



City Research Online

City, University of London Institutional Repository

Citation: Shang, H. & Haberman, S. (2025). Forecasting Age Distribution of Deaths: Cumulative Distribution Function Transformation. *Insurance: Mathematics and Economics*, 122, pp. 249-261. doi: 10.1016/j.insmatheco.2025.03.007

This is the accepted version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/34941/>

Link to published version: <https://doi.org/10.1016/j.insmatheco.2025.03.007>

Copyright: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

Reuse: Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Forecasting Age Distribution of Deaths: Cumulative Distribution Function Transformation

Han Lin Shang  *

Department of Actuarial Studies and Business Analytics
Macquarie University

Steven Haberman 

Bayes Business School
City St George's, University of London

Abstract

Like density functions, period life-table death counts are nonnegative and have a constrained integral, and thus live in a constrained nonlinear space. Implementing established modelling and forecasting methods without obeying these constraints can be problematic for such nonlinear data. We introduce cumulative distribution function transformation to forecast the life-table death counts. Using the Japanese life-table death counts obtained from the [Japanese Mortality Database \(2024\)](#), we evaluate the point and interval forecast accuracies of the proposed approach, which compares favourably to an existing compositional data analytic approach. The improved forecast accuracy of life-table death counts is of great interest to demographers for estimating age-specific survival probabilities and life expectancy and actuaries for determining temporary annuity prices for different ages and maturities.

Keywords: Constrained functional time series; Life-table death counts; Principal component analysis; Quantile density; Single-premium temporary annuity

*Postal address: Department of Actuarial Studies and Business Analytics, 4 Eastern Road, Macquarie University, NSW 2109, Australia; Email: hanlin.shang@mq.edu.au; Telephone: +61(2) 9850 4689.

1 Introduction

Density-valued objects are common, examples of which include income distribution (Kneip & Utikal 2001), financial return distribution (Kokoszka et al. 2019), distributions of the times when bids are submitted in an online auction (Jank et al. 2008), bike sharing network in Paris (Jiang 2023), and age distribution of deaths in demography (Shang & Haberman 2020), to name only a few.

In demography, the age distribution of deaths provides important insights into longevity and lifespan variability that cannot be grasped directly from either the central mortality rate or the survival function. For modelling the age distribution of deaths, Oeppen (2008) demonstrated that using compositional data analysis (CoDa) to forecast the age distribution of death counts does not lead to more pessimistic results than forecasting age-specific mortality. The pessimistic mortality rate forecasts may also be due to the linear extrapolation (constant trend) in the well-known Lee-Carter model; in that case, the mortality rate forecasts are higher than the actual ones. Thus, the age distribution of death counts is more suitable than the central mortality rates for computing period life expectancy and annuity values. In addition to providing an informative description of the mortality experience of a population, the age-at-death distribution yields readily available information on the “central longevity indicators” (e.g., mean, median and modal age at death, see Cheung et al. (2005) and Canudas-Romo (2010)) and lifespan variability (e.g., Robine 2001, Vaupel et al. 2011, van Raalte & Caswell 2013, van Raalte et al. 2014, Aburto & van Raalte 2018).

To model and forecast a time series of density-valued curves, Bergeron-Boucher et al. (2017) have put forward a principal component approach to forecast life-table death counts within a CoDa framework by treating life-table death count (d_x) for each age x as compositional data. The data are constrained to vary between two limits (e.g., 0 and a constant upper bound), which in turn imposes constraints upon the variance-covariance structure of the raw data. To remove these constraints, a standard approach involves breaking the sum constraint using a transformation of the raw data before applying conventional statistical techniques to the transformed data. The most popular transformation is the centred log-ratio (clr) transformation dated back to pioneering work by Aitchison & Shen (1980) and Aitchison (1982, 1986).

An issue with the clr transformation in the CoDa framework is the possible presence of zero counts. Although some ad-hoc ways exist of handling the zero counts (see, e.g., Martín-Fernández et al. 2013, Fry et al. 2000), we introduce a novel transformation that naturally handles zero counts. In our method, we first convert an age distribution of death for each year to a probability by dividing its radix 10^5 . So, the life-table death count at each age can be seen as a composition that sums to one. Via the cumulative sum, we transform a probability distribution function into a cumulative distribution function (CDF). In doing so, it requires a fewer number of constraints than the life-table death counts, with an additional benefit of monotonicity (see also Mayhew & Smith 2013). The inverse of CDF is quantile, which is a key quantity in Wasserstein distance for measuring discrepancy between two distributions (see, e.g., Dubey & Müller 2022). With a time series of CDFs, we model its pattern via a logistic transformation. Within this unconstrained space,

we apply a functional time series forecasting method to obtain the h -step-ahead curve prediction for a chosen horizon h . Taking the inverse logistic transformation, we obtain the h -step-ahead life-table death count forecasts after first-order differencing.

Using the Japanese age- and sex-specific life-table death counts from 1975 to 2022, we evaluate and compare one- to 16-step-ahead point and interval forecast accuracy between the CoDa and CDF transformations. To evaluate point forecast accuracy, we use the Kullback-Leibler divergence and the Jensen-Shannon divergence. To assess the interval forecast accuracy, we utilise the interval score of [Gneiting & Raftery \(2007\)](#) and [Gneiting & Katzfuss \(2014\)](#); see Section 4 for details. The improved forecast accuracy of life-table death counts is of great importance to demographers and actuaries for determining remaining age-specific life expectancies and to actuaries for pricing temporary annuities for various ages and maturities.

The remainder of this paper is organised as follows: Section 2 describes the Japanese data set obtained from the [Japanese Mortality Database \(2024\)](#), which is the age- and sex-specific life-table death counts in Japan from 1975 to 2022. Section 3 introduces the CDF transformation for producing the point and interval forecasts of the age distribution of life-table death counts. Using the point and interval forecast errors in Section 4, we evaluate and compare the point and interval forecast accuracies among the methods considered. Section 5 applies the CDF transformation to estimate the single-premium temporary immediate annuity prices for various ages and maturities for female and male policyholders residing in Japan. Conclusions are presented in Section 6, along with some ideas on how the method presented here can be further extended.

2 Age- and sex-specific life-table death counts in Japan

In many developed nations, such as Japan, increases in longevity risk and an aging population has led to concerns regarding the sustainability of the pension, health and aged care systems (see, e.g., [Coulmas 2007](#), [OECD 2020](#)). Japan has one of the highest average life expectancies in the world, with exceptional longevity in Okinawa ([Poulain 2011](#)). Japan is also one of the countries that have freely available subnational mortality data, alongside Australia, Canada, France and the US.

In Figure 1, we display Japanese age- and sex-specific life-table death counts from 1975 to 2022, obtained from the [Japanese Mortality Database \(2024\)](#). We use data from the period after the First and Second World Wars to obtain a more stable parameter estimate from the historical data. We study life-table death counts, where the life-table radix (i.e., a population experiencing 100,000 births annually) is fixed at 100,000 at age 0 while the remaining number of people alive is 0 in the last age group for each year. For the life-table death counts, there are 111 ages, and these are ages 0, 1, ..., 109, 110+. Due to rounding, there are potentially zero counts for age 110+ at some years. To overcome this problem, we work with the probability of dying (i.e., q_x) and the life-table radix to recalculate our estimated death counts (up to six decimal places). In doing so, we obtain more precise death counts than the ones reported in the [Japanese Mortality Database \(2024\)](#).

Figure 1 demonstrates a decreasing trend in infant death counts and a typical negatively skewed distribution for the life-table death counts, where the peaks shift to higher ages for both sexes. This shift is a main driver of longevity risk, which is a major issue for insurers and pension funds, especially in the selling and risk management of annuity products (see [Denuit et al. 2007](#), for a discussion). In demography, it becomes increasingly popular to model and forecast *period* life-table death counts. By modelling the life-table death counts, we can understand a redistribution of life-table death counts, where deaths at younger ages gradually shift towards older ages.

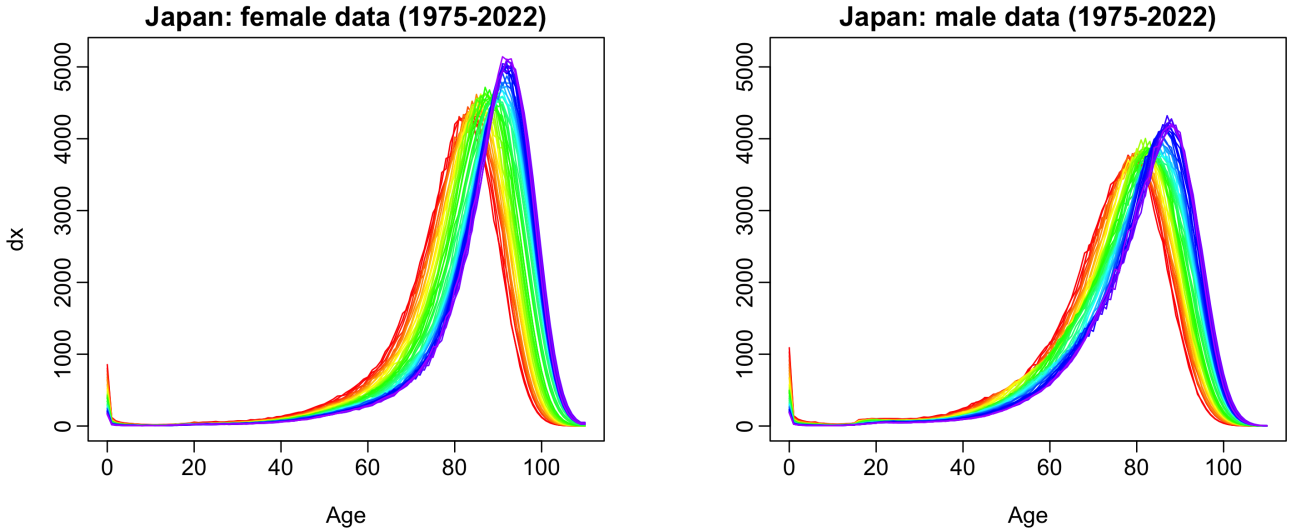


Figure 1: Rainbow plots of the age distribution of deaths from 1975 to 2022 in a single-year group. The life-table radix is 100,000 for each year. The life-table death counts in the oldest years are shown in red, while the most recent years are in violet. Curves are ordered chronologically according to the rainbow colours.

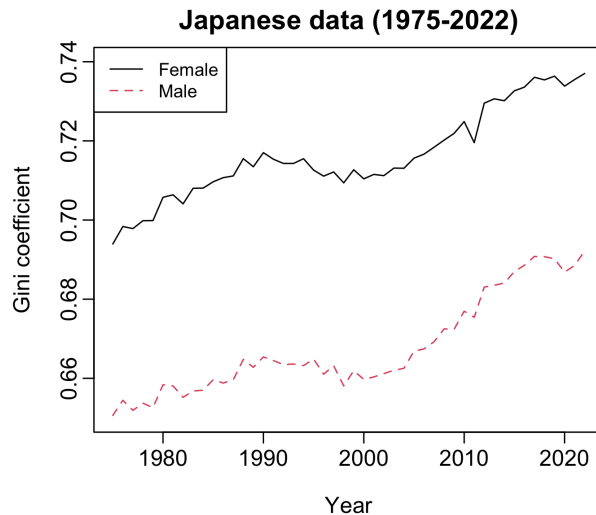


Figure 2: Gini coefficients for Japanese age- and sex-specific life-table death counts from 1975 to 2022. When the Gini coefficient approaches 0, it indicates perfect inequality across ages in the life-table death counts. When the Gini coefficient approaches 1, it indicates perfect equality.

