Institute and Faculty of Actuaries

Gaussian process regression method for forecasting of mortality rates

Ruhao Wu & Bo Wang
Department of Mathematics, University of Leicester

22 October 2015

Existing models in literature for forecasting mortality rates

- Lee-Carter model (1992), Lee-Miller model (2001)
- Hyndman-Ullah model (2007)

Our Gaussian process regression (GPR) model

- Consider the mortality curve of a specific age group over time to follow a Gaussian process
- Predictive results of different age groups converted into complete mortality-over-age picture of a future year using interpolation

Institute and Faculty of Actuaries

Fundamentals of Gaussian process regression

- Definition: Gaussian process is a stochastic process that any finite subset throughout its domain follows a multivariate normal distribution, a less parametric tool

- A nonlinear regression model with noise:

$$y = f(x) + \varepsilon, \varepsilon \sim N(0, \sigma^2)$$

- A mean function $\mu(\cdot)$ and a covariance function $k(\cdot,\cdot)$ defined for $f(x)$

Institute
and Faculty
of Actuaries

Fundamentals of Gaussian process regression

- The covariance function (kernel function) is defined as:

$$Cov\big(f(x), f(x')\big) = k(x, x'; \theta),$$

where θ denotes the set of hyper-parameters, estimated by empirical Bayesian approach

- Gaussian process regression (GPR) model can then be denoted as:

$$f(x) \sim GPR[\mu(\cdot), k(x, x'; \theta) | x].$$

Institute
and Faculty
of Actuaries

Fundamentals of Gaussian process regression

- Given observed data $\mathcal{D} = \{(x_1, y_1), \dots, (x_n, y_n)\}$,

$$\Psi_{ij} = Cov\left(y_i, y_j\right) = k\left(x_i, x_j; \theta\right) + \sigma^2 \delta_{ij},$$

- $\hat{\theta}$ and $\hat{\sigma}$ calculated by maximizing the marginal log-likelihood

- For a new input x^*,

$$E(f(x^*)|\mathcal{D}) = \mu(x^*) + \psi^T(x^*)\Psi^{-1}(y - \mu),$$
$$Var(f(x^*)|\mathcal{D}) = k\left(x^*, x^*; \hat{\theta}\right) - \psi^T(x^*)\Psi^{-1}\psi(x^*).$$

- $\hat{y}^* = E(y^*|\mathcal{D}) = E(f(x^*)|\mathcal{D})$, $\hat{\sigma}^{*2} = Var(y^*|\mathcal{D}) = Var(f(x^*)|\mathcal{D}) + \hat{\sigma}^2$

- 95% confidence interval: $(\hat{y}^* - 1.96\hat{\sigma}^{*2}, \hat{y}^* + 1.96\hat{\sigma}^{*2})$

GPR models in forecasting mortality rates

- Let $y_x(t)$ denote log of mortality rate for age *x* in year *t*
- Underlying function $f_x(t)$ observed with error at discrete points
- $y_x(t_i) = f_x(t_i) + \varepsilon_{i,x}, x = 1, \ldots, n, i = 1, \ldots, m$
- Forecast $y_x(\mathrm{t})$ for $t \in [t_{m+1}, t_{m+h}]$

Basic GPR models

- Mean function: use a linear function obtained from smoothing the past observed data by linear regression.

- Covariance function: squared exponential (SE), Matern (MA), rational quadratic (RQ)

- $$k_{SE}(\tau) = \sigma^2 \left(-\frac{\tau^2}{2l^2}\right),$$
  $$k_{MA}(\tau) = \sigma^2 \left(1 + \frac{\sqrt{3}}{l}\tau\right) exp\left(-\frac{\sqrt{3}}{l}\tau\right),$$
  $$k_{RQ}(\tau) = \sigma^2 \left(1 + \frac{\tau^2}{2\alpha l^2}\right)^{-\alpha}.$$

  where $\tau = t - t'$.

