A Universe of Individual-Level Mortality Models

Xiaochuan Lu, Katja Hanewald and Andrés M. Villegas

UNSW Sydney, School of Risk and Actuarial Studies

Longevity 20 Conference

Introduction

- **Context & Motivation**
  - Increasing importance of individual-level mortality modelling for insurers, pension funds and longevity researchers
  - Need for methods that integrate age and covariate effects
- **Research Gap**
  - Ng (2023) provided a taxonomy of mortality models based on characteristics such as parametric nature
  - Separated statistical from machine learning models, overlooking potential combinations
- **Objective**
  - Present a modular framework and introduce novel combined models (e.g., survival tree partitioning, mob tree partitioning)

Individual-Level Mortality Modelling

- **Data Structure:** Survival data with age at entry, age at exit, event indicator, and covariates (mortality rate is unobservable)

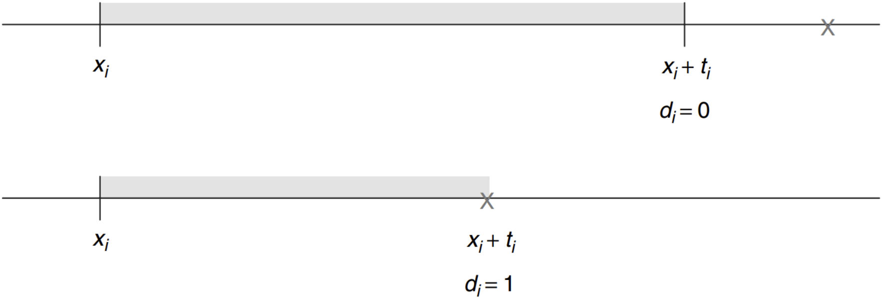
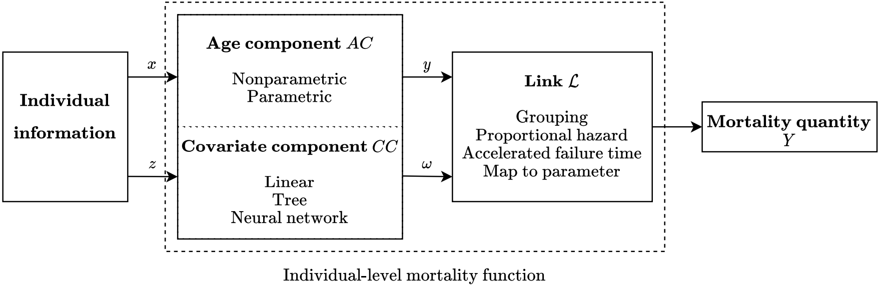


Figure 1: Diagram of survival model setup (Richards, 2012)

Universe of Individual-level Mortality Models



Estimation of AC and CC parameters: simultaneous or sequential Age can be an input for CC, allowing for interaction

Age Component

- **Purpose:** Define the distributional form of age effect
- **Two Approaches:**
  - **Non-Parametric Age Component:**
    - Models y empirically with a step function
    - E.g. Kaplan-Meier / Nelson-Aalen estimators
  - **Parametric Age Component:**
    - Assumes age at death follows a specific distribution
    - E.g. Gompertz, Makeham, Perks, Beard, Kannisto, Thatcher laws
    - Enables interpolation/extrapolation and likelihood-based estimation

Covariate Component

- **Role:** Captures how non-age information (e.g., demographics, health measures) affects mortality
- **Modelling Options:**
  - **Linear:** $\omega = \mathcal{B}z$.
  - **Tree-Based:** Non-linear mapping via binary splits; provides grouping of individuals
  - **Neural Network:** More flexible, continuous mapping suited for high-dimensional data

