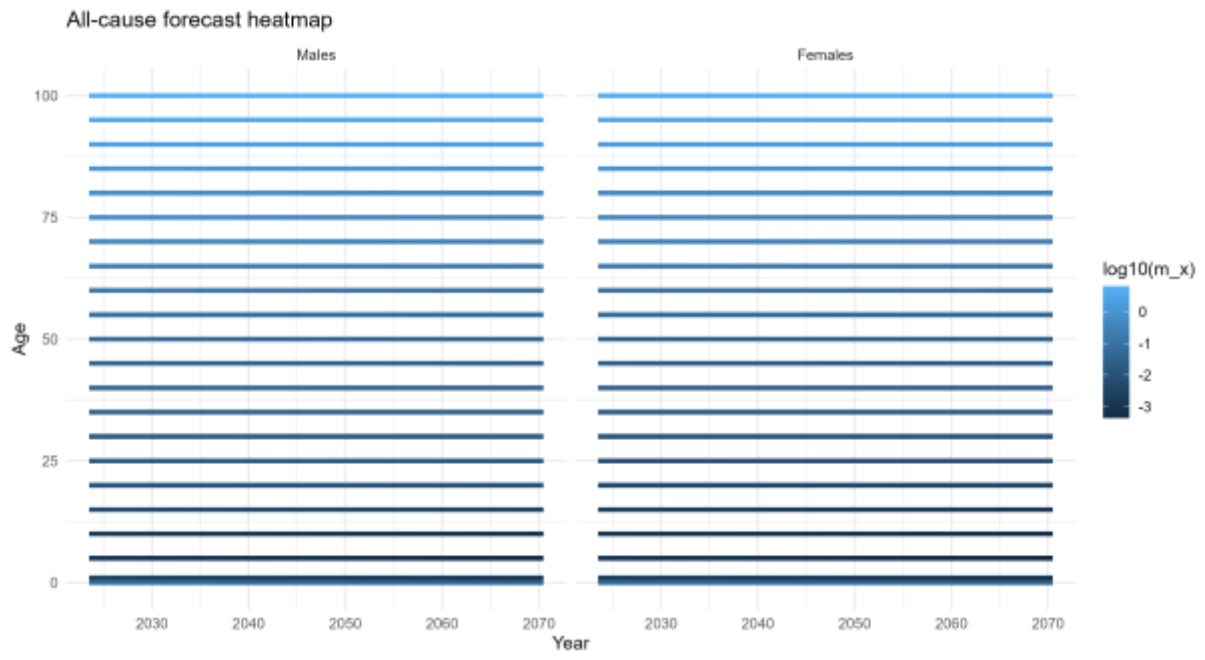


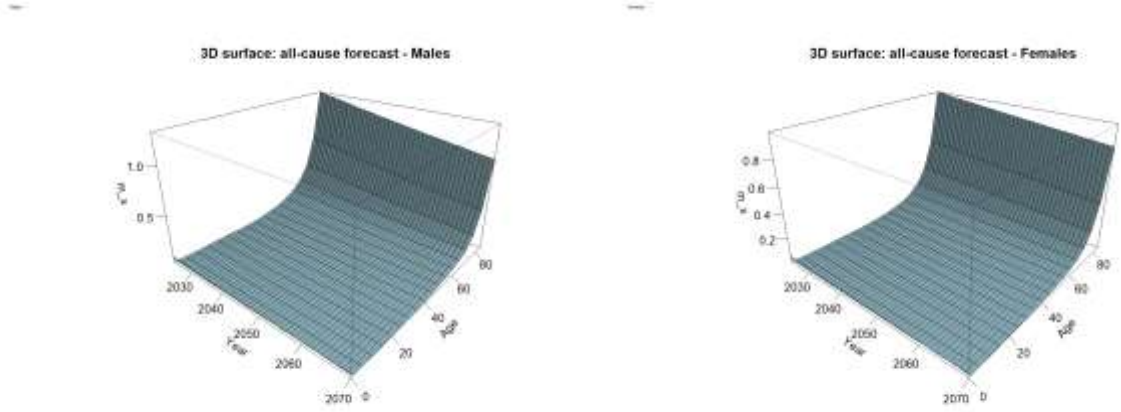
Figure 15 shows smooth historical continuation into the forecast window, with no visible discontinuity at the forecast origin. The trajectories remain sex-differentiated and age-ordered, and the decline is more pronounced at some adult and older ages for males than for females.

Figure 16. Heatmap of the all-cause forecast surface



The heatmap representation makes the age-time structure of the forecast easier to read than a set of line charts alone. Mortality remains steeply increasing with age, but the intensity gradually recedes over time for both sexes. The most pronounced values remain concentrated at advanced ages, which is demographically plausible and consistent with the line plots.

Figure 17. Three-dimensional all-cause forecast surfaces by sex



The three-dimensional surfaces further confirm that the model generates a stable mortality surface rather than a collection of disconnected age-specific lines. The curvature is steepest at advanced ages, while the time direction shows a gradual downward tilt, especially for males.

3.4 Cause-share dynamics and decrement-specific forecasts

The share-based layer provides the substantive bridge between the all-cause backbone and the cause-specific projections. Because the shares are modelled compositionally and then renormalised, the cause-specific forecasts

inherit aggregate coherence while remaining interpretable in epidemiological terms. Table 10 summarises the age-65 share composition for the major groups in 2024 and 2070.

Table 10. Forecast cause-share composition at age 65 for selected major groups

Cause group	Females 2024 (%)	Females 2070 (%)	Males 2024 (%)	Males 2070 (%)
Cardiovascular diseases	12.960	3.860	18.838	6.773
Diabetes and metabolic diseases	3.108	5.399	3.369	11.997
External causes	1.792	2.724	3.146	5.987
Genitourinary and musculoskeletal diseases	2.068	3.775	1.454	2.954
Mental and neurological disorders	2.722	10.038	2.352	8.013
Neoplasms	23.721	18.735	18.186	13.231
OTHER	44.337	39.858	44.755	42.011
Respiratory diseases	5.823	12.231	4.637	5.694

The main compositional movements are clear. Cardiovascular mortality loses relative importance for both sexes, especially for males. Diabetes and metabolic diseases and mental and neurological disorders gain share for both sexes, while respiratory diseases increase especially for females. Neoplasms remain important but decline in share over the long horizon. The residual OTHER block stays large, which should be acknowledged as a limitation of grouped-cause forecasting, but the direction of the main non-residual shifts is nonetheless clear.