Lifespan variability can be assessed using measures like the interquartile range or the Gini

coefficient (see [Wilmoth & Horiuchi 1999](#), [van Raalte & Caswell 2013](#), [Debón et al. 2017](#)). In life-table death counts, a Gini coefficient of 0 represents perfect inequality among ages, while 1 indicates perfect equality, with deaths occurring at the same age. Figure 2 shows higher Gini coefficients for females, reflecting greater inequality among males. While not our focus, quantifying and examining temporal changes in distribution spread is a key advantage of modeling age distributions of deaths.

3 Forecasting methods for density-valued objects

Let us denote age-specific life-table death counts as $d_{t,i}^s$, where t denotes a year, s denotes either the female or male series, and i denotes an age. For each year t , the life-table death counts sum to a radix 10^5 . From the viewpoints of proportions and CDFs, it is convenient to scale the life-table death counts to sum to one.

3.1 Cumulative distribution function transformation

We describe the CDF transformation in the following steps:

- 1) Compute the empirical cumulative distribution function via cumulative sum,

$$D_{t,x}^s = \sum_{i=1}^x d_{t,i}^s, \quad x = 1, \dots, 111, \quad t = 1, \dots, n,$$

where $D_{t,111}^s = 1$ and s represents female or male. In Figure 3, we display a time series of the empirical CDFs of the Japanese data from 1975 to 2022.

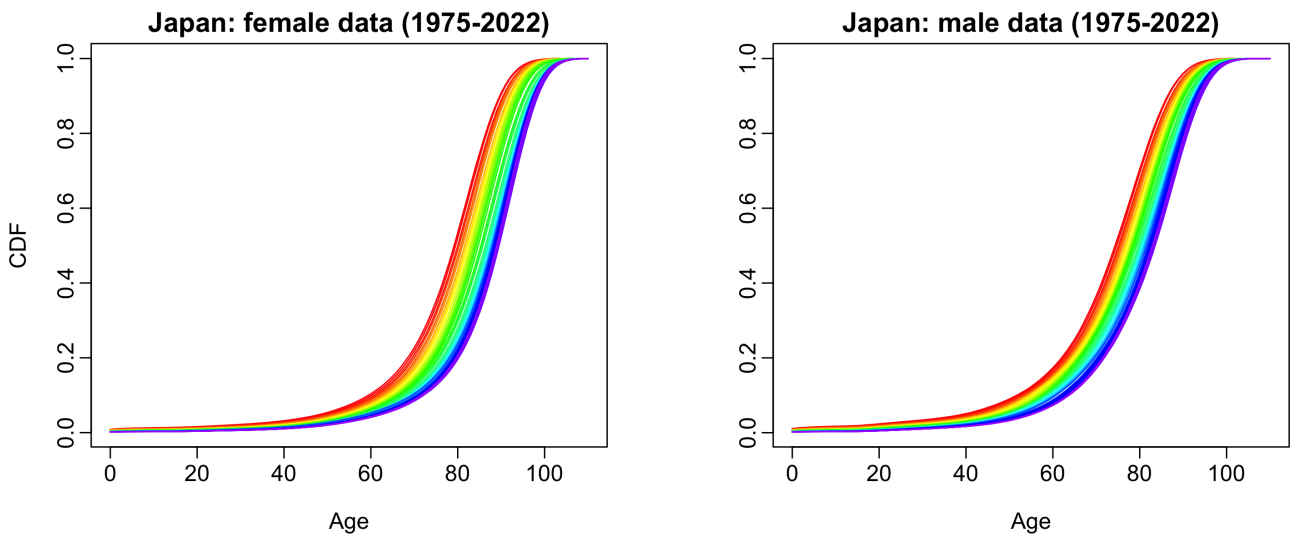


Figure 3: Empirical CDFs of the Japanese age- and sex-specific life-table death counts from 1975 to 2022.

2) Perform the logistic transformation,

$$Z_{t,x}^s = \text{logit}(D_{t,x}^s) = \ln \left(\frac{D_{t,x}^s}{1 - D_{t,x}^s} \right),$$

where $\ln(\cdot)$ denotes the natural logarithm. In Figure 4, we present the logit transformation of the empirical CDFs.

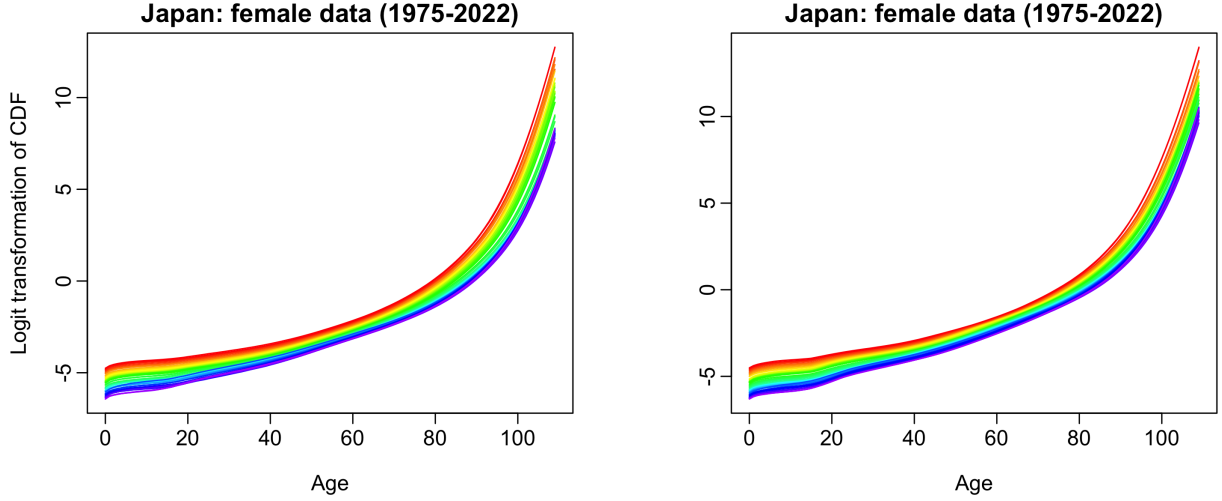


Figure 4: Empirical CDFs of the Japanese age- and sex-specific life-table death counts from 1975 to 2022.

3a) Univariate functional time series (UFTS) method: With a set of unconstrained data $\mathbf{Z}^s = (\mathbf{Z}_1^s, \dots, \mathbf{Z}_n^s)^\top$ and $\mathbf{Z}_1^s = (Z_{1,1}^s, \dots, Z_{n,111}^s)$, we consider a functional time series forecasting technique (see, e.g., Kokoszka & Reimherr 2017, Haghbin & Maadooliat 2024). It is assumed that the observations $\mathbf{Z}^s$ are elements of the Hilbert space equipped with an inner product. Each function is a square-integrable function with a finite variance.

Assuming the existence of the mean and finite variance, its non-negative definite covariance function is given by

$$\begin{aligned} \text{Cov}^s(x, y) &:= \text{Cov}(\mathbf{Z}_x^s, \mathbf{Z}_y^s) \\ &= E(\mathbf{Z}_x^s - \mu_x^s)(\mathbf{Z}_y^s - \mu_y^s), \end{aligned}$$

where $\mu_x^s = \frac{1}{n} \sum_{t=1}^n Z_{t,x}^s$, x or y represents a continuum in a function support range $\mathcal{I} \in \mathbb{R}$ and $\mathbb{R}$ is the real line. Via Mercer's lemma, there exists a set of orthonormal sequence $\psi_{k,x}^s$ of continuous function and a non-increasing sequence λ_k of positive numbers, such that

$$\text{Cov}^s(x, y) = \sum_{k=1}^{\infty} \lambda_k^s \psi_{k,x}^s \psi_{k,y}^s,$$

where λ_k^s and $\psi_{k,x}^s$ denote the k^{th} eigenvalue and eigenfunction, respectively. The Karhunen-

Loève expansion of a stochastic process $Z_{t,x}^s$ can be expressed as

$$Z_{t,x}^s = \sum_{k=1}^{K_s} \eta_{t,k}^s \psi_{k,x}^s + \varepsilon_{t,x}^s,$$

where $\psi_{k,x}^s$ denotes the k^{th} orthonormal principal component for age x and sex s , $\eta_{t,k}^s = \langle Z_{t,x}^s, \psi_{k,x}^s \rangle$ is the estimated principal component score at time t , $\varepsilon_{t,x}^s$ denotes the model residual function for age x and sex s in year t , and K_s denotes the number of retained functional principal components. We consider an eigenvalue ratio (EVR) criterion of [Li et al. \(2020\)](#) to select the optimal value of K_s , as well as $K_s = 6$ used in [Hyndman et al. \(2013\)](#), which is the default opinion in the demography package ([Hyndman 2023](#)) in [R \(R Core Team 2024\)](#).

- 3b) Multivariate functional time series (MFTS) method: The female and male series are stacked in a vector $\mathbf{Z}_{t,\zeta} = (Z_{t,x}^F, Z_{t,x}^M) \in \mathbb{R}^2$ for a given age x . The number of ages in the stacked series is the sum of the number of ages in the two series. Before performing an eigendecomposition to $\mathbf{Z} = (\mathbf{Z}_{1,\zeta}, \dots, \mathbf{Z}_{n,\zeta})^\top$, it is essential to centre each series by removing its mean and standard deviation. Applying eigendecomposition to $\mathbf{Z}$, it gives

$$Z_{t,x}^s = \sum_{k=1}^K \eta_{t,k}^s \psi_{k,x}^s + \epsilon_{t,x}^s,$$

where $\psi_{k,x}^F$ corresponds to the first 111 rows of the estimated principal components, while $\psi_{k,x}^M$ corresponds to the remaining 111 rows, and K denotes the common number of principal components shared by both series. The value of K can be determined by the EVR criterion or fixed at $K = 6$.

- 3c) Multilevel functional time series (MLFTS) method: The joint modelling of the female and male series bears a strong resemblance to the two-way analysis of variance studied by [Zhang \(2013\)](#) and [Jiménez-Varón et al. \(2024\)](#). In the MLFTS method, the essential idea is to decompose each series into its mean function, a common trend across series $R^{(c)}$, and a series-specific residual trend U^s . It can be expressed as

$$Z_{t,x}^s = \mu_x^s + R_{t,x}^c + U_{t,x}^s,$$

where $\mathbf{R}_x^c = (R_{1,x}^c, \dots, R_{n,x}^c)^\top$ and $\mathbf{U}_x^s = (U_{1,x}^s, \dots, U_{n,x}^s)^\top$ can be modeled via two-stage principal component analysis. To model the common patterns effectively, identifying an auxiliary series that both series share is crucial. We suggest using a simple average of the series as a starting point for the first-stage functional principal component analysis. Then, each residual series is independently modelled by the second-stage functional principal component analysis.

- 4) Forecasting $Z_{n+h,x}^s$: By conditioning on the estimated functional principal components $\Psi_x = (\psi_{1,x}, \dots, \psi_{K,x})$ and observed data $Z_x^s = (Z_{1,x}^s, \dots, Z_{n,x}^s)$, the h -step-ahead forecast of $Z_{n+h,x}^s$ can be obtained

$$\hat{Z}_{n+h|n,x}^s = E[Z_{n+h,x}^s | \Psi_x, Z_x^s] = \sum_{k=1}^K \hat{\eta}_{n+h|n,k}^s \psi_{k,x}^s,$$

where $\hat{\eta}_{n+h|n,k}^s$ denotes the h -step-ahead univariate time-series forecast of the principal component scores. Among the univariate time-series methods, we consider the exponential smoothing method of [Hyndman et al. \(2008\)](#) for forecasting scores. In Figure 5, we present the 16 steps ahead forecasts of CDFs from 2023 to 2038 using the UFTS method.

