# A Comprehensive Framework for Mortality Forecasting

Jon Forster
Mathematical Sciences & ESRC Centre for Population Change
University of Southampton, UK

Joint Eurostat/UNECE Work Session on Demographic Projections
Geneva, 18-20 April 2016

Joint work with Erengul Dodd, Jakub Bijak and Peter Smith

# Mortality projection models

Models for central mortality rates $m_{xt}$ over age $x$ and time $t$ include:

- Generalised bilinear (e.g. Lee Carter with cohort)

$$\log m_{xt} = \alpha_x + \beta_x \kappa_t + \gamma_{t-x}$$

- Generalised linear (e.g. APC with age-period interaction)

$$\log m_{xt} = \alpha_x + t\beta_x + \kappa_t + \gamma_{t-x}$$

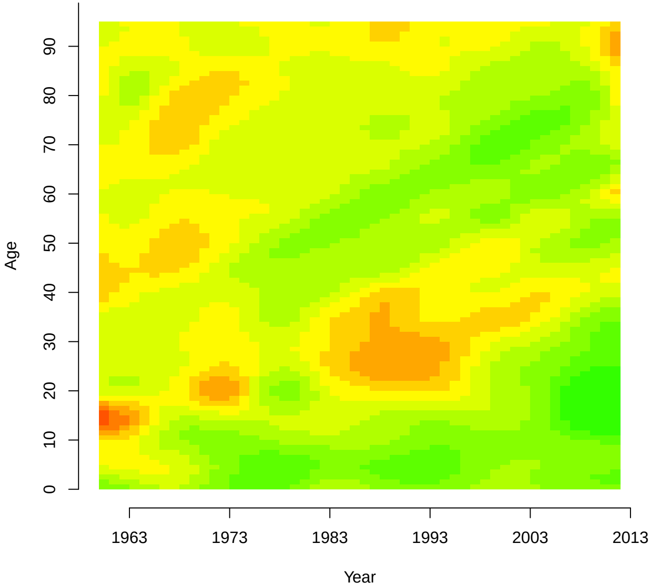
- semi-parametric

$$\log m_{xt} = s(x, t)$$

- generalised additive (GAM)

$$\log m_{xt} = s_\alpha(x) + t\, s_\beta(x) + \kappa_t + s_\gamma(t-x)$$

# Smoothed UK mortality



# The initial proposed model specification

Age-period-cohort (APC) model for mortality improvements

$$\log \frac{m_{xt}}{m_{x\,t-1}} = \alpha_x + \kappa_t + \gamma_{t-x} \tag{1}$$

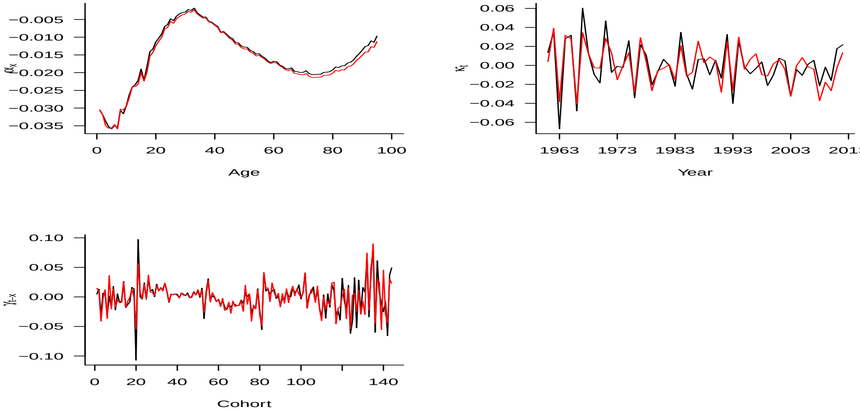
or equivalently APC model for mortality rates, with age-period interaction

$$\log m_{xt} = m_{x0} + \alpha_x t + \kappa_t + \gamma_{t-x} \tag{2}$$

Estimation of the model parameters can be based on the negative binomial log-likelihood: which accounts for over dispersion (relative to Poisson).

