A Warped Failure Time Model for Human Mortality

Extended abstract submitted to the XXVI IUSSP International Population Conference

Carlo G. Camarda* Paul H.C. Eilers† Jutta Gampe‡

15 September 2008

1 Introduction

Traditionally human mortality is studied by comparing hazard functions. However, lifespan distributions can also be characterized by their probability density. Thus, instead of modelling trends in the hazards, we may study how the age-axis would have to be transformed so that one age-at-death distribution conforms to another. In the simplest case the transformation is linear, leading to an accelerated failure time (AFT) model. However, the assumption that all cohorts postpone death or speed up their lives at a constant rate across the age range is too simplistic for human mortality studies. In this paper, we generalize the AFT approach and present a new model for comparing age-at-death distributions assuming only smoothness for the transformation function.

2 Comparing age-at-death distributions

Figure 1 shows the age-at-death distributions, as derived from period life-tables, for Danish women from age 30 to 110 in 1930 and in 2006. Data are derived from the Human Mortality Database (2008). Instead of pure death counts, we use a period life-table approach for adjusting the effect of birth cohort sizes in the age-at-death distributions. In this way, exposure population is excluded in the approach.

To investigate the changes in mortality that lead to the different patterns we want to transform the age-axis such that the two densities coincide. More specifically, we define one

*Corresponding author: Max Planck Institute for Demographic Research, Konrad-Zuse-Strasse 1, 18057 Rostock, Germany; Tel. +49(0)381 2081 172, Fax. +49(0)381 2081 472; email: camarda@demogr.mpg.de

†Methodology and Statistics, Faculty of Social and Behavioural Sciences, Utrecht University & Data Theory Group, Leiden University, The Netherlands; email: p.h.c.eilers@uu.nl

‡Max Planck Institute for Demographic Research, Rostock, Germany; email: gampe@demogr.mpg.de

distribution as the target, with density $f(y)$, and want to obtain the transformation function $y = w(x)$ so that the density of the other distribution, $g(x)$, conforms to the target density on the warped axis, i.e.,

$$g(x) = f(w(x)) \cdot |w'(x)| \,. \tag{1}$$

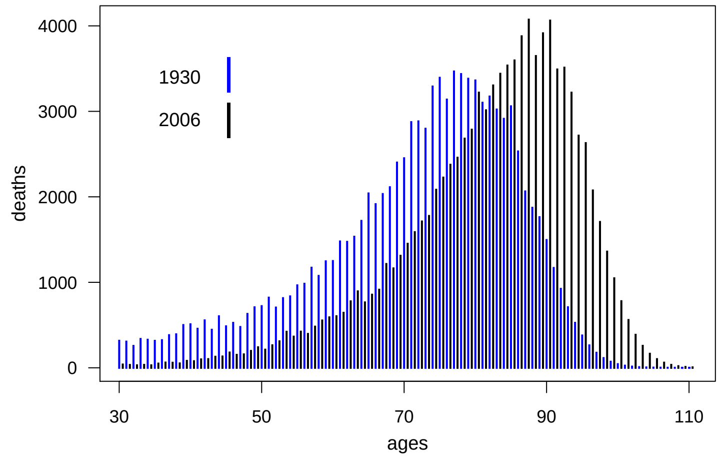

Figure 1: Life-table age-at-death distribution of Danish females for the years 1930 and 2006.

As mentioned, a linear $w(x)$ which shifts and/or stretches the age-axis to conform one distribution to another would be too simplistic for human mortality. Figure 1 clearly shows this issue: for instance, the age-at-death distribution in 1930 is neither a shifted nor a stretched version of the age-at-death distribution in 2006. On the other hand, parametric assumption of the transformation function for the age-axis will impose too much structure on the model and may lead to misleading interpretation.

Following these considerations, we suggest freeing the transformation function $w(x)$ from any rigid shape, assuming only smoothness. This approach leads to a nonlinear transformation of the independent axis which is commonly called "warping". On the other hand, we deal with a transformation of time as in the AFT models. Hence we call our model Warped Failure Time (WaFT) model.