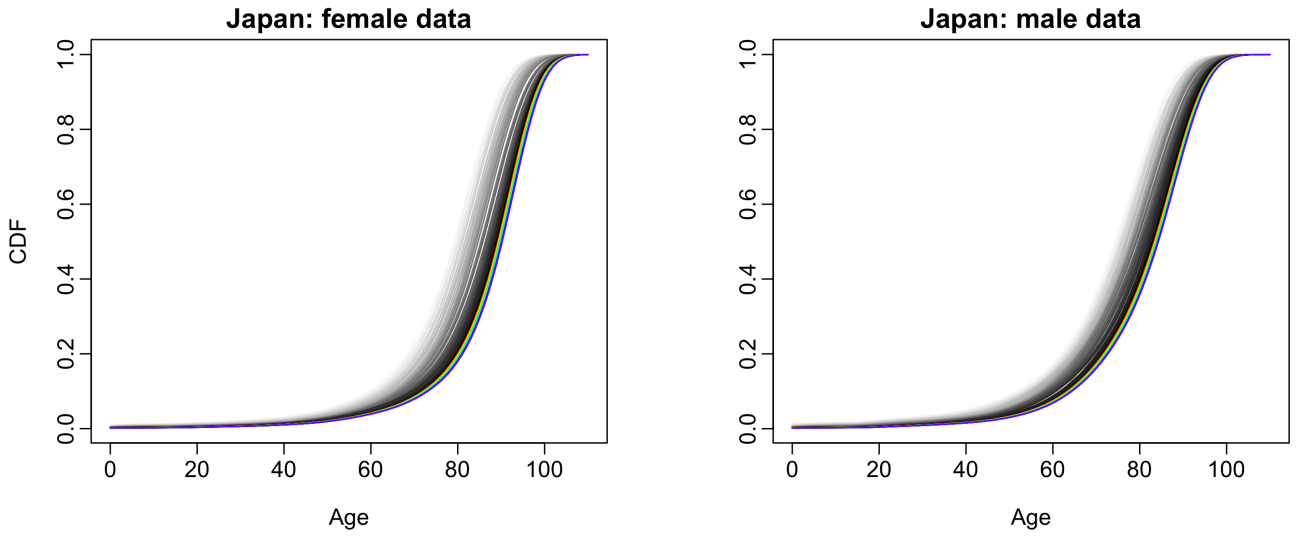


Figure 5: The 16 steps ahead forecasts of CDFs from 2023 to 2038 were obtained by the UFTS method. While the data from the distant past are shown in grey, the forecasts are shown in rainbow colours.

- 5) By taking the inverse logit transformation, we obtain

$$\hat{D}_{n+h|n,x}^s = \frac{\exp^{\hat{Z}_{n+h|n,x}^s}}{1 + \exp^{\hat{Z}_{n+h|n,x}^s}}$$

- 6) By taking the first-order differencing, we obtain

$$\hat{d}_{n+h|n,x}^s = \Delta_{i=1}^x \hat{D}_{n+h|n,i}^s,$$

where Δ represents the first-order differencing. In Figure 6, we display the 16 steps ahead forecasts of life-table death counts.

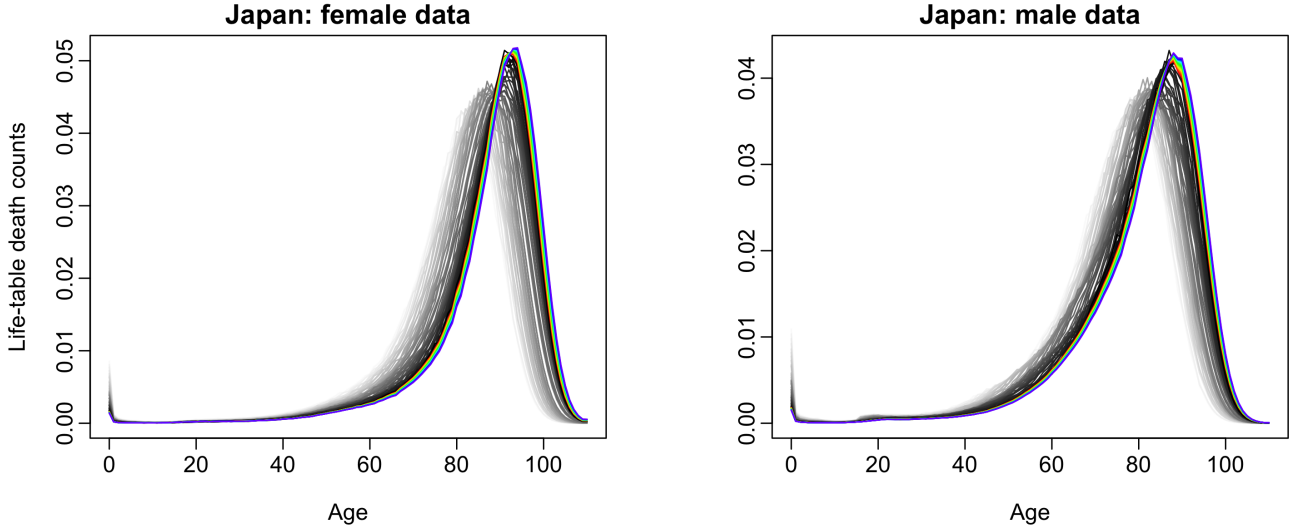


Figure 6: The 16 steps ahead forecasts of the life-table death counts from 2023 to 2038.

3.2 Centered log-ratio transformation

As a benchmark, the centred log-ratio transformation presents one of many possible ways to deal with the non-negativity constraint by transforming the raw data. The clr transformation (see, e.g., [Aitchison 1982, 1986](#)) is a mapping between the simplex to a hyperplane in Euclidean space. PCA can be applied directly to this hyperplane. The algorithm for implementing the CoDa method consists of the following steps:

- 1) Compute the geometric mean function, which can be estimated by a simple average

$$\alpha_{n,x} = \exp \left(\frac{1}{n} \sum_{t=1}^n \ln d_{t,x} \right),$$

where $d_{t,x} > 0$ denotes the age-specific life-table death counts. We compute the centred data by dividing each death count by its geometric mean

$$s_{t,x} = \frac{d_{t,x}}{\alpha_{n,x}}.$$

- 2) By taking the nature logarithm, we obtain

$$\begin{aligned} \beta_{t,x} &= \ln s_{t,x} \\ &= \ln d_{t,x} - \ln \alpha_{n,x} \\ &= \ln d_{t,x} - \frac{1}{n} \sum_{t=1}^n \ln d_{t,x} \end{aligned}$$

where $\beta_x = (\beta_{1,x}, \dots, \beta_{n,x})^\top$ can be viewed as an unconstrained functional time series.

3) Applying eigendecomposition to β_x gives

$$\beta_{t,x} = \sum_{k=1}^K \gamma_{t,k} \phi_{k,x} + \omega_{t,x}, \quad (1.1)$$

where $\phi_{k,x}$ is the k^{th} estimated principal component for age x , $\gamma_{t,k} = \langle \beta_{t,x}, \phi_{k,x} \rangle$ is the k^{th} estimated principal component score at time t , and $\omega_{t,x}$ denotes the model residual function for age x in year t . When $K = 1$, (1.1) mimics the spirit of Lee & Carter's (1992) method. Within the CoDa framework, this method is known as the Lee-Carter-CoDa method.

4) Forecasting $\beta_{n+h,x}$: By conditioning on the estimated functional principal components $\Phi_x = (\phi_{1,x}, \dots, \phi_{K,x})$ and observed data β_x , the h -step-ahead forecast of $\beta_{n+h,x}$ can be obtained

$$\hat{\beta}_{n+h|n,x} = E[\beta_{n+h,x} | \Phi_x, \beta_x] = \sum_{k=1}^K \hat{\gamma}_{n+h|n,k} \phi_{k,x},$$

where $\hat{\gamma}_{n+h|n,k}$ denotes the h -step-ahead univariate time-series forecast of the principal component scores.

5) Transform back to the compositional data. We take the inverse-centered log-ratio transformation given by

$$\hat{s}_{n+h|n,x} = \exp^{\hat{\beta}_{n+h|n,x}}$$

6) We add back the geometric means to obtain the life-table death count forecasts $d_{n+h,x}$,

$$\hat{d}_{n+h|n,x} = \hat{s}_{n+h|n,x} \times \alpha_{n,x},$$

where $\alpha_{n,x}$ denotes the simple mean given in Step 1).

4 Analysis of the Japanese life-table death counts

4.1 Expanding window

An expanding window analysis of a time series model is commonly used to assess model and parameter stability over time, and prediction accuracy. The expanding window analysis determines the constancy of a model's parameter by computing parameter estimates and their forecasts over an expanding window through the sample (for details Zivot & Wang 2006, pp. 313-314). Using the first 32 observations from 1975 to 2006 in the Japanese female and male age-specific life-table death counts, we produce one- to 16-step-ahead forecasts. Through an expanding-window approach, we estimate the parameters in the time series forecasting models using the first 33 observations from 1975 to 2007. Forecasts from the estimated models are then produced for one- to 15-step-ahead

point and interval forecasts. We iterate this process by increasing the sample size by one year until reaching the end of the data period in 2022. This process produces 16 one-step-ahead forecasts, 15 two-step-ahead forecasts, $\dots$, and one 16-step-ahead forecast. We compare these forecasts with the holdout samples to determine the accuracy of the out-of-sample forecast.

4.2 Point forecast evaluation

We consider the expanding window scheme to assess the point forecast as described in [Zivot & Wang \(2006, Chapter 9\)](#). As the age distribution of death counts can be considered a probability density function, we consider density evaluation metrics. They include discrete Kullback-Leibler divergence ([Kullback & Leibler 1951](#)) and the Jensen-Shannon divergence ([Shannon 1948](#)).

The Kullback-Leibler divergence measures information loss by replacing an unknown density with an approximation. For two probability density functions, denoted by $d_{n+\xi}(u)$ and $\hat{d}_{n+\xi|n}(u)$, the discrete Kullback-Leibler divergence is defined as

$$\begin{aligned} \text{KLD}(h) &= D_{\text{KL}}(d_{n+\xi,x} \| \hat{d}_{n+\xi|n,x}) + D_{\text{KL}}(\hat{d}_{n+\xi|n,x} \| d_{n+\xi,x}) \\ &= \frac{1}{111 \times (17-h)} \sum_{\xi=h}^{16} \sum_{x=1}^{111} d_{n+\xi,x} \cdot (\ln d_{n+\xi,x} - \ln \hat{d}_{n+\xi|n,x}) + \\ &\quad \frac{1}{111 \times (17-h)} \sum_{\xi=h}^{16} \sum_{x=1}^{111} \hat{d}_{n+\xi|n,x} \cdot (\ln \hat{d}_{n+\xi|n,x} - \ln d_{n+\xi,x}), \end{aligned}$$

where ξ denotes the forecasting period. The discrete Kullback-Leibler divergence is symmetric and non-negative.

An alternative is given by the Jensen-Shannon divergence, which is a symmetrised and smoothed version of the Kullback-Leibler divergence. It is defined by

$$\text{JSD}(h) = \frac{1}{2} D_{\text{KL}}(d_{n+\xi,x} \| \delta_{n+\xi,x}) + \frac{1}{2} D_{\text{KL}}(\hat{d}_{n+\xi|n,x} \| \delta_{n+\xi,x}),$$

where $\delta_{n+\xi,x}$ measures a common quantity between $d_{n+\xi,x}$ and $\hat{d}_{n+\xi|n,x}$. We consider the geometric mean $\delta_{n+\xi,x} = \sqrt{d_{n+\xi,x} \hat{d}_{n+\xi|n,x}}$.