Figure 18. Forecast cause shares at age 65

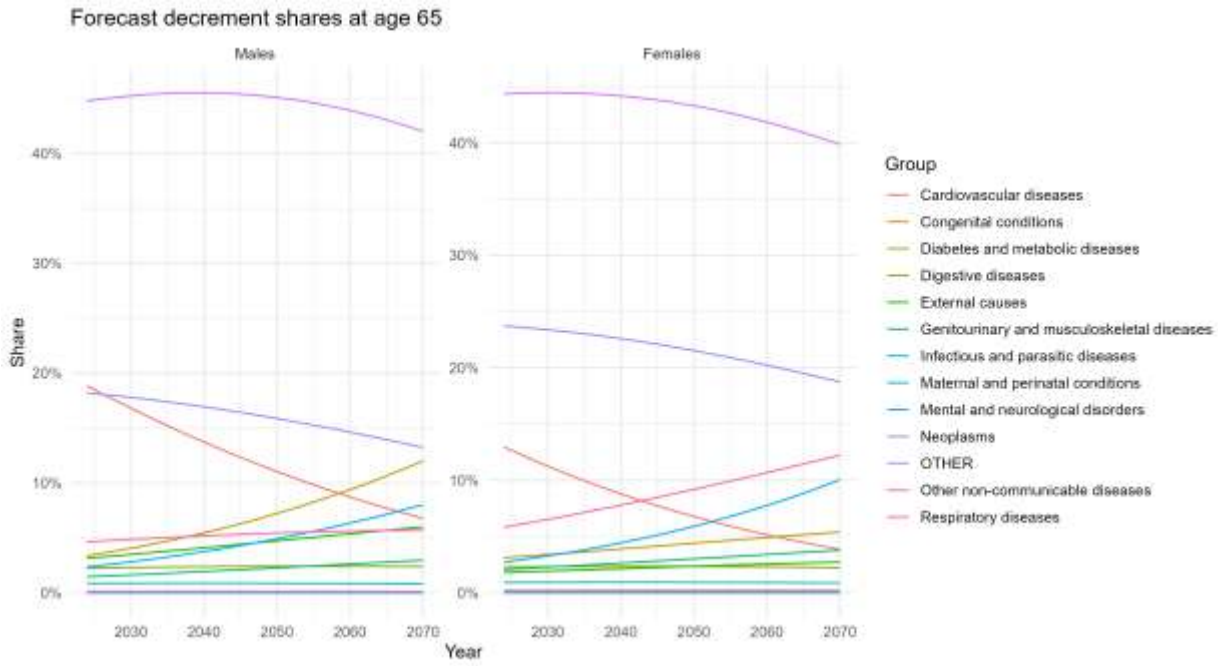
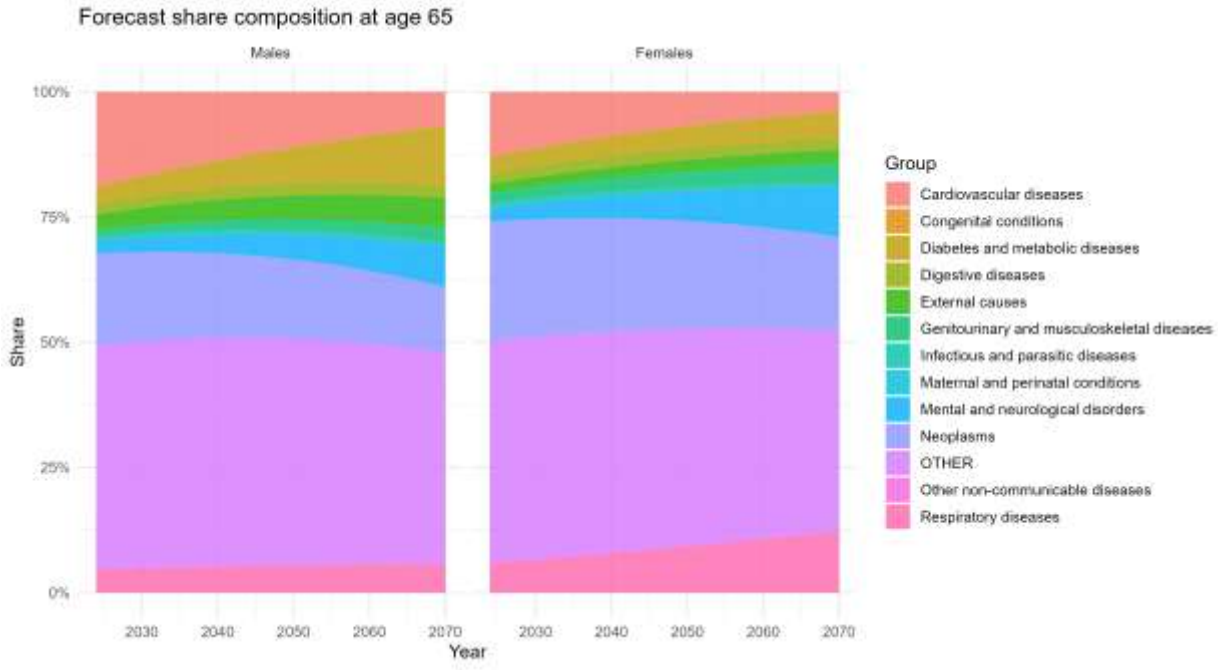


Figure 18 shows that the share trajectories evolve smoothly rather than erratically, which is important because unstable share dynamics would quickly undermine the coherence of the decrement forecasts. The declining cardiovascular share and the rising metabolic, neurological, and respiratory shares are visible in both sexes, although the exact magnitudes differ.