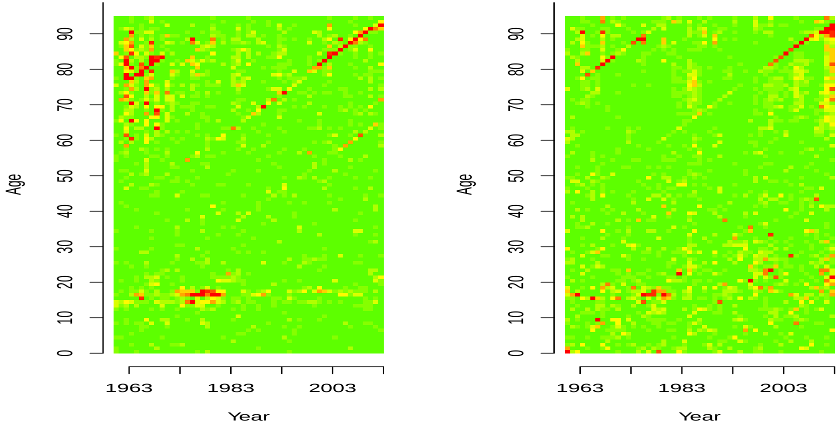
# Parameter estimates





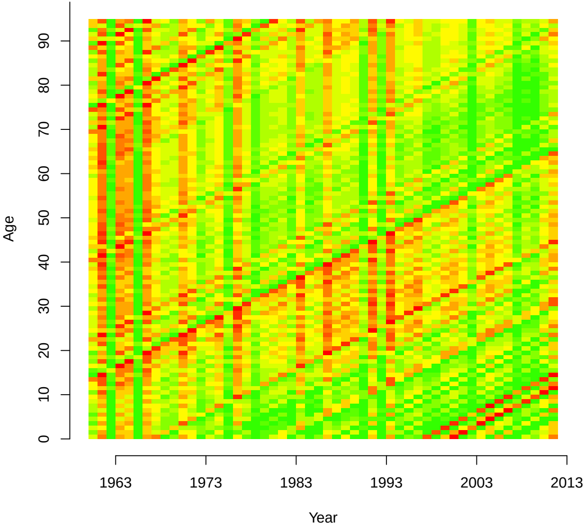

Maximum likelihood estimates of the parameters of model (1) under the Poisson model (black line) and the negative binomial model (red line), data for males 1961-2013.

# Comparison of residuals




The P-spline approach allowing for overdispersion (left panel) and model (1) under the negative binomial distribution (right panel).

# Fitted mortality improvements for model(1)



# Smooth estimates using GAM

UNIVERSITY OF Southampton

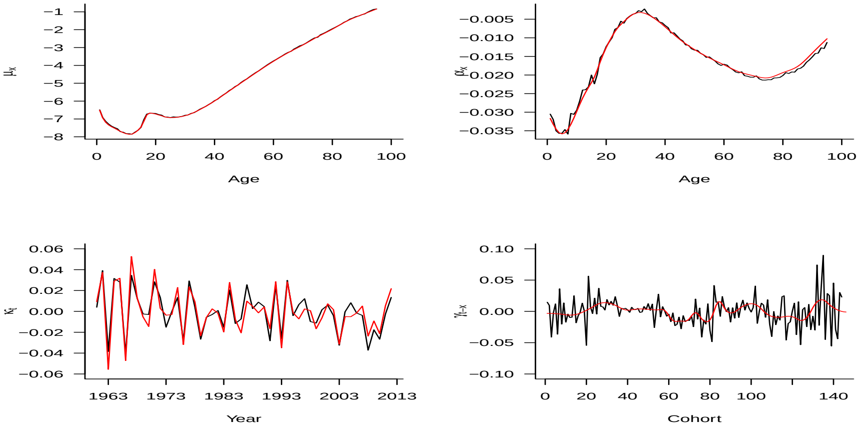
To obtain smoother estimates modify (2) to a generalised additive model (GAM):

$$\log m_{xt} = s_\mu(x) + s_\alpha(x)t + \kappa_t + s_\gamma(t - x). \qquad (3)$$

where $s_\mu$, $s_\alpha$ and $s_\gamma$ are arbitrary smooth functions.