Link Function

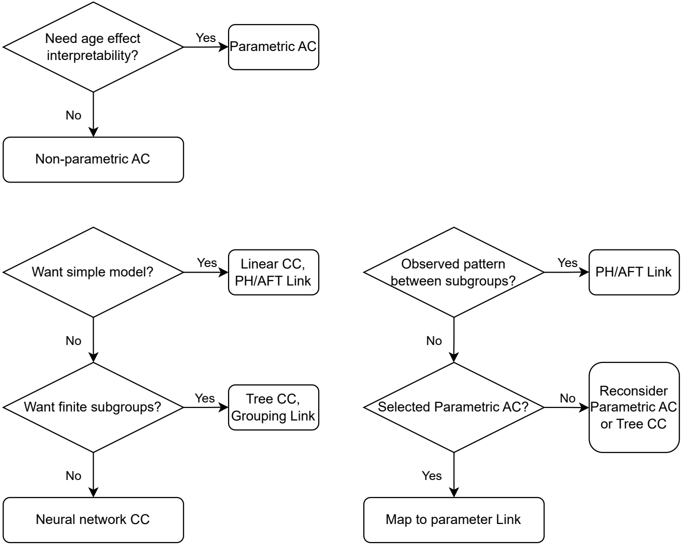
- **Function of $\mathcal{L}$:** Integrate the outputs of age and covariate components to yield the overall mortality output Y
- **Different Forms:**
  - **Grouping Link:** $Y = \sum_{i=1}^{k} y_i \, 1\{\omega = i\}$
  - **Proportional Hazards (PH) Link:** $Y = \mu = \mu_0 \exp(\omega)$
  - **Accelerated Failure Time (AFT) Link:**

$$X \stackrel{d}{=} \exp(\omega) X_0$$

$$Y = \mu(x) = \exp(\omega)\mu_0(\exp(\omega)x), \omega \in \mathbb{R}$$

  - **Map-to-Parameter Link:** $Y = y|(\theta = \omega), \omega \in \Theta$
    - Example: $AC =$ Gompertz$(\alpha, \beta), CC : z \to \theta = (\alpha, \beta)$.

Component Selection Guide



Case Study: CLHLS Dataset

- Chinese Longitudinal Healthy Longevity Survey (CLHLS)
- 7 waves spanning 1998–2018
- Includes data on oldest-old (aged 80+) with over 59,000 observations
- Rich covariate information (demographics, health, cognitive abilities, lifestyle)
- Left-truncated and right-censored survival data

Combined Models

- **Survival Tree:** $AC =$ non-parametric, $CC =$ tree-based, $L =$ grouping
- **Survival Tree Partitioning (STP):** Estimate survival tree and then apply a parametric model to each terminal node
- **Mob Tree Partitioning (MTP):** Estimate AC and CC simultaneously to maximize the likelihood
- **Proportional Hazards (PH) Model:** Extends the Cox model with parametric age component

Evaluation Metrics

Out of sample metrics:

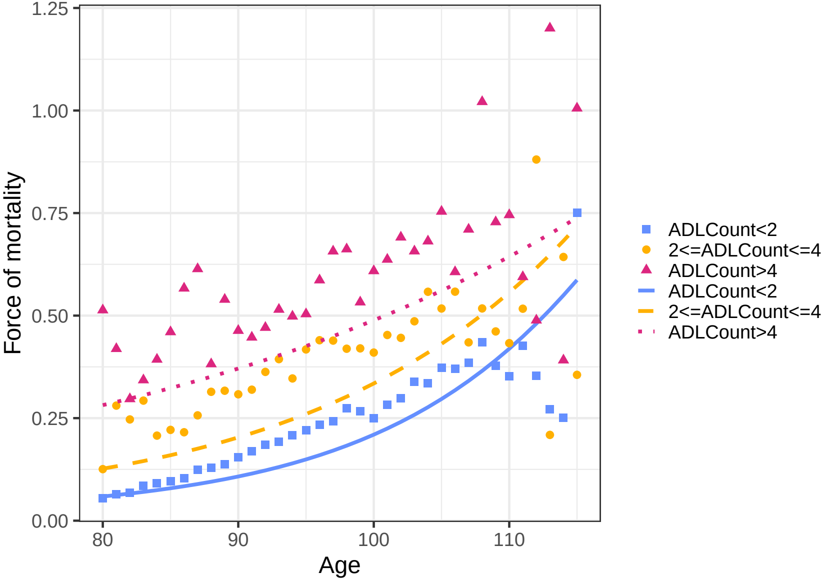
- Integrated Brier Score (IBS):

$$IBS = \int \left(\hat{S}(t|x) - S(t|x)\right)^2 dt$$

- log Score: log likelihood evaluated on test/validation set

$$\mathcal{S}_L(S(x|z), (L, C, \delta)) = \sum_i [\delta_i \log \mu(C_i) - H(C_i) + H(L_i)]$$

Partial Dependence Plot



Application to Life Expectancy

- **Using the Compound Gompertz Model:**
  - Life expectancy is derived by integrating the force of mortality
  - Example equation using the exponential integral:

$$e_G(x) = \frac{1}{b} \exp\left(\frac{a}{b} e^{bx}\right) E_1\left(\frac{a}{b} e^{bx}\right)$$

- **Findings:**
  - Life expectancy differs substantially based on covariate strata (e.g., ADL Count, Interviewer Rated Health)
  - E.g., at age 80, overall life expectancy is around 8 years, but varies significantly among groups

Life Expectancy by ADL Count

Table 1: Life expectancy in years for each cohort at age 80, 90, 100, 110.