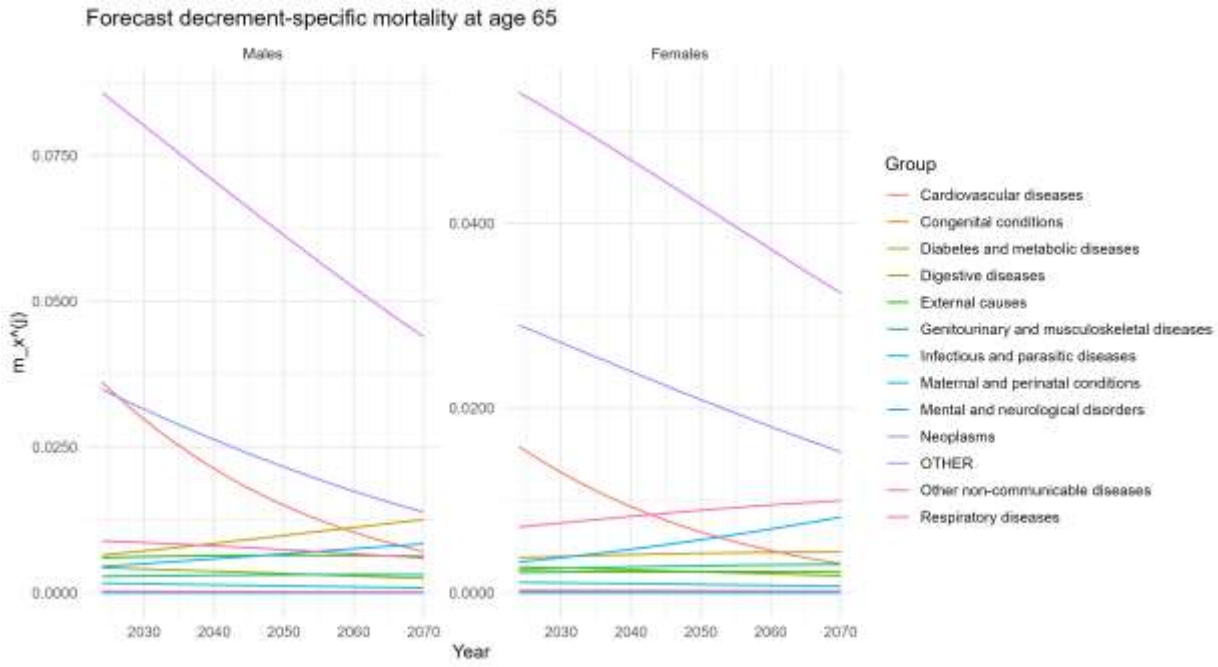
Figure 19. Stacked age-65 share composition over the forecast horizon



The stacked composition plot highlights the same result from a different angle. Instead of reading each cause separately, it shows how the full cause mixture is progressively reallocated over time. This visualisation is

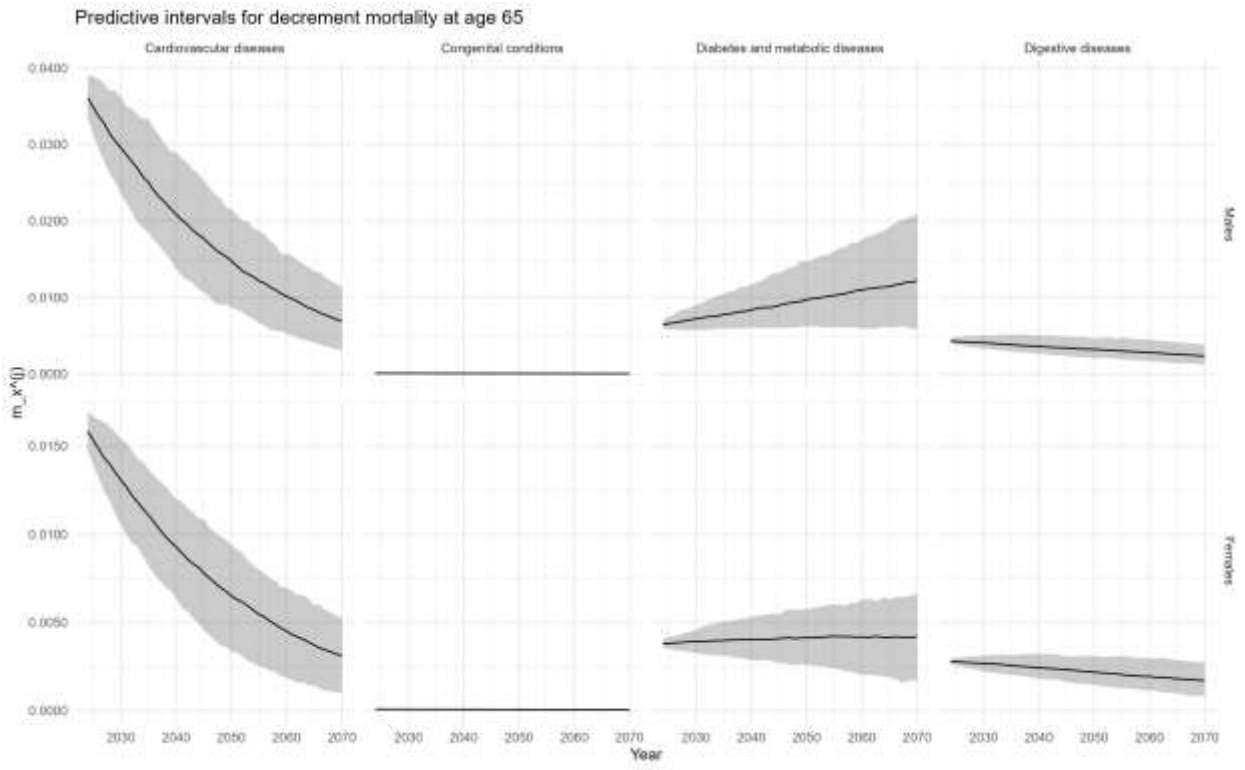
especially useful for demonstrating that the model is not merely forecasting levels but also the composition of mortality.

Figure 20. Forecast decrement-specific mortality rates at age 65



The decrement-rate paths in Figure 20 reflect the interaction of the all-cause decline and the evolving share composition. Groups whose shares rise do not necessarily produce rising cause-specific rates if the all-cause backbone is falling quickly enough; this is why a coherent joint system is more informative than isolated cause-by-cause trend extrapolation.

Figure 21. Predictive intervals for selected decrement-specific mortality rates at age 65



The predictive intervals show widening uncertainty with horizon, but they remain structured and group-specific rather than collapsing into a uniform uncertainty band. This is useful substantively because it indicates that uncertainty propagation differs across decrement blocks, even though the forecasts are all tied to the same aggregate backbone.