3 The Warped Failure Time Model and its estimation

Our model is not restricted to any particular target distribution, $f(x; \boldsymbol{\theta})$, and it will mostly be estimated from data. In the following we consider the parameters $\boldsymbol{\theta}$ fixed. The observed death counts at age x_i are denoted by y_i and are realizations from Poisson variables with $E(y_i) = \mu_i$. The values μ_i derive from the density $g(x)$ that generated the data as in equation (1). The proposed model is

$$
\begin{aligned}
\mu_i = E(y_i) &= \gamma \cdot f(w(x_i; \boldsymbol{\alpha}), \boldsymbol{\theta}) \cdot \frac{\partial}{\partial x} w(x_i; \boldsymbol{\alpha}) \\
&= \gamma \cdot f(w(x_i; \boldsymbol{\alpha}), \boldsymbol{\theta}) \cdot v(x_i; \boldsymbol{\alpha}) \,, \qquad (2)
\end{aligned}
$$

where γ is a normalizing constant such that $\sum_i y_i = \sum_i \mu_i$ and $f(\cdot)$ is the target density. The warping function $w(x_i; \boldsymbol{\alpha})$ is to be determined such that, after transforming the age-axis, the density matches the specified target.

To allow for arbitrary shape of $w(x)$, we represent the warping function by a linear combination of B-splines and a penalized Poisson likelihood approach is implemented to estimate the coefficients $\boldsymbol{\alpha}$ in equation (2). Concisely, we adapt the iteratively reweighted least squares (IWLS) algorithm and follow a P-spline approach (Eilers and Marx, 1996) for smoothing the warping function $w(\cdot)$. In matrix notation:

$$
(\tilde{\boldsymbol{X}}'\tilde{\boldsymbol{W}}\tilde{\boldsymbol{X}} + \lambda \boldsymbol{P})\boldsymbol{\beta} = \tilde{\boldsymbol{X}}'\tilde{\boldsymbol{W}}(\tilde{\boldsymbol{W}}^{-1}(\boldsymbol{y} - \tilde{\boldsymbol{\mu}}) + \tilde{\boldsymbol{X}}\tilde{\boldsymbol{\beta}})
$$

where $\boldsymbol{P} = \begin{pmatrix} 0 & 0 \\ 0 & \breve{\boldsymbol{P}} \end{pmatrix}$ and $\breve{\boldsymbol{P}} = \boldsymbol{D}_d'\boldsymbol{D}_d$. The matrix $\boldsymbol{D}_d$ calculates d-th order differences. The model matrix is $\tilde{\boldsymbol{X}} = [\mathbf{1}, \tilde{\boldsymbol{Q}}]$, the coefficient vector $\boldsymbol{\beta}' = [\ln(\gamma), \boldsymbol{\alpha}']$ and the matrix $\tilde{\boldsymbol{W}} = \texttt{diag}(\tilde{\boldsymbol{\mu}})$. The matrix $\boldsymbol{Q}$ includes the B-splines basis and its differences. The target function $f(\cdot)$ with its derivatives are also incorporated in $\boldsymbol{Q}$. Via the value of the parameter λ the smoothness the warping function can be controlled. λ is optimized by minimizing the Bayesian Information Criterion. Camarda et al. (2008) presents a detailed description of the model and its penalized Poisson likelihood estimation.

4 Applications

4.1 Gompertz target distribution

For the Danish age-at-death distributions introduced in Section 2, we use the year 2006 as the target density. Dealing with mortality over age 30, a possible choice for representing age-at-death distribution in year 2006 is the Gompertz distribution. The estimated parameters for the Gompertz are $\hat{\boldsymbol{\theta}} = (\hat{a}, \hat{b})' = (1.14e^{-5}, 0.11)'$.

Fixed the target Gompertz distribution for year 2006, the WaFT model has been used to warp the age-axis and fit the Danish age-at-death distribution observed in 1930. Figure 2 (left panel) shows the target distribution with its Gompertz estimates as well as the fitted values from the WaFT model. The BIC profile for this example is presented in the right panel of Figure 2, and the smoothing parameter λ was selected equal to 47.9. Figure 3 shows the resulting transformation function $w(\boldsymbol{x}, \hat{\boldsymbol{\alpha}})$ along with its derivative. The identity transformation is indicated by a dashed line. The warping function is clearly nonlinear, that is, neither a simple shift nor a uniform scaling of the age-axis can map one density on to the other. This feature can be easily acknowledge looking at the derivative of the warping function.