4.3 Interval forecast evaluation

For each year in the forecasting period, the h -step-ahead prediction intervals were computed at the $(1 - \alpha)$ nominal coverage probabilities, where α denotes a level of significance. Let us denote $\hat{d}_{n+\xi|n,x}^{\text{lb}}$ and $\hat{d}_{n+\xi|n,x}^{\text{ub}}$ as the lower and upper bounds, respectively. We compute the empirical coverage probability, defined as the proportion of observations that fall into the calculated

prediction intervals. It can be expressed as

$$\text{ECP}_\alpha(h) = 1 - \frac{\sum_{\xi=h}^{16} \sum_{x=1}^{111} \left[\mathbb{1}(d_{n+\xi,x} < \hat{d}_{n+\xi|n,x}^{\text{lb}}) + \mathbb{1}(d_{n+\xi,x} > \hat{d}_{n+\xi|n,x}^{\text{ub}}) \right]}{111 \times (17 - h)},$$

where the denominator is the total number of observations in the forecasting period, and $\mathbb{1}(\cdot)$ represents the binary indicator function.

We compute the coverage probability difference (CPD), which is the absolute difference between the nominal and empirical coverage probabilities, and use this measure to evaluate the accuracy of the interval forecast (see also [Swanson & Beck 1994](#), [Tayman et al. 2007](#)). The CPD can be expressed as

$$\text{CPD}_\alpha(h) = |\text{ECP}_\alpha(h) - (1 - \alpha)|.$$

The smaller the value of CPD is, the more accurate the interval forecast one method produces.

Although the ECP and CPD are measures of interval forecast accuracy, neither consider the sharpness of the prediction intervals, i.e., the distance between the lower and upper bounds. We also consider the interval score of [Gneiting & Raftery \(2007\)](#), which is defined as

$$S_\alpha(\hat{d}_{n+\xi|n,x}^{\text{lb}}, \hat{d}_{n+\xi|n,x}^{\text{ub}}, d_{n+\xi,x}) = (\hat{d}_{n+\xi|n,x}^{\text{ub}} - \hat{d}_{n+\xi|n,x}^{\text{lb}}) + \frac{2}{\alpha}(\hat{d}_{n+\xi|n,x}^{\text{lb}} - d_{n+\xi,x})\mathbb{1}(d_{n+\xi,x} < \hat{d}_{n+\xi|n,x}^{\text{lb}}) \\ + \frac{2}{\alpha}(d_{n+\xi,x} - \hat{d}_{n+\xi|n,x}^{\text{ub}})\mathbb{1}(d_{n+\xi,x} > \hat{d}_{n+\xi|n,x}^{\text{ub}}).$$

The interval score rewards a narrow prediction interval if and only if the true observation lies within the prediction interval. The optimal interval score is achieved when $d_{n+\xi,x}$ lies between $\hat{d}_{n+\xi|n,x}^{\text{lb}}$ and $\hat{d}_{n+\xi|n,x}^{\text{ub}}$ and the distance between $\hat{d}_{n+\xi|n,x}^{\text{lb}}$ and $\hat{d}_{n+\xi|n,x}^{\text{ub}}$ is minimal for age x .

For different ages and years in the forecasting period, the mean interval score is defined by

$$\bar{S}_\alpha(h) = \frac{\sum_{\xi=h}^{16} \sum_{x=1}^{111} S_\alpha(\hat{d}_{n+\xi|n,x}^{\text{lb}}, \hat{d}_{n+\xi|n,x}^{\text{ub}}, d_{n+\xi,x})}{111 \times (17 - h)},$$

where $S_\alpha(\hat{d}_{n+\xi|n,x}^{\text{lb}}, \hat{d}_{n+\xi|n,x}^{\text{ub}}, d_{n+\xi,x})$ denotes the interval score at the ξ^{th} curve in the testing period.

4.4 Comparisons of point forecast accuracy

In Table 1, we present the one- to 16-step-ahead point forecast errors of the life-table death counts in Japan. The errors are measured by the symmetrised Kullback-Leibler Divergence (KLD) and Jensen-Shannon Divergence (JSD), which are commonly used for evaluating density forecasts. Since the errors were minuscule, we multiplied them by 100 to highlight the differences among the methods. Between the CDF and clr transformations, the CDF transformation coupled with the MLFTS method produces the smallest errors for almost all forecast horizons. Between the two

methods of selecting the number of components, using $K = 6$ proves advantageous, confirming the early finding of [Hyndman et al. \(2013\)](#). Additionally, between the female and male series, the latter yields smaller errors. For male data, the clr transformation performs better at longer horizons. Although that could be due to data randomness, the finding motivates us to consider a model-averaging approach by assigning weights to forecasts obtained from the two transformations.

Table 1: Comparison of the point forecast errors ($\times 100$) of the age distribution of deaths between the CDF and clr transformations, when the number of components is fixed at $K = 6$. Among the CDF transformation, we also consider the UFTS, MFTS and MLFTS methods. The smallest errors are shown in bold.

Sex	h	UFTS		MFTS		MLFTS		clr	
		KLD	JSD	KLD	JSD	KLD	JSD	KLD	JSD
F	1	0.0762	0.0193	0.1010	0.0255	0.0812	0.0206	0.0992	0.0252
	2	0.1313	0.0334	0.1617	0.0408	0.1175	0.0300	0.1629	0.0414
	3	0.1994	0.0506	0.2129	0.0530	0.1492	0.0380	0.2342	0.0593
	4	0.3039	0.0774	0.2997	0.0746	0.2105	0.0539	0.3373	0.0857
	5	0.4225	0.1078	0.4010	0.0991	0.2769	0.0709	0.4466	0.1136
	6	0.5365	0.1370	0.4911	0.1203	0.3245	0.0832	0.5582	0.1420
	7	0.6890	0.1749	0.5649	0.1369	0.3702	0.0945	0.7334	0.1840
	8	0.9445	0.2375	0.5236	0.1297	0.3956	0.1021	0.9441	0.2344
	9	1.1893	0.2994	0.5401	0.1376	0.4295	0.1126	1.0786	0.2672
	10	1.2043	0.3103	0.5681	0.1499	0.4729	0.1251	1.0099	0.2562
	11	1.0678	0.2842	0.5243	0.1409	0.5909	0.1578	0.7761	0.2050
	12	1.2257	0.3368	0.7628	0.2056	0.7742	0.2088	1.0570	0.2827
	13	1.7173	0.4787	1.0650	0.2924	0.9434	0.2570	1.3935	0.3759
	14	2.0706	0.5826	1.3167	0.3667	1.0168	0.2760	1.7221	0.4647
	15	3.0846	0.8887	1.6464	0.4619	1.6175	0.4472	2.8137	0.7742
	16	4.0159	1.1744	2.3856	0.6839	3.3945	0.9802	4.0931	1.1474
	Mean	1.1799	0.3246	0.7228	0.1949	0.6978	0.1911	1.0912	0.2912
M	1	0.0882	0.0222	0.0894	0.0225	0.0799	0.0202	0.0830	0.0209
	2	0.1212	0.0305	0.1346	0.0337	0.1116	0.0282	0.1270	0.0319
	3	0.1606	0.0403	0.1863	0.0463	0.1399	0.0352	0.1863	0.0465
	4	0.2093	0.0524	0.2575	0.0639	0.1889	0.0475	0.2652	0.0659
	5	0.2706	0.0677	0.3378	0.0834	0.2404	0.0605	0.3546	0.0880
	6	0.3500	0.0875	0.4231	0.1038	0.2931	0.0736	0.4575	0.1131
	7	0.4321	0.1075	0.4988	0.1215	0.3495	0.0873	0.5953	0.1459
	8	0.5309	0.1320	0.5326	0.1309	0.4075	0.1018	0.7716	0.1866
	9	0.6153	0.1538	0.5494	0.1354	0.4164	0.1046	0.9298	0.2233

Continued on next page

Sex	h	UFTS		MFTS		MLFTS		clr	
		KLD	JSD	KLD	JSD	KLD	JSD	KLD	JSD
	10	0.6590	0.1659	0.5204	0.1296	0.3345	0.0857	0.9376	0.2286
	11	0.6095	0.1550	0.4766	0.1209	0.3148	0.0814	0.6860	0.1706
	12	0.5926	0.1510	0.4342	0.1109	0.3261	0.0846	0.3517	0.0886
	13	0.6250	0.1599	0.4773	0.1230	0.3630	0.0952	0.3436	0.0866
	14	0.6539	0.1677	0.4701	0.1209	0.3914	0.1032	0.3110	0.0788
	15	0.6204	0.1605	0.5733	0.1506	0.5622	0.1523	0.2314	0.0594
	16	0.5630	0.1471	0.6282	0.1666	0.8016	0.2191	0.2417	0.0641
	Mean	0.4439	0.1126	0.4119	0.1040	0.3325	0.0863	0.4296	0.1062

In Table 2, we display a comparison of the point forecast errors of the life expectancy using a terminal age of 111. Based on the forecasted age distribution of deaths from the CDF and clr transformations, we compute their corresponding age-specific life expectancies and compare these values with the holdout data. To measure the forecast accuracy, we consider the root mean squared forecast error (RMSFE) and mean absolute forecast error (MAFE). For forecasting age-specific life expectancy, the CDF transformation coupled with the MLFTS method is recommended.

Table 2: Comparison of the point forecast errors of the life expectancy between the CDF and clr transformations, when the number of components is fixed at $K = 6$.

Sex	h	UFTS		MFTS		MLFTS		clr	
		RMSFE	MAFE	RMSFE	MAFE	RMSFE	MAFE	RMSFE	MAFE
F	1	0.2358	0.1756	0.2555	0.1920	0.2142	0.1458	0.2861	0.1953
	2	0.3061	0.2186	0.3127	0.2485	0.2451	0.1770	0.3732	0.2715
	3	0.3776	0.2724	0.3323	0.2885	0.2566	0.1987	0.4529	0.3458
	4	0.4738	0.3446	0.3790	0.3395	0.3041	0.2502	0.5512	0.4318
	5	0.5598	0.4129	0.4247	0.3708	0.3461	0.2770	0.6354	0.5152
	6	0.6224	0.4559	0.4467	0.3726	0.3340	0.2716	0.7042	0.5829
	7	0.7139	0.5445	0.4581	0.3654	0.3442	0.2786	0.8125	0.6685
	8	0.8542	0.6361	0.4172	0.3565	0.3764	0.2907	0.9254	0.7777
	9	0.9556	0.7346	0.4002	0.3355	0.3956	0.3159	0.9673	0.8301
	10	0.9274	0.7281	0.4143	0.3663	0.3068	0.2382	0.9263	0.8147
	11	0.7744	0.6485	0.3338	0.2892	0.3109	0.2383	0.7865	0.7195
	12	0.7065	0.6729	0.4767	0.4390	0.4646	0.3803	0.9765	0.8982
	13	0.8900	0.8316	0.5998	0.5229	0.5241	0.4265	1.1285	1.0424
	14	1.0003	0.9321	0.6986	0.5830	0.5202	0.4625	1.2804	1.1860

Continued on next page

Sex	h	UFTS		MFTS		MLFTS		clr	
		RMSFE	MAFE	RMSFE	MAFE	RMSFE	MAFE	RMSFE	MAFE
M	15	1.2642	1.1776	0.7784	0.7053	0.7311	0.7071	1.6265	1.5007
	16	1.4548	1.3989	0.9476	0.9200	1.2572	1.2121	1.9369	1.8094
	Mean	0.7573	0.6366	0.4797	0.4184	0.4332	0.3669	0.8981	0.7869
	1	0.2373	0.1898	0.2303	0.1726	0.2002	0.1331	0.2476	0.1918
	2	0.2788	0.2259	0.2738	0.2140	0.2343	0.1625	0.3250	0.2475
	3	0.3236	0.2649	0.3104	0.2506	0.2720	0.1960	0.4160	0.3042
	4	0.3754	0.3038	0.3615	0.3010	0.3219	0.2311	0.5160	0.3647
	5	0.4408	0.3529	0.4128	0.3444	0.3699	0.2762	0.6170	0.4460
	6	0.5136	0.4187	0.4738	0.4105	0.4122	0.3181	0.7188	0.5161
	7	0.5854	0.4713	0.5268	0.4606	0.4651	0.3713	0.8369	0.5934
	8	0.6681	0.5376	0.5696	0.4904	0.5314	0.4379	0.9646	0.6740
	9	0.7655	0.6147	0.6034	0.5063	0.5573	0.4675	1.0710	0.7540
	10	0.8418	0.7122	0.6248	0.5235	0.4731	0.3893	1.0932	0.8219
	11	0.8411	0.7125	0.6244	0.5209	0.4381	0.3526	0.9364	0.7041
	12	0.8274	0.6877	0.6399	0.5116	0.3994	0.3322	0.6205	0.4960
	13	0.8338	0.6825	0.6620	0.5316	0.3853	0.3115	0.6218	0.5038
	14	0.8406	0.6753	0.6649	0.5246	0.3871	0.3180	0.5815	0.4670
	15	0.7650	0.6142	0.6044	0.4757	0.3640	0.3010	0.4227	0.3554
	16	0.5948	0.5013	0.4569	0.4068	0.3638	0.2777	0.1855	0.1570
	Mean	0.6083	0.4978	0.5025	0.4153	0.3859	0.3048	0.6359	0.4748

In Appendix A, we present the comparisons of point forecast accuracy of the age distribution of deaths and life expectancy obtained from the CDF and clr transformations when the number of components is selected by the EVR criterion.