	80	**90**	**100**	**110**
Age-only model (Gompertz)	8.0 (5.7)	4.7 (3.8)	2.7 (2.3)	1.4 (1.3)
ADLCount<2	10.0 (6.9)	6.3 (4.8)	3.7 (3.1)	2.1 (1.9)
2<=ADLCount<=4	6.1 (4.9)	4.1 (3.5)	2.6 (2.4)	1.7 (1.6)
ADLCount>4	3.3 (3.0)	2.5 (2.4)	1.9 (1.8)	1.5 (1.4)
Interviewer-Rated Health=1	11.0 (7.2)	6.8 (5.0)	3.9 (3.2)	2.1 (1.9)
Interviewer-Rated Health=2	9.7 (6.7)	6.0 (4.6)	3.6 (3.0)	2.0 (1.8)
Interviewer-Rated Health>2	6.0 (4.8)	4.0 (3.4)	2.6 (2.3)	1.6 (1.5)
Interviewer-Rated Health missing	9.9 (6.8)	6.2 (4.7)	3.6 (3.0)	2.0 (1.8)

Note:
Life expectancy is rounded to 1 decimal places (1 day is 0.0027).

Conclusion

- **Summary of Contributions:**
  - Novel framework for individual mortality modelling
  - Unified age and covariate components
  - New combined models with strong performance
- **Insights from CLHLS:**
  - Key predictors: ADL Count, Interviewer Rated Health
  - MTP and PH models show top accuracy
- **Practical Implications:**
  - Distributional mortality predictions for insurers/pensions
  - Clearer covariate effects on longevity
  - Mitigate adverse selection
- **Future Directions:**
  - Deep learning integration
  - Competing risks, multi-state models

Thank you!

Contact: xiaochuan.lu@unsw.edu.au

Link to paper:
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5256615

References I

Aalen, O. (1978). Nonparametric Inference for a Family of Counting Processes. *The Annals of Statistics*, *6*, 701–726.

Beard, R. E. (1971). Some aspects of theories of mortality, cause of death analysis, forecasting and stochastic processes. *Biological Aspects of Demography*, *10*, 57–68.

Cox, D. R. (1972). Regression Models and Life-Tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, *34*, 187–220.

Fu, W., & Simonoff, J. S. (2017). Survival trees for left-truncated and right-censored data, with application to time-varying covariate data. *Biostatistics*, *18*, 352–369.

Gompertz, B. (1825). On the Nature of the Function Expressive of the Law of Human Mortality, and on a New Mode of Determining the Value of Life Contingencies. *Philosophical Transactions of the Royal Society of London*, *115*, 513–583.

Kannisto, V., Lauritsen, J., Thatcher, A. R., & Vaupel, J. W. (1994). Reductions in Mortality at Advanced Ages: Several Decades of Evidence from 27 Countries. *Population and Development Review*, *20*, 793–810.

Kaplan, E. L., & Meier, P. (1958). Nonparametric Estimation from Incomplete Observations. *Journal of the American Statistical Association*, *53*, 457–481.

References II

Makeham, W. M. (1859). On the Law of Mortality and the Construction of Annuity Tables. *The Assurance Magazine, and Journal of the Institute of Actuaries*, *8*, 301–310.

Nelson, W. (1972). Theory and Applications of Hazard Plotting for Censored Failure Data. *Technometrics*, *14*, 945–966.

Ng, J. (2023). Multivariable Mortality Modeling, Survival Analysis, and Machine Learning. *Longevity Bulletin*, 5–11.