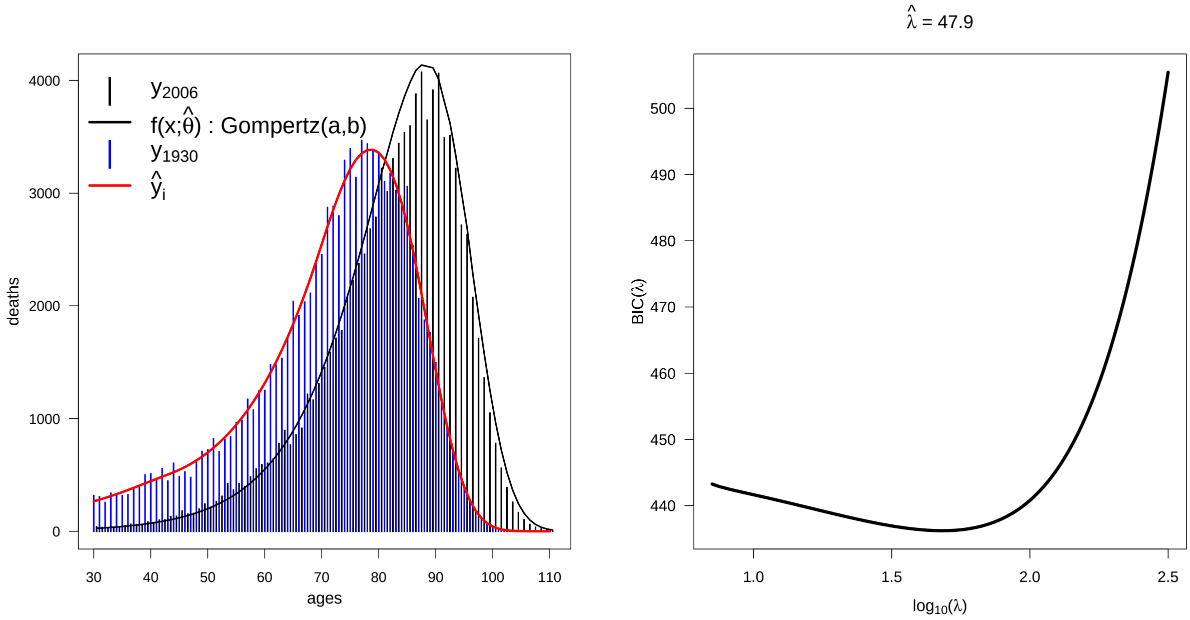

Figure 2: Left panel: Life-table age-at-death distributions for the Danish data over age 30. Data from 2006 are fitted with a Gompertz function and used as target distribution. Data from 1930 are estimated with the WaFT model. Right panel: BIC profile.

4.1.1 Non-parametric target distribution

Often the Gompertz distribution with only two parameters cannot properly describe more complex patterns of adult mortality. Alternatively we can use a non-parametric estimate of the target distribution to improve the model. Again we choose a P–spline approach to obtain the estimated target density (Eilers & Marx, 1996).

The Gompertz distribution plays a prominent role in the study of adult human mortality, but sometimes such parametric distribution cannot properly describe more complex patterns of adult mortality. Instead of searching alternative parametric distributions for portraying the target density, we can free the WaFT from any parametric assumption even regarding the estimation of the target distribution. In particular, we estimate a target distribution using a P-spline approach introduced by Eilers and Marx (1996). Once a target distribution is fitted, the WaFT model can be easily adopted to estimate the warping function.

Figure 4 shows outcomes from a P-spline approach for the Danish women above age 10 over which Gompertz distribution with only two parameters is likely inappropriate.

Figure 4 presents also the fitted values from the WaFT model. Since we do not assume any parametric distribution, the WaFT model actually warps the age-axis such that the Danish age-at-death distribution in 1930 conforms the age-at-death distribution in 2006. Figure 5 shows