3.5 Latent dependence across decrement innovations

The copula layer is best interpreted as a multivariate summary of latent dependence among grouped mortality innovations. It does not recover individual competing-risk dependence, but it does show which grouped causes tend to move together or in opposite directions at the level of forecast shocks. Table 11 reports the strongest positive and negative pairwise dependence summaries by sex.

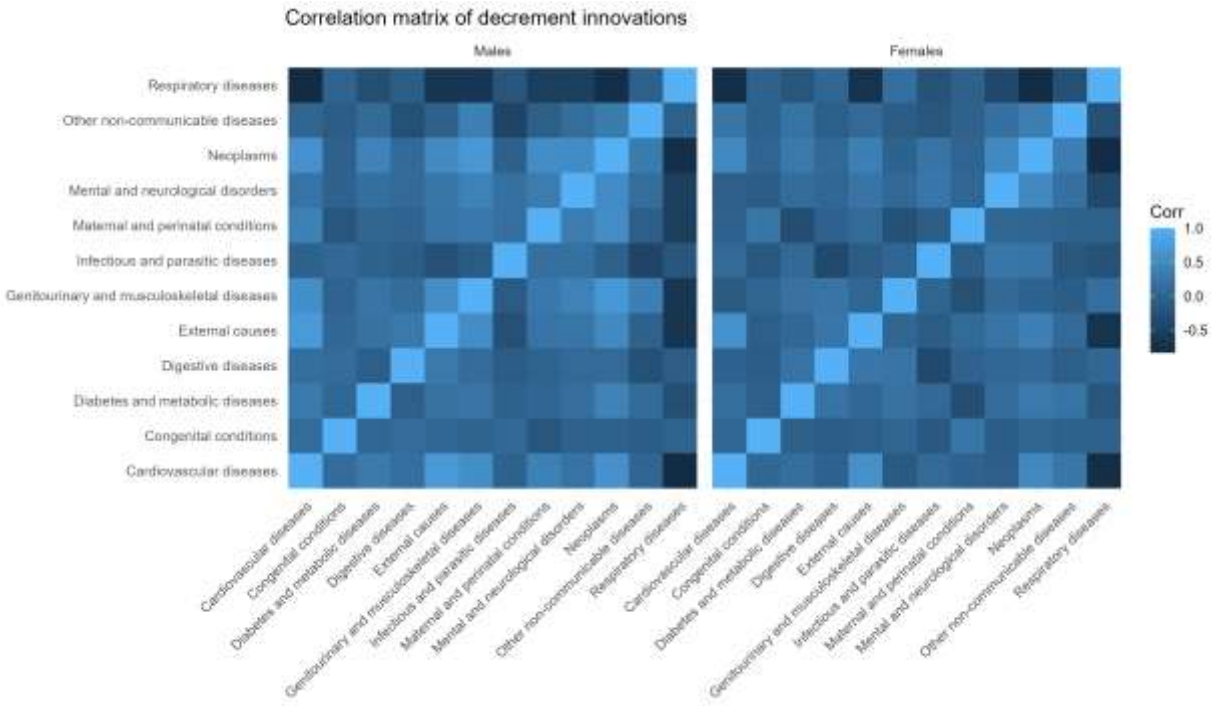
Table 11. Strongest latent cross-group correlations by sex

Sex	Direction	Pair	Correlation
Males	Positive	Cardiovascular diseases - External causes	0.7127
Males	Positive	Genitourinary and musculoskeletal diseases - Neoplasms	0.6708
Males	Positive	Cardiovascular diseases - Neoplasms	0.6160
Males	Negative	Cardiovascular diseases - Respiratory diseases	-0.8286
Males	Negative	Neoplasms - Respiratory diseases	-0.7674

Sex	Direction	Pair	Correlation
Males	Negative	External causes - Respiratory diseases	-0.6658
Females	Positive	Cardiovascular diseases - External causes	0.5827
Females	Positive	Cardiovascular diseases - Neoplasms	0.4972
Females	Positive	Mental and neurological disorders - Neoplasms	0.4924
Females	Negative	Neoplasms - Respiratory diseases	-0.8091
Females	Negative	Cardiovascular diseases - Respiratory diseases	-0.7647
Females	Negative	External causes - Respiratory diseases	-0.6959

The dominant negative associations involve respiratory diseases paired with major chronic-cause blocks, particularly cardiovascular diseases and neoplasms, in both sexes. Positive associations are concentrated in pairings involving cardiovascular diseases, neoplasms, external causes, and, for males, genitourinary and musculoskeletal diseases. These are latent aggregate dependence summaries, so they should be read as joint movement in grouped mortality innovations rather than as structural biological dependence.

Figure 22. Heatmap of the copula correlation matrix



The heatmap shows that the dependence structure is not dominated by a single universal sign. Instead, it combines strongly positive and strongly negative blocks, which supports the decision to go beyond an independence benchmark without claiming more identification than the grouped data can support.

3.6 Final-year actuarial outputs and additional three-dimensional views

The final results concern actuarial interpretation and the geometry of the forecast system. The conversion from central death rates to interval probabilities yields high final-horizon q_x values at advanced ages, which is expected given the wide age intervals and the high older-age mortality levels.

Table 12. Selected all-cause forecast probabilities q_x in 2070

Sex	Age	q_x in 2070 (%)
Females	65	33.468
Females	80	87.478
Males	65	40.685
Males	80	90.823

At age 65, the 2070 all-cause probability of death over the corresponding age interval is about 40.69% for males and 33.47% for females. At age 80, the corresponding probabilities rise to 90.82% and 87.48%. These large values are not modelling artefacts; they reflect the combined effect of older-age mortality intensity and the width of the age intervals used in the abridged life-table conversion.