4.5 Comparisons of interval forecast accuracy

The interval forecast errors are measured by the mean interval score and coverage probability deviation (CPD). In Table 3, we present the one- to 16-step-ahead interval forecast errors of the life-table death counts in Japan under a nominal coverage probability of 80% when the number of components is fixed at $K = 6$. Between the CDF and clr transformations, the CDF transformation coupled with the MLFTS method produces the smallest interval scores for the shorter forecast horizon. Conversely, for the female series, the clr transformation results in the smallest errors over a relatively longer forecast horizon. Regarding the CPD, the CDF transformation with the MLFTS method is advantageous for shorter forecast horizons. For relatively longer forecast horizons,

the CDF transformation using the univariate functional time series method is recommended. Additionally, between the female and male series, the latter yields smaller errors.

Table 3: Comparison of the interval forecast errors between the CDF and clr transformations at the 80% nominal coverage probability.

Sex	h	UFTS		MFTS		MLFTS		clr (EVR)		clr (K = 6)	
		score	CPD	score	CPD	score	CPD	score	CPD	score	CPD
F	1	0.0016	0.0761	0.0019	0.1499	0.0013	0.0218	0.0017	0.0153	0.0015	0.0080
	2	0.0018	0.0787	0.0021	0.1303	0.0016	0.0174	0.0020	0.0006	0.0018	0.0186
	3	0.0022	0.0662	0.0023	0.1402	0.0019	0.0143	0.0024	0.0385	0.0022	0.0314
	4	0.0024	0.0517	0.0026	0.1279	0.0023	0.0162	0.0027	0.0219	0.0024	0.0344
	5	0.0029	0.0416	0.0030	0.1159	0.0028	0.0335	0.0031	0.0544	0.0029	0.0634
	6	0.0032	0.0583	0.0035	0.1230	0.0031	0.0051	0.0034	0.1181	0.0031	0.0903
	7	0.0036	0.0505	0.0038	0.1198	0.0035	0.0018	0.0036	0.1162	0.0035	0.1126
	8	0.0039	0.0368	0.0042	0.1279	0.0039	0.0248	0.0042	0.1299	0.0038	0.1229
	9	0.0041	0.0255	0.0047	0.1178	0.0043	0.0108	0.0044	0.1572	0.0040	0.1347
	10	0.0045	0.0417	0.0052	0.1228	0.0046	0.0391	0.0049	0.1562	0.0045	0.1447
	11	0.0048	0.0604	0.0059	0.1204	0.0049	0.0423	0.0052	0.1655	0.0047	0.1384
	12	0.0057	0.1009	0.0066	0.1495	0.0058	0.0793	0.0059	0.1586	0.0056	0.1351
	13	0.0067	0.1009	0.0074	0.1459	0.0069	0.0806	0.0070	0.1662	0.0067	0.1279
	14	0.0076	0.1009	0.0083	0.1369	0.0073	0.0589	0.0074	0.1610	0.0075	0.1159
	15	0.0085	0.1009	0.0094	0.1414	0.0089	0.0063	0.0083	0.1640	0.0086	0.1279
	16	0.0098	0.1009	0.0110	0.1459	0.0109	0.0432	0.0097	0.1550	0.0099	0.1279
	Mean	0.0046	0.0682	0.0051	0.1322	0.0046	0.0310	0.0047	0.1112	0.0045	0.0959
M	1	0.0016	0.1459	0.0019	0.1662	0.0012	0.0078	0.0017	0.0367	0.0014	0.0654
	2	0.0018	0.1345	0.0020	0.1604	0.0014	0.0018	0.0019	0.0505	0.0016	0.0775
	3	0.0021	0.1421	0.0021	0.1505	0.0016	0.0076	0.0023	0.0346	0.0019	0.0726
	4	0.0024	0.1362	0.0022	0.1612	0.0019	0.0141	0.0026	0.0420	0.0023	0.0884
	5	0.0027	0.1354	0.0024	0.1542	0.0023	0.0087	0.0029	0.0551	0.0025	0.0649
	6	0.0030	0.1181	0.0027	0.1443	0.0026	0.0075	0.0033	0.0452	0.0028	0.0739
	7	0.0033	0.0919	0.0029	0.1171	0.0030	0.0063	0.0036	0.0252	0.0031	0.0622
	8	0.0036	0.0709	0.0032	0.1149	0.0032	0.0268	0.0040	0.0238	0.0034	0.0569
	9	0.0037	0.0412	0.0034	0.0919	0.0033	0.0333	0.0042	0.0232	0.0034	0.0311
	10	0.0039	0.0211	0.0037	0.0829	0.0033	0.0842	0.0037	0.0520	0.0031	0.0327
	11	0.0040	0.0012	0.0041	0.0784	0.0035	0.0784	0.0034	0.1024	0.0032	0.0814
	12	0.0043	0.0234	0.0044	0.0667	0.0039	0.0973	0.0040	0.1820	0.0036	0.1459
	13	0.0048	0.0477	0.0048	0.0514	0.0044	0.1189	0.0043	0.1865	0.0038	0.1212

Continued on next page

Sex	h	UFTS		MFTS		MLFTS		clr (EVR)		clr (K = 6)	
		score	CPD	score	CPD	score	CPD	score	CPD	score	CPD
	14	0.0050	0.0372	0.0050	0.0859	0.0045	0.1069	0.0042	0.1880	0.0040	0.1489
	15	0.0053	0.0207	0.0057	0.1054	0.0050	0.1189	0.0047	0.1910	0.0046	0.1595
	16	0.0058	0.0072	0.0062	0.1550	0.0056	0.1459	0.0053	0.1820	0.0054	0.1640
	Mean	0.0036	0.0734	0.0035	0.1179	0.0032	0.0540	0.0035	0.0888	0.0031	0.0904

When the number of components is selected by the EVR criterion, the additional results are presented in Appendix B. In Appendix C, we also display the one- to 16-step-ahead interval forecast errors under a nominal coverage probability of 95%.

5 Application to a single-premium temporary immediate annuity

An important application of mortality forecasts for those individuals at ages over 60 is in the pension and superannuation industries, whose profitability and solvency crucially depend on accurate forecasts to appropriately hedge longevity risks. When an individual retires, an optimal way of guaranteeing one individual's financial income in retirement is to purchase an annuity (Yaari 1965). An annuity is a financial contract offered by insurers that guarantees a steady stream of income payments for either a temporary or lifetime of the annuitants in exchange for an initial premium fee.

We study temporary annuities, which have gained popularity in developed countries (e.g., Australia and the USA), because lifetime immediate annuities, where rates are locked in for life, have been shown to deliver poor value for money (see, e.g., Cannon & Tonks 2008, Chapter 6). These temporary annuities pay a pre-determined and guaranteed income level higher than the income provided by a lifetime annuity for a similar premium. In combination with deferred annuities, fixed-term annuities offer an attractive alternative to lifetime immediate annuities.

Using the CDF transformation, we apply the multilevel functional time series method to obtain forecasts of age-specific life-table death counts and then determine the corresponding age-specific survival probabilities. In Figure 7, we display the age-specific life-table death count forecasts from 2023 to 2072 for Japanese females and males.

With the forecasts of life-table death counts, we input them into the calculation of single-premium term immediate annuities, and we adopt a cohort approach to calculating survival probabilities. The τ year survival probability of a person aged x currently at $t = 0$ is determined by

$${}_{\tau}p_x = \prod_{j=1}^{\tau} p_{x+j-1} = \prod_{j=1}^{\tau} (1 - q_{x+j-1}) = \prod_{j=1}^{\tau} \left(1 - \frac{d_{x+j-1}}{l_{x+j-1}}\right),$$

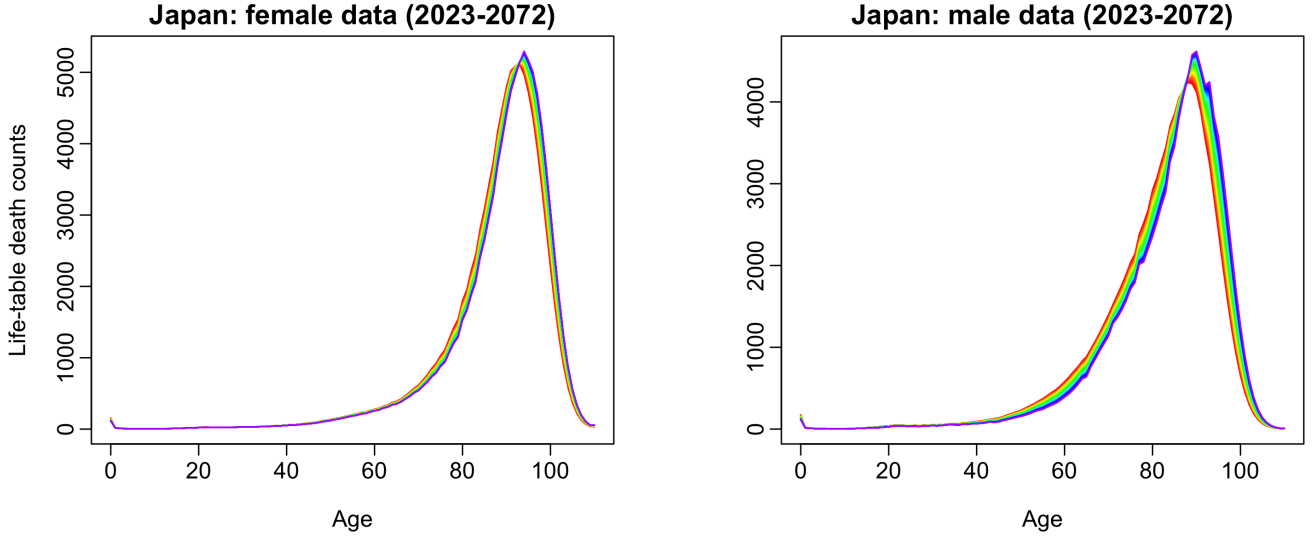


Figure 7: Age distribution of death count forecasts from 2023 to 2072 for Japanese females and males.

where d_{x+j-1} denotes the number of death counts between ages $x + j - 1$ and $x + j$ and l_{x+j-1} denotes the number of lives alive at age $x + j - 1$.

The price of an annuity with a maturity term of a T year is a random variable since it depends on the value of zero-coupon bond price and mortality forecasts. The annuity price written for an x -year-old with the benefit of one Japanese Yen per year is given by

$$a_x^T = \sum_{\tau=1}^T B(0, \tau) {}_{\tau}p_x,$$

where $B(0, \tau)$ is the τ -year bond price and ${}_{\tau}p_x$ denotes the survival probability.