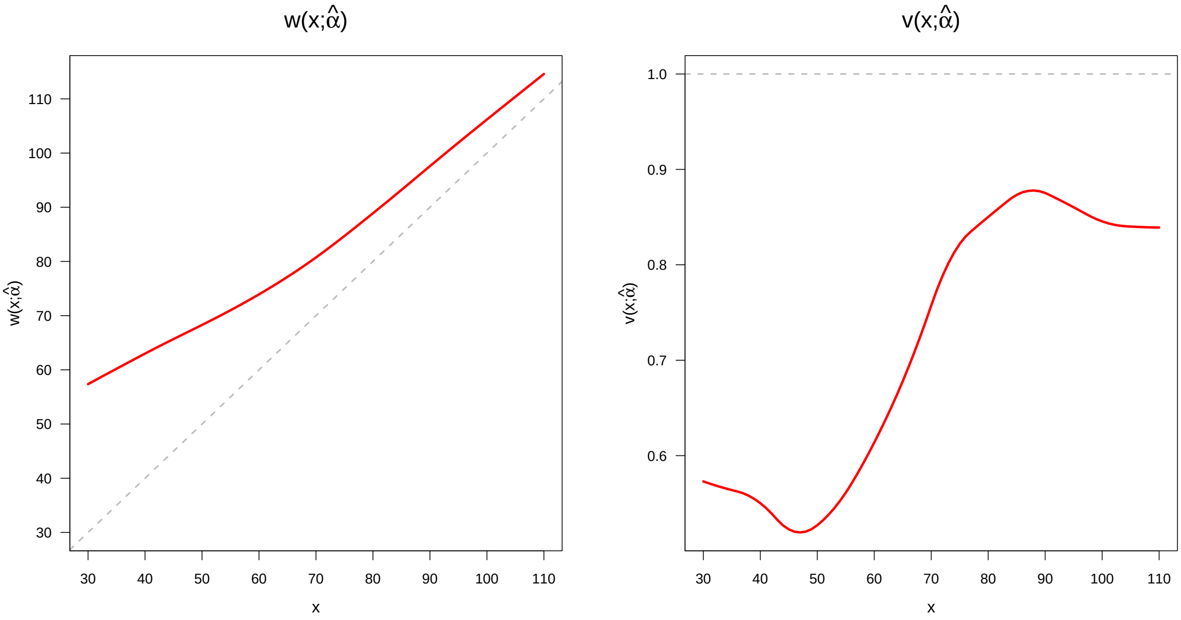

Figure 3: Outcomes from the Danish female population over age 30. Left panel: estimated warping function $w(\boldsymbol{x}, \hat{\boldsymbol{\alpha}})$. The identity transformation is indicated by a dashed grey line. Right panel: estimated derivative of the warping function. The grey dotted lines represents any simple shift transformation of the x-axis.

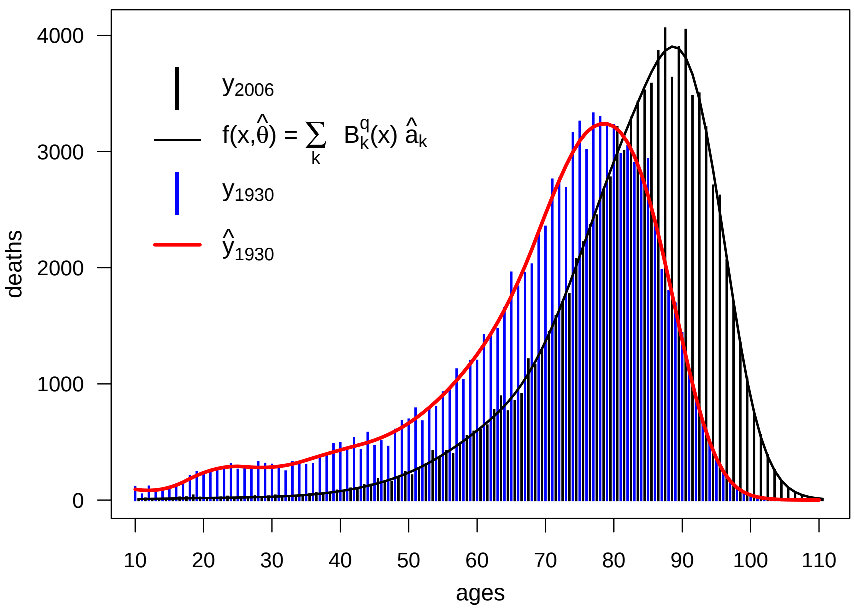

Figure 4: Life-table age-at-death distributions for the Danish data over age 10. Non-parametric P-splines estimate for the target distribution (year 2006). Data from 1930 are estimated with the WaFT model.

both the fitted transformation function and its derivative. Also in this case, the derivative clearly shows that a simple linear warping of the age-axis would not be enough to account for the differences in the age-at-death distributions between these two years.

