Some comments on "A Hermite spline approach for modelling population mortality" by Tang, Li & Tickle (2022)

Richards, Stephen J.*

April 17, 2023

Abstract

Tang et al. [2022] propose a new class of models for stochastic mortality modelling using Hermite splines. There are four useful features of this class that are worth emphasising. First, for single-sex datasets this new class of projection models can be fitted as a Generalized Linear Model (GLM). Second, these models can automatically extrapolate mortality rates to ages above the maximum age of the data set. Third, simpler sub-variants of the models exist for forecasting when one of the variables lacks a clear drift. Finally, a minor reparameterisation increases the quality of long-range forecasts of period mortality.

Regarding the barrier-fitting algorithm in Tang et al. [2022], it is worth noting that for a single-sex data set the penalty is not required and the model in equation (8) is then just a Generalized Linear Model (GLM) [McCullagh and Nelder, 1989]. Furthermore, the models specified in equation (8) require no identifiability constraints. This can be proved by showing that the rank of the model matrix, $\boldsymbol{X}$, equals the length of the parameter vector, $\boldsymbol{B}$ [Currie, 2020]. This is a feature shared only with the stochastic mortality model of Cairns et al. [2006], which is itself also a GLM. The models of Tang et al. [2022] are therefore easily implemented for single-sex datasets.

Like the Gompertz model of Cairns et al. [2006], the Hermite-spline models of Tang et al. [2022] can extrapolate mortality rates beyond the upper age of the available data. This is a particularly useful feature for actuarial calculations involving annuities and pensions. For example, in Figure 1 the calibrating data stop at age 105, but extrapolation to higher ages was achieved simply by setting $x_1 = 120$. In the case of females for England & Wales, Table 1 shows that using $x_1 = 120$ also markedly reduces the AIC [Akaike, 1987] compared to using $x_1 = 105$. The Gompertz model of Cairns et al. [2006] extrapolates an ever-increasing mortality hazard with age, as advocated by Gavrilov and Gavrilova [2015]. In contrast, the Hermite-spline models of Richards [2020] and Tang et al. [2022] extrapolate to a mortality plateau, as advocated by Gampe [2010]. In Figure 1 the limiting mortality hazard is around 1.089, corresponding to a limiting annual mortality rate of 66%. This compares with annual mortality rates of 61-63% at age 119 in the mortality tables used by UK actuaries for pension and annuity calculations (CMI, 2020). In contrast, Gampe [2010] found a limiting annual mortality rate of 50%, which is the rate assumed from age 112 by US actuaries for similar calculations (PBGC, 2023).

For mortality projections it is necessary to have a clear time signal in the parameters. However, not every parameter vector for every data set will exhibit this. An example is shown in Figure 2 for females in England & Wales — there is a clear time signal for $\{\hat{\alpha}_t\}$, and a relatively clear signal for $\{\hat{s}_{0,t}\}$ after the mid-1990s, but not for $\{\hat{\omega}_t\}$. The estimated drift term, $\hat{\mu}$, for the $\{\hat{\omega}_t\}$ process is 0.002

*stephen@longevitas.co.uk, Heriot-Watt University, Edinburgh, EH14 4AS, United Kingdom. http://orcid.org/0000-0001-6859-6347

Figure 1: Observed, fitted and extrapolated mortality rates in 2019 for females in England & Wales, ages 50–105. Source: own calculations for HS2 GLM of Tang et al. [2022] using data at ages 50–105, 1971-2019.