Table 4: Estimates of annuity prices with different ages and maturities (T) for female and male populations residing in Japan. These estimates are based on forecast mortality from 2023 to 2072 and a constant interest rate $\eta = 0.25\%$. We consider only contracts with maturity so that age + maturity ≤ 110 .

Age	Female						Male					
	$T = 5$	10	15	20	25	30	5	10	15	20	25	30
60	4.916	9.671	14.215	18.467	22.274	25.364	4.859	9.430	13.606	17.258	20.224	22.285
65	4.895	9.574	13.951	17.871	21.052	23.133	4.804	9.193	13.032	16.150	18.316	19.418
70	4.855	9.397	13.464	16.766	18.924	19.851	4.712	8.832	12.178	14.504	15.687	16.050
75	4.784	9.067	12.544	14.818	15.794	15.996	4.578	8.295	10.878	12.192	12.596	12.652
80	4.635	8.398	10.859	11.915	12.133	12.150	4.352	7.376	8.914	9.387	9.453	9.457
85	4.324	7.152	8.366	8.617	8.636		3.901	5.885	6.494	6.580	6.585	
90	3.720	5.316	5.647	5.672			3.187	4.165	4.303	4.310		
95	2.829	3.415	3.461				2.335	2.664	2.682			
100	1.851	1.993					1.551	1.636				
105	1.176						1.043					

In Table 4, to provide an example of the annuity calculations, we compute the point estimate of the annuity prices for various ages and maturities for female and male populations in Japan. We assume a constant current interest rate at $\eta = 0.25\%$ (<https://tradingeconomics.com/japan/interest-rate>) and hence zero-coupon bond is given as $B(0, \tau) = \exp^{-\eta\tau}$. In Appendix D, we present the best estimate of annuity prices under a constant interest rate of 3%. Although the difference in annuity price might appear small, any mispricing can result in a significant risk when considering a large contract value and annuity portfolio.

6 Conclusion

We introduce a CDF transformation to model and forecast the age distribution of deaths. Using the Japanese age- and sex-specific life-table death counts, we evaluate and compare the point and interval forecast accuracies between the CDF and clr transformations. Based on the error measures, the CDF transformation coupled with the MLFTS method is generally recommended for forecasting age distribution of deaths and the corresponding age-specific life expectancy. The superiority of this method is also driven by the ability to model female and male age distribution of deaths jointly.

We apply the CDF transformation to forecast life-table death counts from 2023 to 2072. We then calculate the cumulative survival probability and obtain the best estimates of temporary annuity prices. As expected, we find that the cumulative survival probability has a pronounced impact on annuity prices. For reproducibility, the computer  code is available at https://github.com/hanshang/CDF_transformation.

There are a few ways in which this paper could be extended, and we briefly discussed four.

- 1) A robust MLFTS method may be considered in the presence of outlying years.
- 2) A model-averaging approach can improve accuracy for forecasting the age distribution of death counts.
- 3) The methodology can be applied to calculate other annuity prices, such as the whole-life immediate or deferred annuity.
- 4) We may also consider a *cohort* life table that depicts the life history of a specific group of individuals (see, e.g., [Hatzopoulos & Haberman 2015](#), [Basellini et al. 2020](#)), but is dependent on projected mortality rates for those cohorts born more recently.

Acknowledgements

The first author acknowledges the financial support from an Australian Research Council Future Fellowship, FT240100338.

A Additional results of point forecast accuracy comparison

In Table 5, we compare the point forecast errors between the CDF and clr transformations when the number of components is selected by the EVR criterion. As measured by the KLD and JSD, the CDF transformation coupled with the MLFTS method performs the best with the smallest averaged errors.

Table 5: Comparison of the point forecast errors ($\times 100$) between the CDF and clr transformations when the number of components is selected by the EVR criterion.

Sex	h	UFTS		MFTS		MLFTS		clr	
		KLD	JSD	KLD	JSD	KLD	JSD	KLD	JSD
F	1	0.1011	0.0255	0.5286	0.1411	0.0863	0.0219	0.1524	0.0385
	2	0.1649	0.0419	0.6040	0.1621	0.1277	0.0325	0.2389	0.0605
	3	0.2410	0.0611	0.6589	0.1771	0.1443	0.0365	0.3365	0.0850
	4	0.3594	0.0915	0.7364	0.1988	0.1906	0.0482	0.4785	0.1210
	5	0.4908	0.1252	0.7784	0.2110	0.2535	0.0640	0.6321	0.1598
	6	0.6162	0.1573	0.7728	0.2099	0.2889	0.0729	0.7832	0.1980
	7	0.7847	0.1993	0.7584	0.2061	0.3512	0.0880	1.0116	0.2535
	8	1.0604	0.2672	0.7425	0.2014	0.3394	0.0855	1.2974	0.3225
	9	1.3333	0.3363	0.7935	0.2155	0.3098	0.0791	1.5357	0.3815
	10	1.3743	0.3539	0.7971	0.2163	0.3083	0.0790	1.5183	0.3823
	11	1.2208	0.3236	0.8236	0.2226	0.3827	0.0990	1.2251	0.3155
	12	1.2978	0.3568	1.1603	0.3199	0.4369	0.1126	1.0567	0.2824
	13	1.8033	0.5035	1.2738	0.3536	0.5649	0.1475	1.4206	0.3817
	14	2.1521	0.6062	1.3081	0.3635	0.5291	0.1368	1.7751	0.4758
	15	3.2044	0.9241	1.8190	0.5143	0.8501	0.2245	2.8910	0.7893
	16	4.1412	1.2129	2.0189	0.5716	1.9506	0.5419	4.1828	1.1635
	Mean	1.2716	0.3491	0.9734	0.2678	0.4446	0.1169	1.2835	0.3382
M	1	0.0983	0.0248	0.1563	0.0401	0.0973	0.0245	0.1919	0.0480
	2	0.1328	0.0334	0.1923	0.0494	0.1285	0.0323	0.2710	0.0676
	3	0.1718	0.0430	0.2268	0.0582	0.1588	0.0398	0.3579	0.0888
	4	0.2196	0.0550	0.2728	0.0700	0.2026	0.0509	0.4641	0.1149
	5	0.2774	0.0694	0.3164	0.0812	0.2572	0.0646	0.5811	0.1434
	6	0.3530	0.0881	0.3584	0.0919	0.3213	0.0806	0.7151	0.1759
	7	0.4318	0.1074	0.4015	0.1028	0.3788	0.0945	0.8803	0.2148
	8	0.5290	0.1315	0.4478	0.1144	0.4497	0.1122	1.0840	0.2614
	9	0.6148	0.1535	0.4962	0.1266	0.4929	0.1230	1.2672	0.3035

Continued on next page

Sex	h	UFTS		MFTS		MLFTS		clr	
		KLD	JSD	KLD	JSD	KLD	JSD	KLD	JSD
	10	0.6564	0.1652	0.5062	0.1294	0.4918	0.1234	1.2893	0.3123
	11	0.6125	0.1557	0.4831	0.1237	0.3435	0.0891	0.9209	0.2268
	12	0.5962	0.1519	0.5057	0.1301	0.3467	0.0902	0.3806	0.0957
	13	0.6304	0.1613	0.5326	0.1376	0.3710	0.0975	0.3724	0.0937
	14	0.6669	0.1711	0.5431	0.1402	0.3699	0.0979	0.3355	0.0848
	15	0.6473	0.1673	0.5767	0.1509	0.5385	0.1457	0.2462	0.0630
	16	0.5901	0.1537	0.5596	0.1473	0.9476	0.2628	0.2468	0.0653
	Mean	0.4518	0.1145	0.4110	0.1059	0.3685	0.0956	0.6003	0.1475

In Table 6, we compare the point forecast errors (as measured by RMSFE and MAFE) of the age-specific life expectancy between the CDF and clr transformation under the EVR criterion. The CDF transformation, coupled with the MLFTS method, gives the smallest forecast errors.

Table 6: Comparison of the point forecast errors of the life expectancy between the CDF and clr transformations, when the number of components is selected by the EVR criterion.

Sex	h	UFTS		MFTS		MLFTS		clr	
		RMSFE	MAFE	RMSFE	MAFE	RMSFE	MAFE	RMSFE	MAFE
F	1	0.2682	0.1967	0.5308	0.4200	0.2231	0.1736	0.3551	0.2779
	2	0.3410	0.2346	0.5630	0.4416	0.2826	0.2191	0.4607	0.3637
	3	0.4157	0.2898	0.5761	0.4703	0.2842	0.2202	0.5608	0.4324
	4	0.5179	0.3669	0.6045	0.4920	0.3333	0.2586	0.6870	0.5321
	5	0.6072	0.4358	0.6064	0.4968	0.3987	0.3272	0.8003	0.6251
	6	0.6727	0.4818	0.5728	0.4628	0.4095	0.3102	0.8896	0.6969
	7	0.7669	0.5696	0.5407	0.4323	0.4863	0.3920	1.0176	0.7927
	8	0.9082	0.6603	0.5043	0.4149	0.4628	0.3614	1.1597	0.9066
	9	1.0160	0.7600	0.5088	0.4446	0.4111	0.3412	1.2673	1.0041
	10	0.9950	0.7612	0.4736	0.4184	0.3350	0.2842	1.2661	1.0068
	11	0.8346	0.6767	0.4471	0.4188	0.3282	0.2755	1.1243	0.9277
	12	0.7088	0.6739	0.6328	0.5486	0.3454	0.2885	0.9909	0.9087
	13	0.8919	0.8341	0.6776	0.5731	0.3651	0.3169	1.1649	1.0740
	14	1.0003	0.9352	0.6847	0.5788	0.2570	0.1972	1.3319	1.2290
	15	1.2775	1.1902	0.8604	0.7441	0.3391	0.3009	1.6866	1.5502
	16	1.4662	1.4099	0.8776	0.8514	0.8056	0.7786	1.9981	1.8592
	Mean	0.7930	0.6548	0.6038	0.5130	0.3792	0.3153	1.0476	0.8867

Continued on next page

Sex	h	UFTS		MFTS		MLFTS		clr	
		RMSFE	MAFE	RMSFE	MAFE	RMSFE	MAFE	RMSFE	MAFE
M	1	0.2597	0.2116	0.2852	0.2209	0.2101	0.1201	0.4447	0.3361
	2	0.3028	0.2512	0.3380	0.2605	0.2373	0.1506	0.5448	0.4084
	3	0.3516	0.2899	0.3856	0.2931	0.2597	0.1721	0.6407	0.4573
	4	0.4054	0.3298	0.4415	0.3400	0.3000	0.2195	0.7420	0.5138
	5	0.4701	0.3777	0.4974	0.3922	0.3546	0.2720	0.8428	0.5894
	6	0.5429	0.4422	0.5555	0.4509	0.4083	0.3149	0.9455	0.6657
	7	0.6127	0.4910	0.6137	0.4999	0.4539	0.3633	1.0582	0.7498
	8	0.6956	0.5553	0.6664	0.5372	0.5178	0.4233	1.1791	0.8306
	9	0.7952	0.6374	0.7103	0.5757	0.5914	0.4869	1.2803	0.8994
	10	0.8703	0.7289	0.7219	0.5946	0.6313	0.5057	1.3014	0.9483
	11	0.8723	0.7369	0.6997	0.5893	0.4865	0.3936	1.0912	0.7971
	12	0.8512	0.7069	0.6846	0.5588	0.4453	0.3665	0.6330	0.5167
	13	0.8529	0.6996	0.7035	0.5708	0.4059	0.3321	0.6240	0.5190
	14	0.8601	0.6916	0.7061	0.5677	0.3328	0.2808	0.5718	0.4672
	15	0.7898	0.6364	0.6313	0.5049	0.3066	0.2528	0.4091	0.3462
	16	0.6301	0.5333	0.4684	0.4113	0.4724	0.4278	0.1757	0.1455
Mean		0.6352	0.5200	0.5693	0.4605	0.4009	0.3176	0.7803	0.5744

B Additional results of interval forecast accuracy comparison

In Table 7, we present the comparison of the interval forecast accuracy between the CDF and clr transformation at the 80% nominal coverage probability when the number of components is selected by the EVR criterion. For the female data, the clr transformation produces the smallest averaged interval score, while the CDF transformation coupled with the MLFTS method gives the smallest averaged CPD. For the male data, the clr transformation gives the smallest interval forecast errors. The mixed results motivate us to consider a model-averaging approach by assigning weights to forecasts obtained from the two transformations.