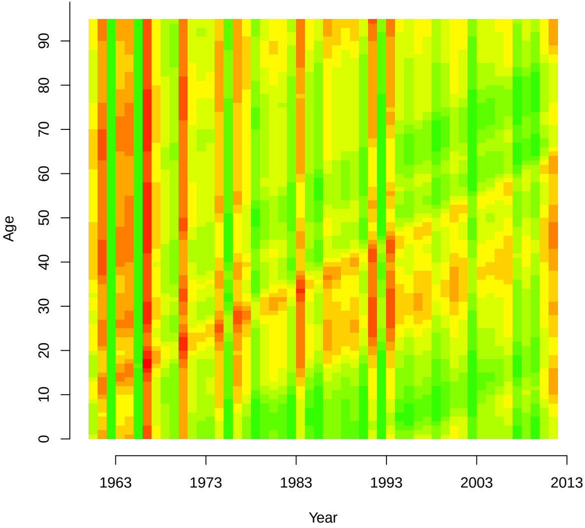
This is a *generalised additive model* which can be robustly fitted in standard statistical software (Wood, 2006).

# Parameter estimates using GAM




Estimates of the parameters of model (3), data for males 1961-2013, (red lines) superimposed over the corresponding estimates for model (2) (black lines).

# Fitted mortality improvements for model(3)



# Model specification for old ages

For the highest ages $x$, use parametric models:

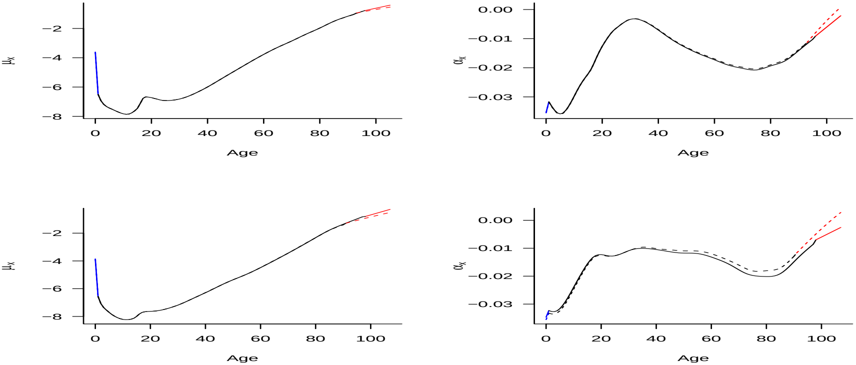
$$\log m_{xt} = \mu_0 + \mu_1 x + (\alpha_0 + \alpha_1 x)t + \kappa_t + s_\gamma(t - x) \qquad x > x_0 \qquad (4)$$

or

$$\log \frac{m_{xt}}{\beta - m_{xt}} = \mu_0 + \mu_1 x + (\alpha_0 + \alpha_1 x)t + \kappa_t + s_\gamma(t - x) \qquad x > x_0 \quad (5)$$

where $\kappa_t$, $s_\gamma(t - x)$ are estimates obtained from fitting (3) to the main body of the data $(0 < x \leq x_0)$.

# Parameter estimates




Estimates of the parameters of models (3), (4) and (5), 1961-2013, for males (upper panels; $x_0$=95 for log-linear model (solid line) and $x_0$=92 for logistic model (dashed line)) and females (lower panels; $x_0$=97 for log-linear model (solid line) and $x_0$=90 for logistic model (dashed line)).