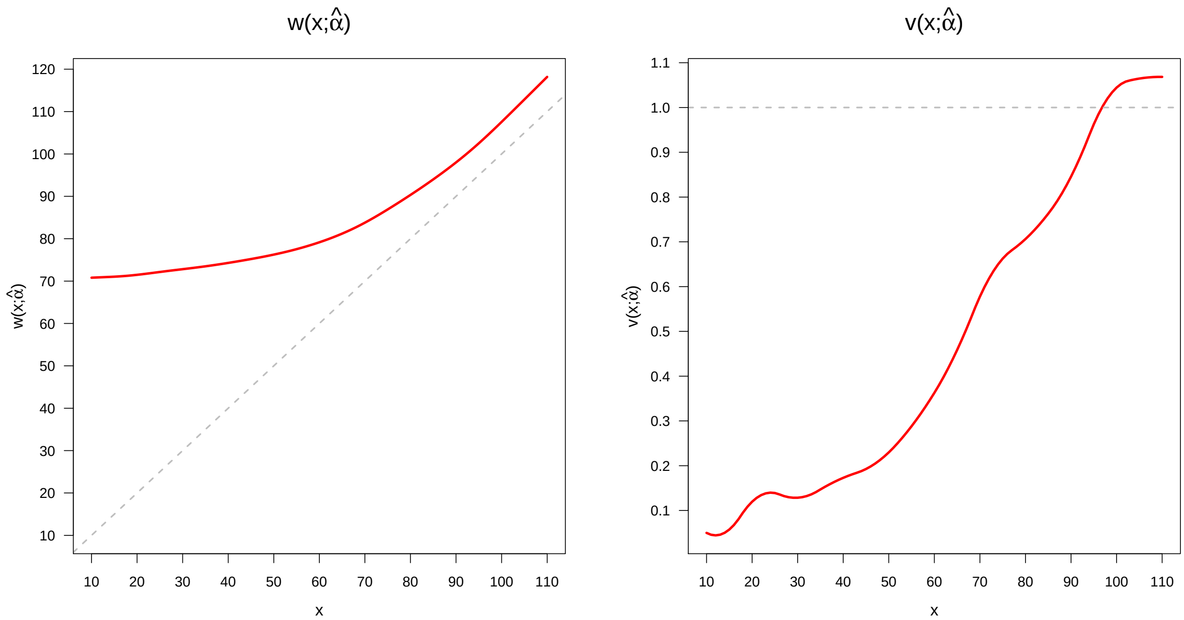

Figure 5: Outcomes from the Danish female population over age 10. Left panel: estimated warping function $w(\boldsymbol{x}, \hat{\boldsymbol{\alpha}})$. The identity transformation is indicated by a dashed grey line. Right panel: estimated derivative of the warping function $v(\boldsymbol{x}, \hat{\boldsymbol{\alpha}})$. The grey dotted lines represent any simple shift transformation of the x-axis.

5 Interpretation and outlook

In this abstract, we briefly present a new approach for dealing with the estimation of a nonlinear transformation to align age-at-death distributions. The proposed WaFT model is a rather general tool and brings together the ideas of warping and smoothing. Starting from a specific target distribution, the model allows estimation of the warping function of the age-axis that can map one distribution onto the other.

The only assumption that is made about the warping function is smoothness. By using a P-spline approach, not only can the warping function be estimated, but we may also directly express its derivative via B-splines. A penalized Poisson likelihood approach is then employed to estimate the model. The target function can be estimated either with parametric or non-parametric approaches which provides to great flexibility.

The fitted warping function can be easily interpreted as the differences in aging between the two groups or the gain (or loss) in longevity. Differently, we can use the warping function as a measure of postponement of dying of a population with respect to the target one. In other words, in Figure 5 (left panel), one can consider that, on average, a 10 years old Danish female in 1930 would have postpone her death to age 70 if she would have experience mortality conditions of 2006.

Furthermore, the derivative of the warping function describe the rate of change of the mentioned process and we clearly see that different ages postpone death or speed up their lives at different rate across the age range. Explicitly, since x represents age, the first derivative of of the warping function can be interpreted as speed of aging at each age x. Moreover, the non-linearity of the warping function shows how the WaFT model extend the common accelerated failure model which is too simplistic for capturing mortality dynamics.

The WaFT model applied on actual populations shows that the fitted warping function gives reasonable and interesting outcomes. Comparisons between countries and sexes are also possible as well as applications on non-human data. The WaFT model is also appropriate for comparison of any two densities. We therefore envision alternative applications of the WaFT model in which nonlinear transformation of the x-axis is a suitable and reasonable idea.

Furthermore, in case of mortality data, we plan a generalization of the WaFT can account in a two-dimensional setting. Warping functions between two subsequent years are expected to change smoothly. Therefore, one can cope with a sequence of warping functions over time by an additional penalty that controls the temporal pattern in the age-axis transformation. Two-dimensional smoothing methodology such as presented by Currie et al. (2006) can be used to generalize the WaFT model to two dimensions, as well.

References

Camarda, C. G., P. H. C. Eilers, and J. Gampe (2008). A Warped Failure Time Model for Human Mortality. In P. H. C. Eilers (Ed.), *Proceedings of the 23rd International Workshop of Statistical Modelling*.

Currie, I. D., M. Durban, and P. H. C. Eilers (2006). Generalized linear array models with applications to multidimentional smoothing. *Journal of the Royal Statistical Society 68*, 259–280.

Eilers, P. H. C. and B. D. Marx (1996). Flexible Smoothing with B-splines and Penalties (with discussion). *Statistical Science 11*(2), 89–102.

Human Mortality Database (2008). University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at `www.mortality.org` (Data downloaded on September 2008).