Table 7: Comparison of the interval forecast errors between the CDF and clr transformations at the 80% nominal coverage probability when the number of components is selected by the EVR criterion.

Sex	h	UFTS		MFTS		MLFTS		clr	
		score	CPD	score	CPD	score	CPD	score	CPD
F	1	0.0037	0.1842	0.0019	0.1482	0.0017	0.0654	0.0017	0.0153
	2	0.0039	0.1820	0.0021	0.1279	0.0021	0.0517	0.0020	0.0006
	3	0.0041	0.1910	0.0023	0.1337	0.0024	0.0629	0.0024	0.0385
	4	0.0043	0.1931	0.0027	0.1272	0.0029	0.0621	0.0027	0.0219
	5	0.0046	0.1925	0.0030	0.1092	0.0033	0.0363	0.0031	0.0544
	6	0.0049	0.1943	0.0034	0.1247	0.0039	0.0943	0.0034	0.1181
	7	0.0051	0.1964	0.0038	0.1261	0.0042	0.1000	0.0036	0.1162
	8	0.0054	0.1970	0.0043	0.1239	0.0048	0.1089	0.0042	0.1299
	9	0.0055	0.1944	0.0047	0.1144	0.0051	0.1020	0.0044	0.1572
	10	0.0059	0.1858	0.0053	0.1254	0.0055	0.1215	0.0049	0.1562
	11	0.0061	0.1895	0.0059	0.1264	0.0060	0.1249	0.0052	0.1655
	12	0.0067	0.1802	0.0065	0.1477	0.0069	0.1207	0.0059	0.1586
	13	0.0076	0.1595	0.0073	0.1392	0.0078	0.1099	0.0070	0.1662
	14	0.0082	0.1610	0.0081	0.1369	0.0082	0.1069	0.0074	0.1610
	15	0.0092	0.1550	0.0097	0.1369	0.0090	0.0784	0.0083	0.1640
	16	0.0103	0.1550	0.0097	0.1459	0.0102	0.0018	0.0097	0.1550
	Mean	0.0060	0.1819	0.0050	0.1309	0.0052	0.0842	0.0047	0.1112
M	1	0.0041	0.1803	0.0019	0.1583	0.0015	0.0722	0.0017	0.0367
	2	0.0042	0.1784	0.0020	0.1568	0.0018	0.0703	0.0019	0.0505
	3	0.0043	0.1910	0.0021	0.1537	0.0021	0.0835	0.0023	0.0346
	4	0.0045	0.1986	0.0022	0.1626	0.0024	0.0829	0.0026	0.0420
	5	0.0047	0.2000	0.0024	0.1512	0.0027	0.0679	0.0029	0.0551
	6	0.0049	0.2000	0.0026	0.1271	0.0031	0.0935	0.0033	0.0452
	7	0.0051	0.1964	0.0029	0.1297	0.0033	0.0766	0.0036	0.0252
	8	0.0053	0.2000	0.0032	0.1179	0.0038	0.0729	0.0040	0.0238
	9	0.0054	0.2000	0.0034	0.0953	0.0038	0.0716	0.0042	0.0232
	10	0.0055	0.2000	0.0038	0.0867	0.0040	0.0945	0.0037	0.0520
	11	0.0055	0.2000	0.0041	0.0664	0.0041	0.1204	0.0034	0.1024
	12	0.0057	0.2000	0.0043	0.0541	0.0045	0.1441	0.0040	0.1820
	13	0.0061	0.2000	0.0047	0.0514	0.0049	0.1482	0.0043	0.1865
	14	0.0061	0.2000	0.0050	0.0739	0.0053	0.1489	0.0042	0.1880
	15	0.0065	0.2000	0.0057	0.0919	0.0057	0.1414	0.0047	0.1910
	16	0.0068	0.2000	0.0058	0.1099	0.0064	0.1550	0.0053	0.1820
	Mean	0.0053	0.1965	0.0035	0.1117	0.0037	0.1027	0.0035	0.0888

C Comparison of the interval forecast errors between the CDF and clr transformations at the 95% nominal coverage probability

In Table 8, we present the comparison of the interval forecast accuracy between the CDF and clr transformation at the 95% nominal coverage probability. We consider two ways of selecting the number of components: EVR criterion or $K = 6$. When the number of components is selected by the EVR criterion, the CDF transformation with the MFTS method gives the smallest average scores, while the CPD transformation with the MLFTS method produces the smallest averaged CPD for the female and male series. When the number of components is fixed at $K = 6$, the CDF transformation with the UFTS method produces the smallest averaged interval scores and CPD for the female series. For the male series, the CDF transformation with the MLFTS method gives the smallest averaged interval scores and CPD.

Table 8: Comparison of the interval forecast errors between the CDF and clr transformation at the 95% nominal coverage probability. The number of components is fixed at $K = 6$ or selected by the EVR criterion.

Criterion	Sex	h	UFTS		MFTS		MLFTS		clr	
			score	CPD	score	CPD	score	CPD	score	CPD
EVR	F	1	0.0064	0.0477	0.0028	0.0382	0.0025	0.0021	0.0025	0.0226
		2	0.0065	0.0494	0.0030	0.0368	0.0029	0.0047	0.0028	0.0161
		3	0.0069	0.0500	0.0034	0.0378	0.0035	0.0034	0.0034	0.0144
		4	0.0071	0.0493	0.0039	0.0368	0.0041	0.0103	0.0039	0.0082
		5	0.0075	0.0500	0.0044	0.0335	0.0046	0.0146	0.0044	0.0072
		6	0.0082	0.0500	0.0051	0.0353	0.0058	0.0032	0.0056	0.0312
		7	0.0081	0.0500	0.0056	0.0320	0.0064	0.0041	0.0057	0.0338
		8	0.0086	0.0500	0.0061	0.0330	0.0071	0.0100	0.0063	0.0400
		9	0.0091	0.0500	0.0066	0.0354	0.0082	0.0173	0.0074	0.0399
		10	0.0095	0.0500	0.0072	0.0307	0.0090	0.0243	0.0083	0.0474
		11	0.0103	0.0500	0.0080	0.0380	0.0105	0.0260	0.0097	0.0485
		12	0.0103	0.0500	0.0088	0.0356	0.0116	0.0212	0.0095	0.0464
		13	0.0116	0.0500	0.0101	0.0252	0.0130	0.0095	0.0113	0.0410
		14	0.0122	0.0500	0.0112	0.0170	0.0140	0.0140	0.0119	0.0350
		15	0.0132	0.0500	0.0127	0.0185	0.0147	0.0050	0.0133	0.0275
		16	0.0141	0.0500	0.0148	0.0050	0.0163	0.0041	0.0151	0.0230
		Mean	0.0094	0.0498	0.0071	0.0305	0.0084	0.0108	0.0076	0.0301
	M	1	0.0062	0.0444	0.0029	0.0348	0.0022	0.0168	0.0022	0.0162
		2	0.0063	0.0470	0.0030	0.0320	0.0025	0.0230	0.0025	0.0200

Continued on next page

Criterion	Sex	h	UFTS		MFTS		MLFTS		clr	
			score	CPD	score	CPD	score	CPD	score	CPD
K = 6	F	3	0.0065	0.0494	0.0030	0.0455	0.0028	0.0268	0.0029	0.0172
		4	0.0067	0.0500	0.0033	0.0465	0.0032	0.0278	0.0035	0.0061
		5	0.0070	0.0500	0.0035	0.0477	0.0035	0.0267	0.0037	0.0155
		6	0.0075	0.0500	0.0039	0.0492	0.0044	0.0361	0.0044	0.0344
		7	0.0075	0.0500	0.0041	0.0491	0.0047	0.0311	0.0046	0.0014
		8	0.0080	0.0500	0.0044	0.0500	0.0051	0.0290	0.0050	0.0090
		9	0.0083	0.0500	0.0047	0.0455	0.0058	0.0230	0.0058	0.0018
		10	0.0086	0.0500	0.0050	0.0423	0.0062	0.0217	0.0061	0.0358
		11	0.0090	0.0500	0.0055	0.0410	0.0070	0.0305	0.0069	0.0470
		12	0.0090	0.0500	0.0057	0.0392	0.0073	0.0320	0.0066	0.0482
		13	0.0096	0.0500	0.0061	0.0410	0.0077	0.0297	0.0075	0.0477
		14	0.0098	0.0500	0.0064	0.0440	0.0083	0.0230	0.0075	0.0440
		15	0.0103	0.0500	0.0069	0.0320	0.0086	0.0185	0.0081	0.0455
		16	0.0109	0.0500	0.0075	0.0230	0.0097	0.0050	0.0090	0.0410
		Mean	0.0082	0.0494	0.0047	0.0414	0.0056	0.0250	0.0054	0.0269
K = 6	F	1	0.0023	0.0151	0.0028	0.0370	0.0018	0.0429	0.0021	0.0271
		2	0.0025	0.0068	0.0030	0.0344	0.0021	0.0419	0.0023	0.0323
		3	0.0030	0.0037	0.0034	0.0384	0.0026	0.0562	0.0029	0.0189
		4	0.0035	0.0027	0.0040	0.0361	0.0031	0.0567	0.0035	0.0179
		5	0.0039	0.0168	0.0045	0.0305	0.0036	0.0784	0.0038	0.0108
		6	0.0046	0.0081	0.0050	0.0361	0.0043	0.0385	0.0046	0.0140
		7	0.0047	0.0050	0.0055	0.0293	0.0043	0.0356	0.0046	0.0239
		8	0.0053	0.0061	0.0061	0.0350	0.0051	0.0211	0.0053	0.0350
		9	0.0055	0.0016	0.0067	0.0376	0.0053	0.0311	0.0053	0.0387
		10	0.0059	0.0101	0.0072	0.0333	0.0060	0.0195	0.0058	0.0384
		11	0.0064	0.0170	0.0081	0.0365	0.0066	0.0131	0.0063	0.0410
		12	0.0074	0.0248	0.0090	0.0356	0.0072	0.0032	0.0073	0.0374
		13	0.0087	0.0117	0.0102	0.0230	0.0085	0.0095	0.0086	0.0252
		14	0.0096	0.0170	0.0111	0.0200	0.0092	0.0020	0.0098	0.0170
		15	0.0110	0.0005	0.0127	0.0140	0.0110	0.0221	0.0112	0.0095
		16	0.0127	0.0050	0.0135	0.0050	0.0165	0.1302	0.0138	0.0050
		Mean	0.0061	0.0095	0.0070	0.0301	0.0061	0.0376	0.0061	0.0245
K = 6	M	1	0.0024	0.0370	0.0029	0.0359	0.0016	0.0176	0.0019	0.0089
		2	0.0026	0.0350	0.0029	0.0332	0.0019	0.0233	0.0022	0.0164
		3	0.0029	0.0436	0.0031	0.0423	0.0021	0.0079	0.0027	0.0140

Continued on next page

Criterion	Sex	h	UFTS		MFTS		MLFTS		clr	
			score	CPD	score	CPD	score	CPD	score	CPD
		4	0.0032	0.0375	0.0033	0.0500	0.0025	0.0103	0.0030	0.0202
		5	0.0036	0.0312	0.0036	0.0500	0.0028	0.0168	0.0032	0.0080
		6	0.0041	0.0320	0.0038	0.0459	0.0034	0.0049	0.0038	0.0181
		7	0.0044	0.0248	0.0041	0.0482	0.0037	0.0167	0.0040	0.0113
		8	0.0051	0.0230	0.0044	0.0470	0.0041	0.0081	0.0046	0.0150
		9	0.0053	0.0140	0.0047	0.0444	0.0043	0.0041	0.0047	0.0027
		10	0.0056	0.0165	0.0050	0.0397	0.0043	0.0191	0.0045	0.0024
		11	0.0054	0.0080	0.0054	0.0425	0.0045	0.0140	0.0046	0.0200
		12	0.0059	0.0023	0.0059	0.0374	0.0048	0.0140	0.0050	0.0338
		13	0.0065	0.0018	0.0062	0.0410	0.0051	0.0185	0.0054	0.0477
		14	0.0067	0.0131	0.0064	0.0440	0.0054	0.0200	0.0054	0.0440
		15	0.0070	0.0041	0.0070	0.0275	0.0056	0.0365	0.0061	0.0365
		16	0.0078	0.0050	0.0073	0.0230	0.0063	0.0230	0.0072	0.0230
		Mean	0.0049	0.0205	0.0047	0.0407	0.0039	0.0159	0.0043	0.0201

D Annuity price calculation under an interest rate $\eta = 3\%$

In Table 9, we display the annuity price calculation for female and male populations residing in Japan under a constant interest rate of $\eta = 3\%$ in line with the US, UK and Australia. In contrast to $\eta = 0.25\%$, the following annuity premiums are lower for each age and maturity period.