# Model specification for infants

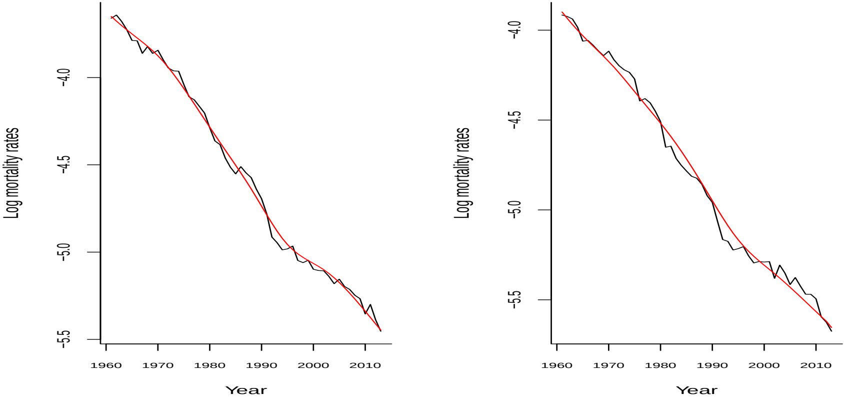
For infants (age 0) we use:

$$\log \mu_{0t} = \mu_0 + \alpha_0 t + s_\gamma(t - x) \tag{6}$$

where $s_\gamma(t - x)$ are estimates obtained from fitting model (3) to the main body of the data $(0 < x \leq x_0)$.

# Estimates of infant mortality rates

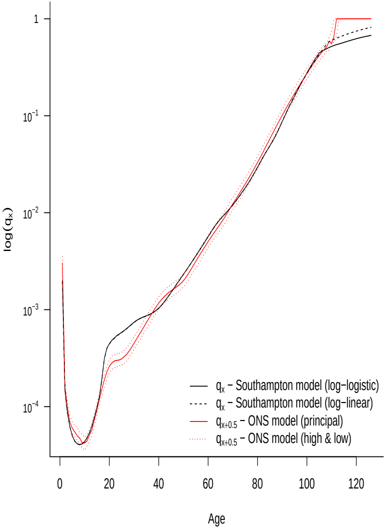
UNIVERSITY OF Southampton



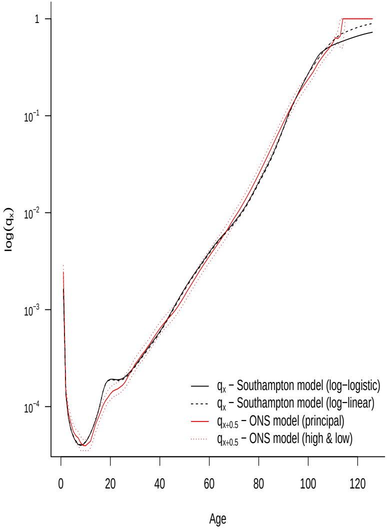

Estimates of infant mortality rates, 1961-2013, for males (left panel) and females (right panel) using model (6; red lines), compared with observed rates (black lines).

# 2035 Projections Males and Females





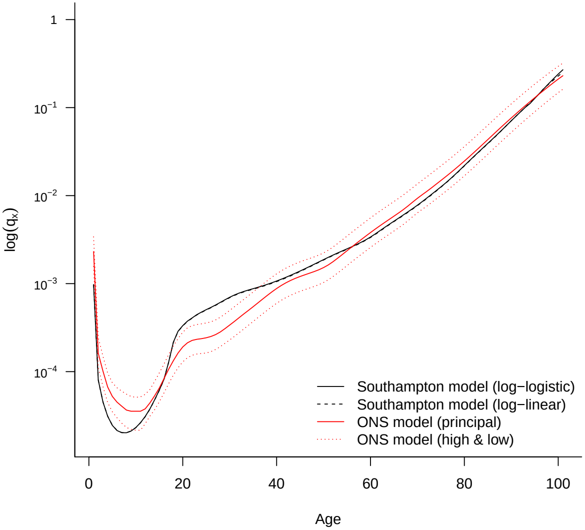

(a) Males




(b) Females

# 2055 Projections (Males)

UNIVERSITY OF Southampton



# Further developments

UNIVERSITY OF Southampton

- Completely integrated estimation
- Bayesian approach with expert opinion and full uncertainty quantification
- Time-varying old-age threshold and/or mortality asymptote
- Models by causes of death

Thank you for listening