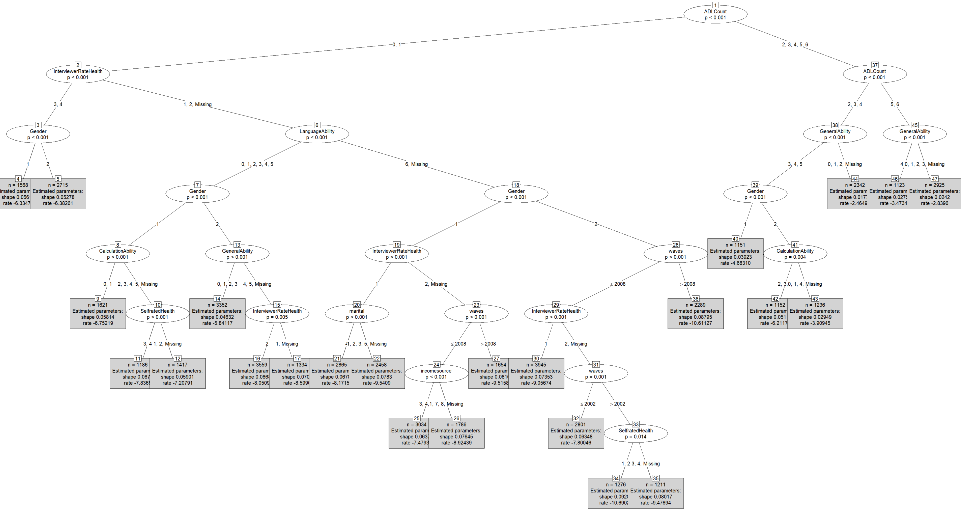
Perks, W. (1932). On Some Experiments in the Graduation of Mortality Statistics. *Journal of the Institute of Actuaries (1886-1994)*, *63*, 12–57.

Richards, S. J. (2012). A handbook of parametric survival models for actuarial use. *Scandinavian Actuarial Journal*, *2012*, 233–257.

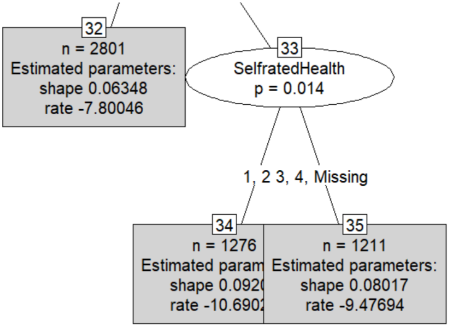
Thatcher, A. R. (1999). The long-term pattern of adult mortality and the highest attained age. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *162*, 5–43.

Zeileis, A., & Hothorn, T. (2015). *Parties , Models , Mobsters : A New Implementation of Model-Based Recursive Partitioning in R*. Retrieved from https://cran.r-project.org/web/packages/partykit/vignettes/mob.pdf

Mob Tree Partitioning results



Mob Tree Partitioning results



Model Comparison (IBS)

Table 2: Test IBS for each model in the 8 samples.

	Gompertz	Beard	STP_Gompertz	STP_Beard	MTP	GPH
1998-2000	0.033	0.033	0.034	0.035	0.034	0.034
2000-2002	0.03	0.03	0.031	0.031	0.031	0.031
2002-2005	0.037	0.036	0.037	0.037	0.038	0.038
2005-2008	0.037	0.036	0.038	0.037	0.038	0.038
2008-2011	0.041	0.04	0.041	0.04	0.041	0.041
2011-2014	0.041	0.038	0.038	0.036	0.038	0.04
2014-2018	0.051	0.049	0.052	0.05	0.052	0.054
1998-2018	0.039	0.038	0.039	0.038	0.039	0.04

Model Comparison (log Score)

Table 3: Test log score for each model in the 8 samples.

	Gompertz	Beard	STP_Gompertz	STP_Beard	MTP	GPH
1998-2000	-1562.7	-1562.6	-1513.1	-1513.9	-1517.4	-1502.7
2000-2002	-1741.7	-1740	-1681.5	-1680.9	-1687.6	-1669.5
2002-2005	-2912.1	-2908.8	-2866.9	-2868.9	-2877.3	-2855.1
2005-2008	-2250.2	-2243	-2207.1	-2199.9	-2210	-2200.9
2008-2011	-2701.2	-2693.1	-2615.5	-2613.7	-2615.8	-2610.9
2011-2014	-1752.5	-1736.5	-1636	-1631	-1637.1	-1643.6
2014-2018	-1692.5	-1677.8	-1626.2	-1612.7	-1624.4	-1619.5
1998-2018	-15235	-15173.3	-14773.3	-14747.9	-14719.4	-14734.2