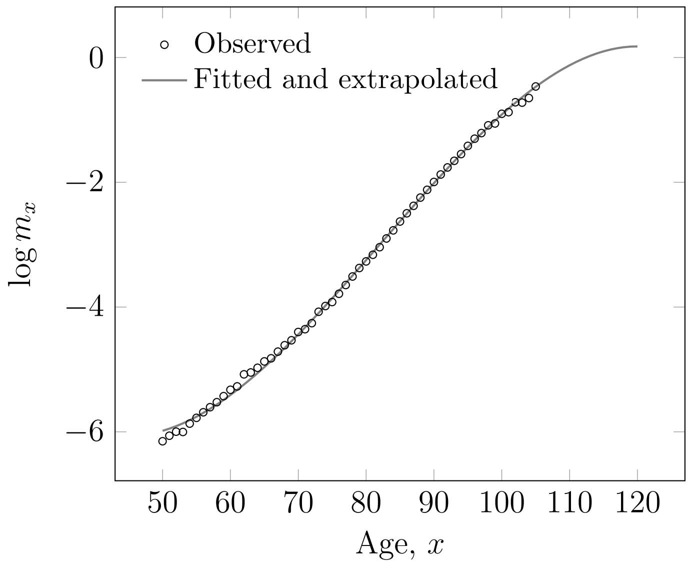
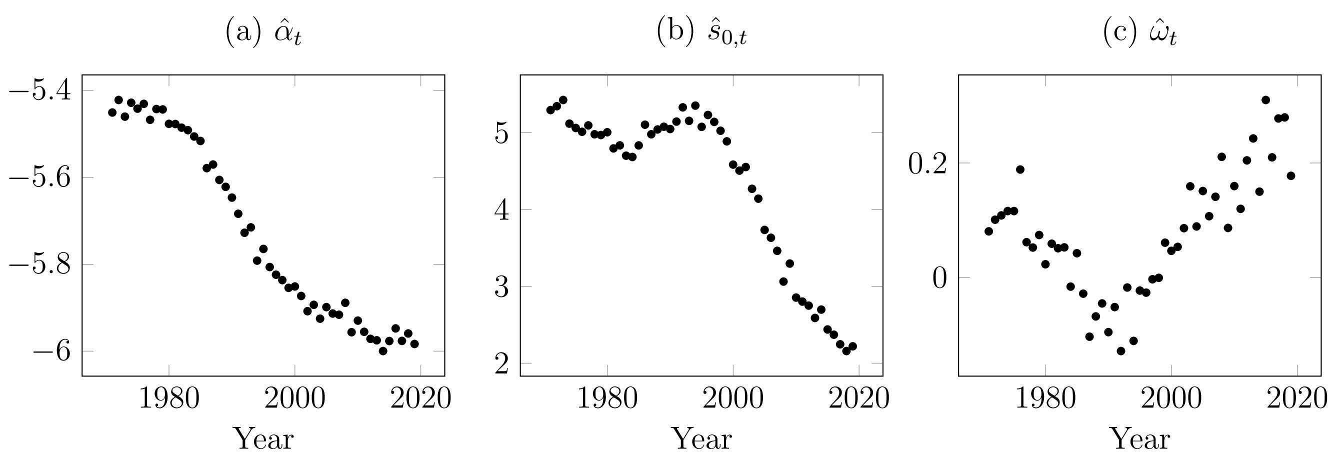


Figure 2: Parameters for HS2 GLM behind Figure 1.



with a standard error of 0.0095, suggesting that $\{\hat{\omega}_t\}$ is merely a random walk without drift. For a long-term forecast it therefore makes sense to adopt a simplifying assumption of $\omega_t = \omega$ as follows:

$$\log \boldsymbol{M} = \boldsymbol{X}\boldsymbol{B} = [\boldsymbol{I}_{\boldsymbol{n_y}} \otimes \boldsymbol{h_{00}} : \boldsymbol{I}_{\boldsymbol{n_y}} \otimes : \boldsymbol{h_{10}} : \boldsymbol{h_{01}}]\boldsymbol{B} \quad (1)$$

where $\boldsymbol{h_{00}}$, $\boldsymbol{h_{10}}$ and $\boldsymbol{h_{01}}$ denote the column vectors of Hermite splines in Tang et al. [2022, equation (3)]. For a single-sex data set equation (1) is also a GLM that requires no identifiability constraints. Like the corresponding CBD model, the mortality rates in equation (1) would also be forecast with a bivariate random walk with drift for $(\alpha_t, s_{0,t})$. The long-term projection of mortality rates under equation (1) will be simpler and more stable than the trivariate HS2 model, albeit at the cost of a poorer fit to the data, as shown in Table 1.

Table 1: AICs for various GLMs fitted to data for females in England & Wales, ages 50–105, 1971-2019.

		AIC:	
Model	Parameters	$x_1 = 105$	$x_1 = 120$
Trivariate HS2	$3n_y = 147$	89,841.8	42,020.4
Bivariate HS2 with constant ω	$2n_y + 1 = 99$	96,355.8	48,010.8
Bivariate Gompertz (CBD)	$2n_y = 98$	73,562.5	73,562.5

However, closeness of fit to data is not the sole criterion (or even necessarily the best one) when choosing a forecasting model. The quality of the forecast [Cairns et al., 2009] is also a consideration, and the forecast values of $s_{0,t}$ in Figure 2 will eventually turn negative, thus causing projected period mortality rates at young ages to reduce with increasing age, as shown in Figure 3. This minor defect in the forecast can be corrected by replacing the $s_{0,t}$ multiplier of h_{10} with $e^{s^*_{0,t}}$. This adjusted model is not a GLM, but as long as $\hat{s}_{0,t} > 0$ in equation (1), then we can derive $\hat{s}^*_{0,t} = \log \hat{s}_{0,t}$. The other parameters and the model fit overall are unchanged, but the bivariate random walk with drift applied to $(\alpha_t, s^*_{0,t})$ leads to non-decreasing mortality rates at all periods in the forecast, as shown in Figure 3. Thus, forecast quality can be improved at no change to the fit as long as $\hat{s}_{0,t} > 0$.