Table 9: Estimates of annuity prices with different ages and maturities (T) for female and male policyholders residing in Japan. These estimates are based on forecast mortality from 2023 to 2072 and a constant interest rate $\eta = 3\%$. We consider only contracts with maturity so that age + maturity ≤ 110 .

Age	T = 5	10	Female				5	10	Male			
			15	20	25	30			15	20	25	30
60	4.532	8.352	11.535	14.131	16.158	17.594	4.481	8.154	11.081	13.313	14.894	15.854
65	4.513	8.273	11.340	13.734	15.431	16.402	4.431	7.960	10.652	12.559	13.717	14.233
70	4.477	8.128	10.979	12.999	14.155	14.590	4.347	7.662	10.011	11.438	12.073	12.244
75	4.412	7.858	10.300	11.696	12.221	12.317	4.226	7.220	9.038	9.848	10.066	10.092
80	4.278	7.310	9.043	9.696	9.815	9.823	4.022	6.464	7.551	7.844	7.880	7.882
85	3.998	6.284	7.144	7.301	7.311		3.614	5.223	5.657	5.711	5.713	
90	3.450	4.749	4.985	5.002			2.965	3.764	3.863	3.868		
95	2.641	3.122	3.154				2.188	2.459	2.472			
100	1.743	1.861					1.465	1.536				
105	1.117						0.993					

References

- Aburto, J. & van Raalte, A. A. (2018), 'Lifespan dispersion in times of life expectancy fluctuation: the case of central and eastern Europe', *Demography* **55**, 2071–2096.
- Aitchison, J. (1982), 'The statistical analysis of compositional data', *Journal of the Royal Statistical Society: Series B* **44**(2), 139–177.
- Aitchison, J. (1986), *The Statistical Analysis of Compositional Data*, Chapman & Hall, London.
- Aitchison, J. & Shen, S. M. (1980), 'Logistic-normal distributions: Some properties and uses', *Biometrika* **67**(2), 261–272.
- Basellini, U., Kjaergaard, S. & Camarda, C. G. (2020), 'An age-at-death distribution approach to forecast cohort mortality', *Insurance: Mathematics and Economics* **91**, 129–143.
- Bergeron-Boucher, M.-P., Canudas-Romo, V., Oeppen, J. & Vaupel, J. W. (2017), 'Coherent forecasts of mortality with compositional data analysis', *Demographic Research* **37**, 527–566.
- Cannon, E. & Tonks, I. (2008), *Annuity Markets*, Oxford University Press, Oxford.
- Canudas-Romo, V. (2010), 'Three measures of longevity: Time trends and record values', *Demography* **47**(2), 299–312.
- Cheung, S. L. K., Robine, J.-M., Tu, E. J.-C. & Caselli, G. (2005), 'Three dimensions of the survival curve: Horizontalization, verticalization, and longevity extension', *Demography* **42**(2), 243–258.
- Coulmas, F. (2007), *Population Decline and Ageing in Japan – the Social Consequences*, Routledge, New York.
- Debón, A., Chaves, L., Haberman, S. & Villa, F. (2017), 'Characterization of between-group inequality of longevity in EU countries', *Insurance: Mathematics and Economics* **75**, 151–165.
- Denuit, M., Devolder, P. & Goderniaux, A.-C. (2007), 'Securitization of longevity risk: Pricing survivor bonds with Wang transform in the Lee-Carter framework', *The Journal of Risk and Insurance* **74**(1), 87–113.
- Dubey, P. & Müller, H.-G. (2022), 'Modeling time-varying random objects and dynamic networks', *Journal of the American Statistical Association: Theory and Methods* **117**(540), 2252–2267.
- Fry, J. M., Fry, T. R. L. & McLaren, K. R. (2000), 'Compositional data analysis and zeros in micro data', *Applied Economics* **32**(8), 953–959.
- Gneiting, T. & Katzfuss, M. (2014), 'Probabilistic forecasting', *Annual Review of Statistics and Its Application* **1**, 125–151.

- Gneiting, T. & Raftery, A. E. (2007), ‘Strictly proper scoring rules, prediction and estimation’, *Journal of the American Statistical Association: Review Article* **102**(477), 359–378.
- Haghighi, H. & Maadooliat, M. (2024), ‘A journey from univariate to multivariate functional time series: A comprehensive review’, *WIREs Computational Statistics* **16**(1), e1640.
- Hatzopoulos, P. & Haberman, S. (2015), ‘Modeling trends in cohort survival probabilities’, *Insurance: Mathematics and Economics* **64**, 162–179.
- Hyndman, R. (2023), *demography: Forecasting Mortality, Fertility, Migration and Population Data*. R package version 2.0. URL: <https://CRAN.R-project.org/package=demography>.
- Hyndman, R. J., Booth, H. & Yasmeen, F. (2013), ‘Coherent mortality forecasting: the product-ratio method with functional time series models’, *Demography* **50**(1), 261–283.
- Hyndman, R. J., Koehler, A., Ord, K. & Snyder, R. (2008), *Forecasting with Exponential Smoothing: the State Space Approach*, Springer-Verlag, Berlin, Heidelberg.
- Jank, W., Shmueli, G., Plaisant, C. & Shneiderman, B. (2008), Visualizing functional data with an application to ebay’s online auctions, in C. Chen, W. Härdle & A. Unwin, eds, ‘Handbook of Data Visualization’, Springer-Verlag, Berlin.
- Japanese Mortality Database (2024), *National Institute of Population and Social Security Research*. Available at <https://www.ipss.go.jp/p-toukei/JMD/index-en.asp> (data downloaded on July 3, 2024).
- Jiang, Y. (2023), Wasserstein multivariate auto-regressive models for modeling distributional time series and its application in graph learning, Working paper, arXiv.
URL: <https://arxiv.org/abs/2207.05442>
- Jiménez-Varón, C. F., Sun, Y. & Shang, H. L. (2024), ‘Forecasting high-dimensional functional time series: Application to sub-national age-specific mortality’, *Journal of Computational and Graphical Statistics* **33**(4), 1160–1174.
- Kneip, A. & Utikal, K. (2001), ‘Inference for density families using functional principal component analysis’, *Journal of the American Statistical Association: Theory and Methods* **96**(454), 519–542.
- Kokoszka, P., Miao, H., Petersen, A. & Shang, H. L. (2019), ‘Forecasting of density functions with an application to cross-sectional and intraday returns’, *International Journal of Forecasting* **35**, 1304–1317.
- Kokoszka, P. & Reimherr, M. (2017), *Introduction to Functional Data Analysis*, Chapman and Hall/CRC, Boca Raton.
- Kullback, S. & Leibler, R. A. (1951), ‘On information and sufficiency’, *The Annals of Mathematical Statistics* **22**(1), 79–86.

- Lee, R. D. & Carter, L. (1992), 'Modeling and forecasting the time series of U.S. mortality', *Journal of the American Statistical Association: Applications & Case Studies* **87**(419), 659–671.
- Li, D., Robinson, P. M. & Shang, H. L. (2020), 'Long-range dependent curve time series', *Journal of the American Statistical Association: Theory and Methods* **115**(530), 957–971.
- Martín-Fernández, J. A., Barceló-Vidal, C. & Pawlowsky-Glahn, V. (2013), 'Dealing with zeros and missing values in compositional data sets using nonparametric imputation', *Mathematical Geology* **35**(3), 253–278.
- Mayhew, L. & Smith, D. (2013), 'A new method of projecting populations based on trends in life expectancy and survival', *Population Studies* **67**(2), 157–170.
- OECD (2020), Pensions at a glance 2019: OECD and G20 indicators, Technical report, Organisation for Economic Co-operation and Development. Retrieved from http://https://www.oecd-ilibrary.org/social-issues-migration-health/pensions-at-a-glance-2019_b6d3dcfc-en at November 4, 2020.
- Oeppen, J. (2008), Coherent forecasting of multiple-decrement life tables: A test using Japanese cause of death data, in 'European Population Conference', Barcelona, Spain. Available online at the address <http://epc2008.princeton.edu/papers/80611>.
- Poulain, M. (2011), 'Exceptional longevity in Okinawa: A plea for in-depth validation', *Demographic Research* **25**, 245–284.
- R Core Team (2024), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.
URL: <https://www.R-project.org/>
- Robine, J.-M. (2001), 'Redefining the stages of the epidemiological transition by a study of the dispersion of life spans: the case of France', *Population: An English Selection* **13**(1), 173–193.
- Shang, H. L. & Haberman, S. (2020), 'Forecasting age distribution of death counts: An application to annuity pricing', *Annals of Actuarial Science* **14**(1), 150–169.
- Shannon, C. E. (1948), 'A mathematical theory of communication', *Bell Labs Technical Journal* **27**(3), 379–423.
- Swanson, D. A. & Beck, D. M. (1994), 'A new short-term county population projection method', *Journal of Economic and Social Measurement* **20**(1), 25–50.
- Tayman, J., Smith, S. K. & Lin, J. (2007), 'Precision, bias, and uncertainty for state population forecasts: An exploratory analysis of time series models', *Population Research and Policy Review* **26**(3), 347–369.

- van Raalte, A. A. & Caswell, H. (2013), 'Perturbation analysis of indices of lifespan variability', *Demography* **50**(5), 1615–1640.
- van Raalte, A. A., Martikainen, P. & Myrskylä, M. (2014), 'Lifespan variation by occupational class: Compression or stagnation over time?', *Demography* **51**, 73–95.
- Vaupel, J. W., Zhang, Z. & van Raalte, A. A. (2011), 'Life expectancy and disparity: An international comparison of life table data', *BMJ Open* **1**(1), e000128.
- Wilmoth, J. R. & Horiuchi, S. (1999), 'Rectangularization revisited: Variability of age at death within human populations', *Demography* **36**(4), 475–495.
- Yaari, M. E. (1965), 'Uncertain lifetime, life insurance, and the theory of the consumer', *The Review of Economic Studies* **32**(2), 137–150.
- Zhang, J. T. (2013), *Analysis of Variance for Functional Data*, Chapman and Hall, London.
- Zivot, E. & Wang, J. (2006), *Modeling Financial Time Series with S-PLUS*, Springer, New York.