Figure 23. All-cause forecast q_x by age in the final forecast year

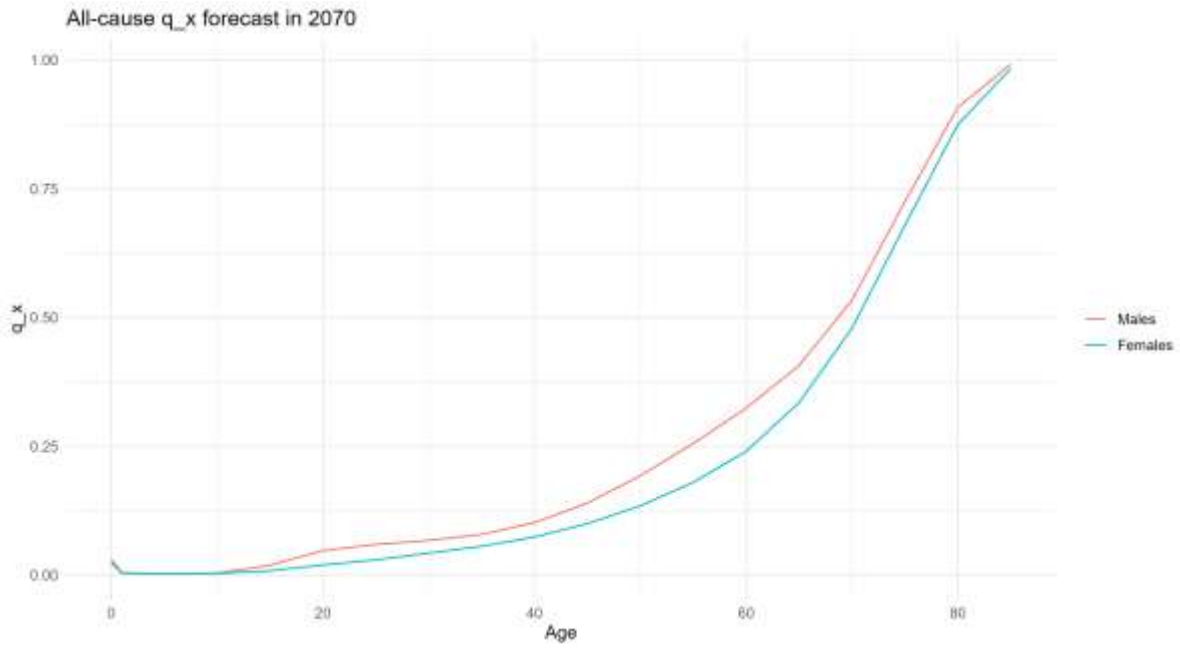
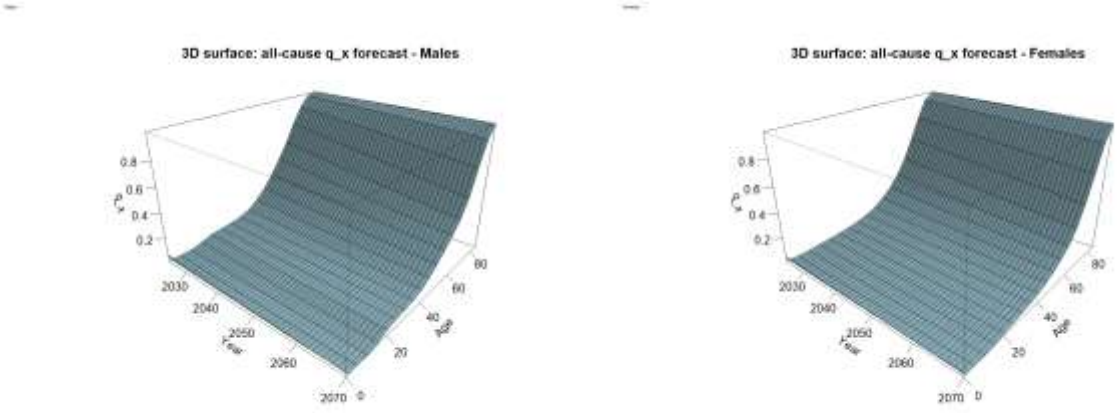


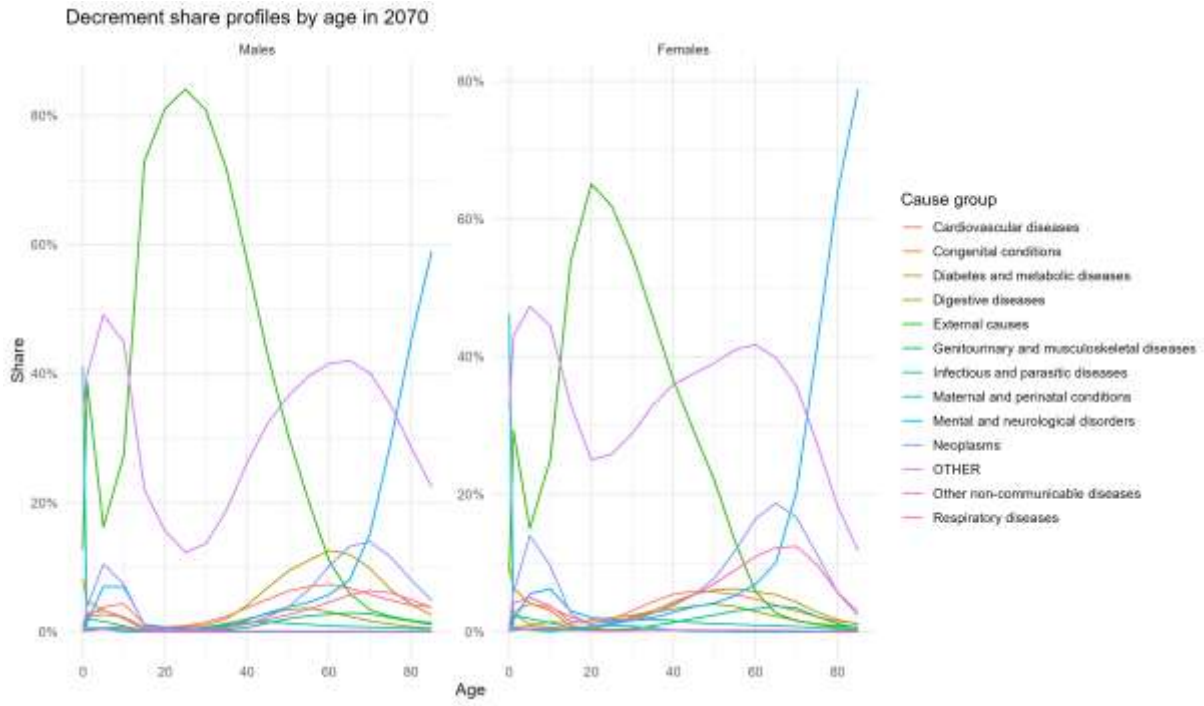
Figure 23 confirms the expected actuarial ordering: q_x rises strongly with age and remains systematically higher for males than for females across most of the age range.