The choice of which HS2 parameterisation to use — the trivariate $(\alpha_t, s_{0,t}, \omega_t)$ model of Tang et al. [2022] or the bivariate model $(\alpha_t, s^*_{0,t})$ based on equation (1) — will depend on the application. For a long-term forecast of period mortality, one would probably use $(\alpha_t, s^*_{0,t})$. However, with short-term value-at-risk calculations for the likes of Solvency II [Richards et al., 2020] it would be important to fully express the short-term variability in ω_t, and so one would probably use the trivariate parameterisation of Tang et al. [2022] for sample paths.

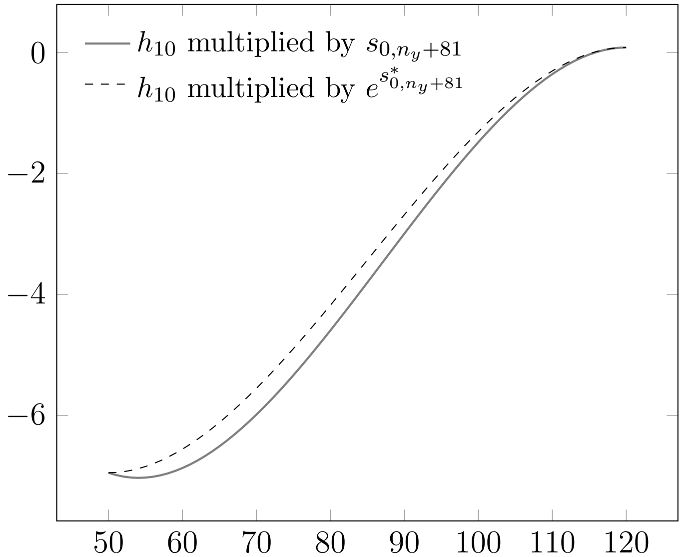
References

H. Akaike. Factor analysis and AIC. *Psychometrica*, 52:317–333, 1987. ISSN 0033–3123. doi: 10.1007/BF02294359.

A. J. G. Cairns, D. Blake, and K. Dowd. A two-factor model for stochastic mortality with parameter uncertainty: theory and calibration. *Journal of Risk and Insurance*, 73:687–718, 2006. doi: 10.1111/j.1539-6975.2006.00195.x.

A. J. G. Cairns, D. Blake, K. Dowd, G. D. Coughlan, D. Epstein, A. Ong, and I. Balevich. A quantitative comparison of stochastic mortality models using data from England and Wales and the United States. *North American Actuarial Journal*, 13(1):1–35, 2009. doi: 10.1080/10920277.2009.10597538.

Figure 3: Forecast mortality rates in 2100 (i.e. time n_y+81) for females in England & Wales using alternative multipliers for the h_{10} Hermite spline. Source: own calculations using data at ages 50–105, 1971-2019.



I. D. Currie. Constraints, the identifiability problem and the forecasting of mortality. *Annals of Actuarial Science*, 14(2):537–566, 2020. doi: 10.1017/S1748499520000020.

J. Gampe. Human mortality beyond age 110. *Supercentenarians. Demographic Research Monographs*, 2010. doi: 10.1007/978-3-642-11520-2_13.

L. A. Gavrilov and N. S. Gavrilova. Biodemography of old-age mortality in humans and rodents. *Journal of Gerontology: Biological Sciences*, 70(1):1–9, 2015.

P. McCullagh and J. A. Nelder. *Generalized Linear Models*, volume 37 of *Monographs on Statistics and Applied Probability*. Chapman and Hall, London, second edition, 1989. ISBN 0-412-31760-5.

S. J. Richards. A Hermite-spline model of post-retirement mortality. *Scandinavian Actuarial Journal*, 2020:2:110–127, 2020. doi: 10.1080/03461238.2019.1642239.

S. J. Richards, I. D. Currie, T. Kleinow, and G. P. Ritchie. Longevity trend risk over limited time horizons. *Annals of Actuarial Science*, 14(2):262–277, 2020. doi: 10.1017/S174849952000007X.

S. Tang, J. Li, and L. Tickle. A Hermite spline approach for modelling population mortality. *Annals of Actuarial Science*, pages 1–42, 2022. doi: 10.1017/S1748499522000173.

Continuous Mortality Investigation Limited (CMI). Final PAIP 16 Series mortality rates v01 2020-07-10. Institute and Faculty of Actuaries, London. https://www.actuaries.org.uk/documents/final-paip-16-series-mortality-rates-v01-2020-07-10; accessed on 3rd April 2023.

Pension Benefit Guaranty Corporation (PBGC). ERISA Section 4044 Mortality Table for 2023 Valuation Dates. https://www.pbgc.gov/prac/mortality-retirement-and-pv-max-guarantee/erisa-mortality-tables/erisa-section-4044-mortality-table-for-2023-valuation-dates; accessed on 3rd April 2023.