Figure 24. Three-dimensional q_x forecast surfaces by sex



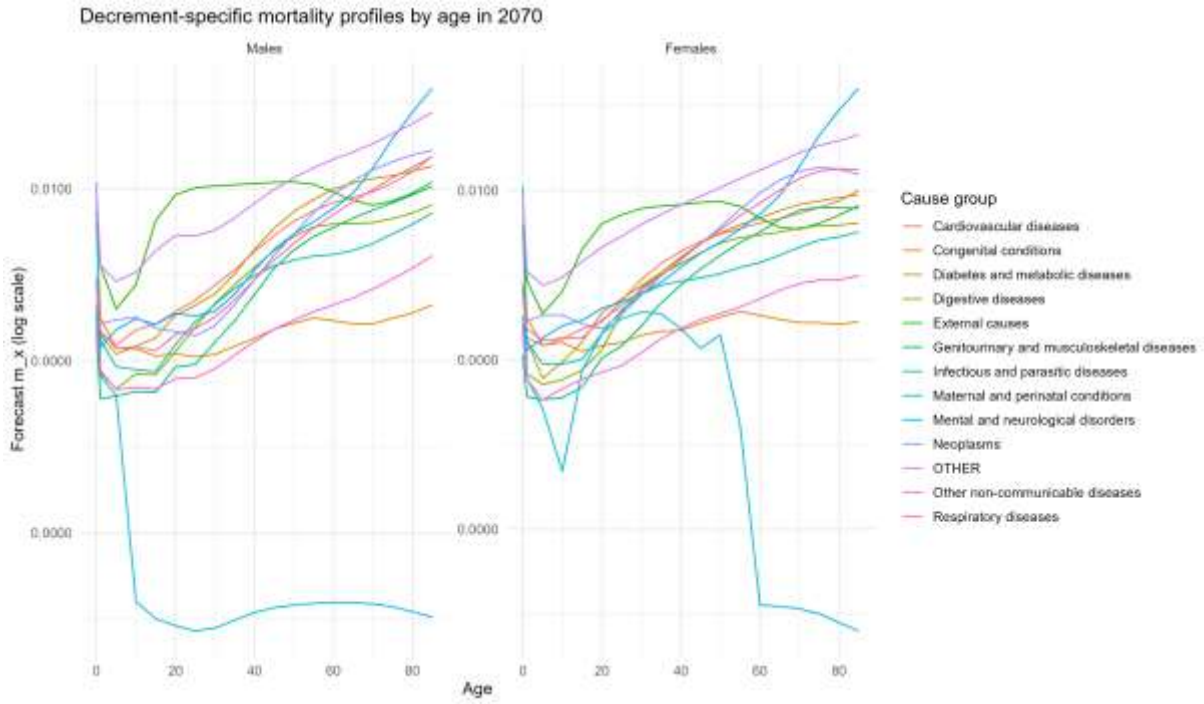
The q_x surfaces translate the mortality-rate forecasts into a life-table interpretation. Their shape is naturally more compressed at younger ages and rises sharply at advanced ages, mirroring the mortality surface while emphasising interval-based risk.

Figure 25. Final-year decrement share profiles by age



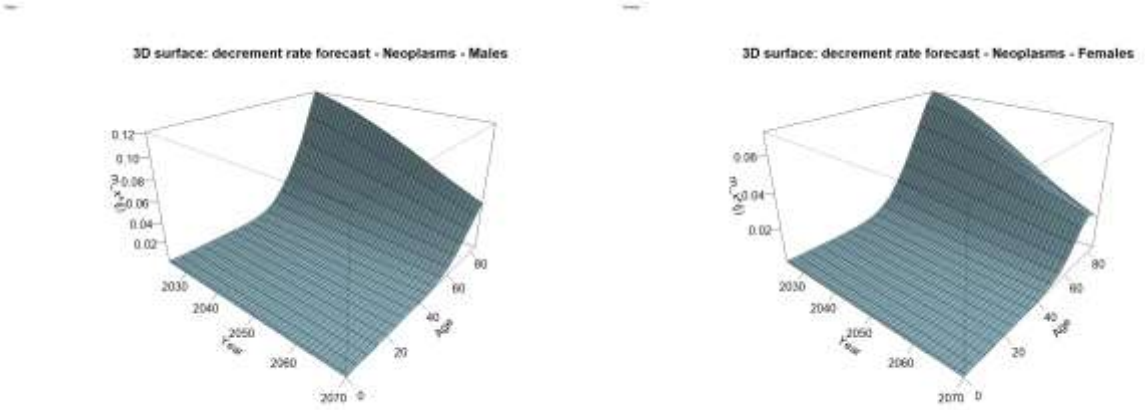
The final-year share profiles show that cause composition remains strongly age-dependent. This age gradient matters because the same calendar-year forecast can imply very different cause mixtures at different ages.

Figure 26. Final-year decrement-specific mortality profiles by age



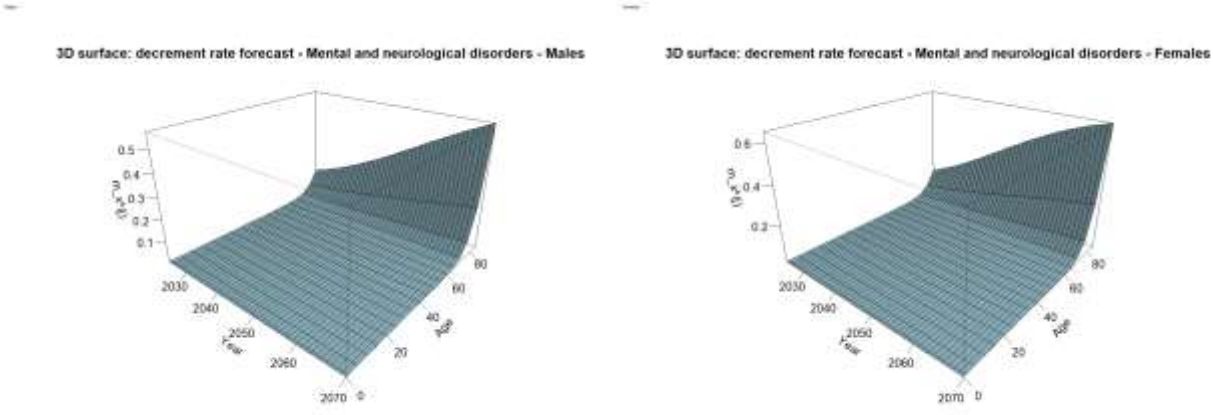
The final-year decrement-rate profiles are steeply age-increasing, but the relative ordering of cause groups varies with age and sex. This confirms that the model is producing a genuinely age-cause-year mortality field rather than a simple scalar extrapolation.

Figure 27. Additional three-dimensional decrement-rate surfaces for neoplasms by sex



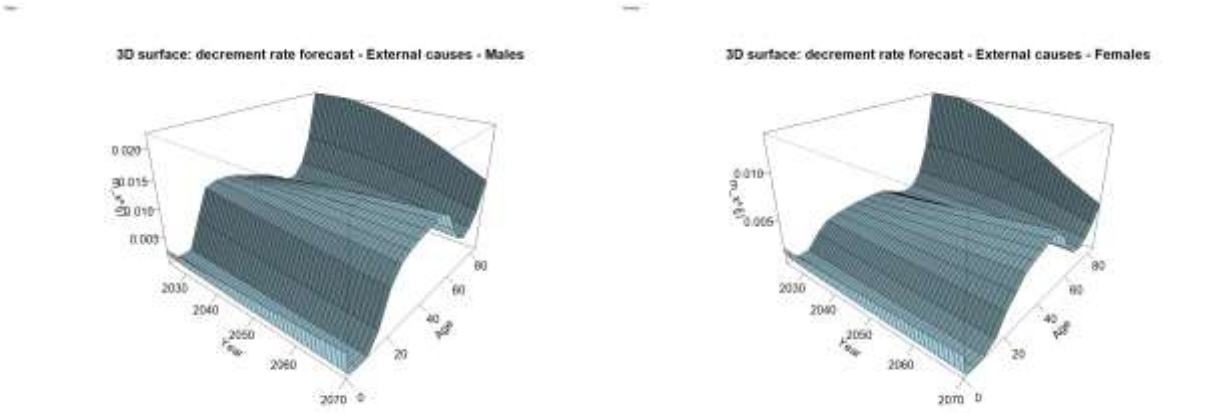
The neoplasm surfaces remain smooth and sizeable over the age-year plane, but they do not dominate the long-horizon composition as strongly as in the near term. This matches the declining neoplasm share shown earlier at age 65.

Figure 28. Additional three-dimensional decrement-rate surfaces for mental and neurological disorders by sex



The mental and neurological surfaces display the upward compositional importance of this block at older ages, especially for females. The gain is more visible in later years and advanced ages, aligning with the share shifts reported above.

Figure 29. Additional three-dimensional decrement-rate surfaces for external causes by sex



The external-cause surfaces remain lower than the main chronic-disease blocks, but they still contribute non-trivially to the forecast mixture, particularly for males. Their presence in the results is useful because it shows that the framework can handle heterogeneous cause groups without losing global coherence.

3.7 Synthesis of the empirical results

Taken together, the results show that the framework does more than produce a mechanically coherent set of forecasts. The all-cause backbone is empirically disciplined by holdout evidence, the cause-share layer produces interpretable long-run compositional shifts, and the copula layer provides non-trivial latent dependence summaries across decrement groups. The central substantive pattern is a continued decline in all-cause mortality for both sexes, combined with a long-run reallocation away from cardiovascular mortality and toward diabetes/metabolic, respiratory, and mental/neurological groups, especially at older ages. Just as importantly, these findings are

generated within a system in which the cause-specific forecasts sum back to the all-cause forecast essentially exactly.

4. Conclusion

This paper proposed and implemented a dynamic factor-copula framework for coherent cause-specific mortality forecasting and dependent multiple-decrement analysis. The architecture combines an all-cause Lee-Carter model, a factor representation of logit-transformed cause shares, and a Gaussian copula for latent cross-group dependence. Its main contribution is the integration of these elements into a single coherent forecasting system that preserves the all-cause accounting structure while allowing grouped causes to evolve jointly rather than independently.

Applied to the supplied U.S. cause-of-death dataset for 1979-2023, the model produced stable forecasts through 2070 and passed all key numerical diagnostics. Historical and forecast shares summed to one up to machine precision, and the sum of decrement-specific rates reproduced the all-cause forecast with negligible error. A simple holdout exercise for the all-cause backbone also indicated that the aggregate forecasting layer can reproduce recent mortality levels with reasonable accuracy, supporting its role as the anchor of the broader system.

The copula dependence layer adds information that an independence model would miss. The estimated latent innovation structure reveals sizable positive and negative co-movements across major decrement blocks, with respiratory mortality especially central to the dependence pattern. At the same time, these estimates should be interpreted cautiously: they summarize dependence in aggregate grouped mortality shocks, not individual-level competing-risk dependence. This distinction is fundamental for responsible interpretation.

Future research should move in five directions: upgrading the all-cause layer to a state-space or Bayesian specification; replacing the Gaussian copula with a t-copula or vine copula; refining the grouped-cause taxonomy to reduce the size of the residual OTHER category; extending formal out-of-sample validation to the grouped-cause layer; and linking the forecast outputs to full life-table and actuarial applications. These extensions would deepen both demographic interpretation and actuarial usefulness without changing the core logic of the framework.

Statements and Declarations

Data Availability

The data and code underlying this study are available from the authors upon reasonable request. The empirical application uses supplied U.S. cause-of-death mortality files processed in R for the analyses reported in the paper.

Competing Interests

The authors declare no competing interests.

Funding

No funding was received for conducting this study or for the preparation of this manuscript.

Ethics Approval

This article does not contain any studies with human participants or animals performed by any of the authors. The study is based exclusively on aggregated secondary mortality data, and ethics approval was therefore not required.

References

Alai, D. H., Arnold, S., & Sherris, M. (2015). Modelling cause-of-death mortality and the impact of cause-elimination. Annals of Actuarial Science, 9(1), 167-186.

Asplund, J., Fritchman, C., & Soyer, R. (2025). Copula modeling of COVID-19 excess mortality. Risks, 13(7), 119. https://doi.org/10.3390/risks13070119

Basellini, U., Camarda, C. G., & Booth, H. (2023). Thirty years on: A review of the Lee-Carter method for forecasting mortality. International Journal of Forecasting, 39(3), 1033-1049. https://doi.org/10.1016/j.ijforecast.2022.11.002

Bett, N., Maina, R., & Irura, Z. (2023). Dependency modeling approach of cause-related mortality and longevity risks: HIV/AIDS. Risks, 11(2), 38. https://doi.org/10.3390/risks11020038

Booth, H., & Tickle, L. (2008). Mortality modelling and forecasting: A review of methods. Annals of Actuarial Science, 3(1-2), 3-43. https://doi.org/10.1017/S1748499500000440

Cairns, A. J. G., Blake, D., & Dowd, K. (2006). A two-factor model for stochastic mortality with parameter uncertainty: Theory and calibration. Journal of Risk and Insurance, 73(4), 687-718.

Chang, L., & Shi, Y. (2023). Forecasting mortality rates with a coherent ensemble averaging approach. ASTIN Bulletin, 53(1), 2-28. https://doi.org/10.1017/asb.2022.23

Currie, I. D., Durban, M., & Eilers, P. H. C. (2004). Smoothing and forecasting mortality rates. Statistical Modelling, 4(4), 279-298.

Debón, A., Haberman, S., & Piscopo, G. (2025). A holistic approach to interpretable modelling and forecasting of human mortality by gender and country. Annals of Operations Research. Advance online publication. https://doi.org/10.1007/s10479-025-06779-2

De Mori, L., Haberman, S., Millossovich, P., & Zhu, R. (2025). Mortality forecasting via multi-task neural networks. ASTIN Bulletin, 55(2), 313-331. https://doi.org/10.1017/asb.2025.10

Diana, A., Wong, J. S. T., & Pittea, A. (2025). A unified Bayesian framework for mortality model selection. Annals of Actuarial Science, 19(3), 462-481. https://doi.org/10.1017/S1748499525100146

Euthum, M., Scherer, M., & Ungolo, F. (2024). A neural network approach for the mortality analysis of multiple populations: A case study on data of the Italian population. European Actuarial Journal, 14(2), 495-524. https://doi.org/10.1007/s13385-024-00377-5

Fu, W., Smith, B. R., Brewer, P., & Droms, S. (2023). Markov-switching Bayesian vector autoregression model in mortality forecasting. Risks, 11(9), 152. https://doi.org/10.3390/risks11090152

Hyndman, R. J., & Ullah, S. (2007). Robust forecasting of mortality and fertility rates: A functional data approach. Computational Statistics & Data Analysis, 51(10), 4942-4956.

Lee, R. D., & Carter, L. R. (1992). Modeling and forecasting U.S. mortality. Journal of the American Statistical Association, 87(419), 659-671.

Lei, X., Spivey, M., & Soyer, R. (2023). Copula models of COVID-19 mortality in Minnesota and Wisconsin. Risks, 11(11), 193. https://doi.org/10.3390/risks11110193

Li, H., Tan, K. S., Tuljapurkar, S., & Zhu, W. (2021). Gompertz law revisited: Forecasting mortality with a multi-factor exponential model. Insurance: Mathematics and Economics, 99, 268-281.

Li, J. S.-H., & Lee, R. (2005). Coherent forecasting of mortality rates: A structural approach. Journal of Econometrics, 126(2), 479-501.

Li, L., Li, H., & Panagiotelis, A. (2025). Boosting domain-specific models with shrinkage: An application in mortality forecasting. International Journal of Forecasting, 41(1), 191-207. https://doi.org/10.1016/j.ijforecast.2024.05.001

Li, R., Zhou, R., & Pitt, D. (2026). Beyond annual data: Mortality forecasting with mixed frequency data. Insurance: Mathematics and Economics, 126, 103172. https://doi.org/10.1016/j.insmatheco.2025.103172

Li, X. (2023). Nonlinear modeling of mortality data and its implications for longevity bond pricing. Risks, 11(12), 207. https://doi.org/10.3390/risks11120207

Plat, R. (2009). On stochastic mortality modeling. Insurance: Mathematics and Economics, 45(3), 393-404.

Renshaw, A. E., & Haberman, S. (2006). A cohort-based extension to the Lee-Carter model for mortality reduction factors. Insurance: Mathematics and Economics, 38(3), 556-570.

Richards, S. J. (2024). Robust mortality forecasting in the presence of outliers. British Actuarial Journal, 29, e19. https://doi.org/10.1017/S1357321724000175

Scognamiglio, S. (2024). Multi-population mortality modelling and forecasting with divergence bounds. Annals of Operations Research. Advance online publication. https://doi.org/10.1007/s10479-023-05808-2

Tanaka, S., & Matsuyama, N. (2025). An interpretable neural network approach to cause-of-death mortality forecasting. Annals of Actuarial Science, 19(3), 1-20. https://doi.org/10.1017/S1748499524000319

Villegas, A. M., Kaishev, V. K., & Millossovich, P. (2018). StMoMo: An R package for stochastic mortality modeling. Journal of Statistical Software, 84(3), 1-38.

Wu, Y., Chen, A., Xu, Y., Pan, G., & Zhu, W. (2024). Modeling mortality with Kernel Principal Component Analysis (KPCA) method. Annals of Actuarial Science, 18(3), 476-493. https://doi.org/10.1017/S1748499524000277

Zheng, H., Wang, H., Zhu, R., & Xue, J.-H. (2025a). A brief review of deep learning methods in mortality forecasting. Annals of Actuarial Science, 20(1), 150-165. https://doi.org/10.1017/S1748499525100110

Zheng, H., Wang, H., Zhu, R., & Xue, J.-H. (2025b). Fine-grained mortality forecasting with deep learning. Annals of Actuarial Science, 20, 1-27. https://doi.org/10.1017/S1748499525100171

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- USAe.csv
- dynamicfactorcopulamortalitymodelfixedv5paperholdout3d.r
- USAdlongidr.csv
- USAmlongidr.csv