Modified GPR model – with weighted mean function

- Extrapolation tends to move to prior mean in the long run.
- Previously model the mean function using equally weighted linear regression.
- Makes sense to use weighted least squares (WLS) to obtain mean function
- Parameters chosen to minimize $e = \sum_{i=1}^{m} z_i \left(y_i - \hat{y}_i\right)^2$, where $z_i = 1/(t_0 - t_i)$

Institute and Faculty of Actuaries

Modified GPR model – with spectral mixture kernels

- Adopt an idea raised by Wilson & Adam (2013): introduces simple closed form kernels derived by modelling a spectral density with a Gaussian mixture.

- $k(\tau) = \int_{R^P} e^{2\pi i s^T \tau} \varphi(ds), \varphi$ is a positive finite measure, has density $S(s)$

- $k(\tau) = \int_{R^P} S(s)\, e^{2\pi i s^T \tau}\, ds,\ \ S(s) = \int_{R^P} k(\tau)\, e^{-2\pi i s^T \tau}\, d\tau$.

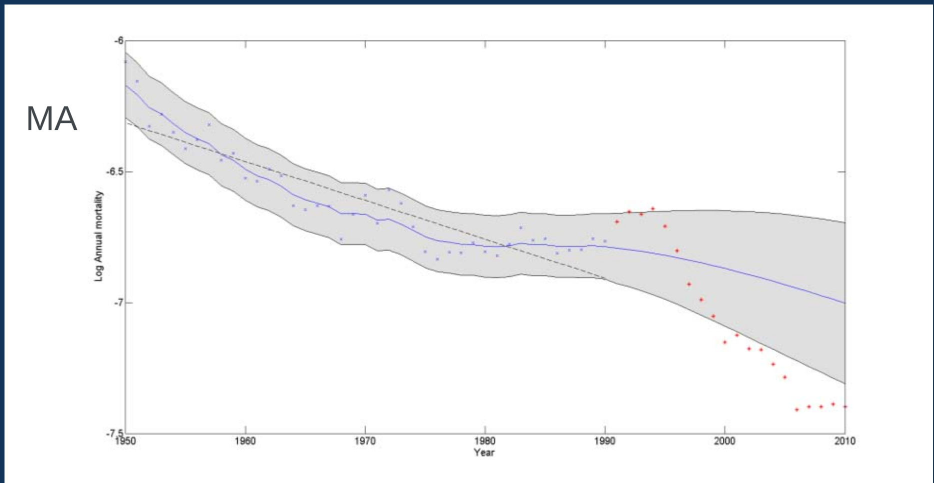
- Wilson and Adam (2013): any stationary covariance kernels can be approximated to arbitrary precision using mixture of Gaussians in spectral density.

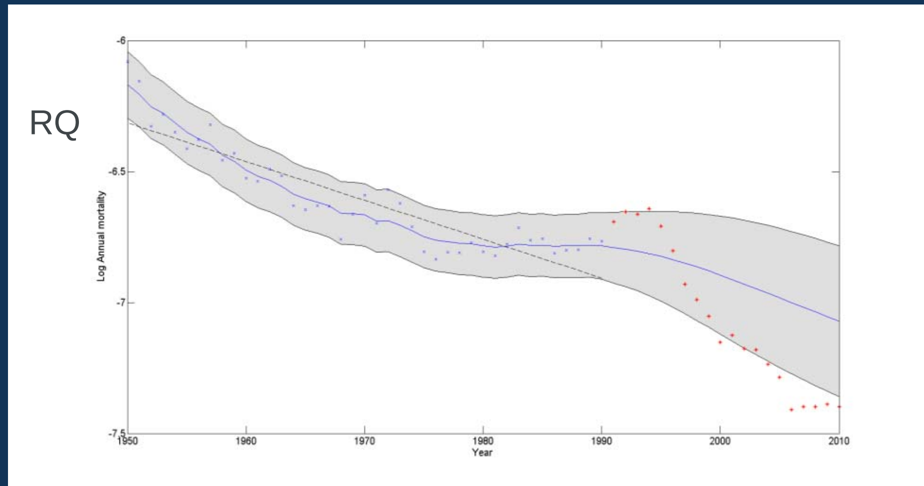
- Model $S(s)$ to be mixture of Gaussian, extend to P dimensions.

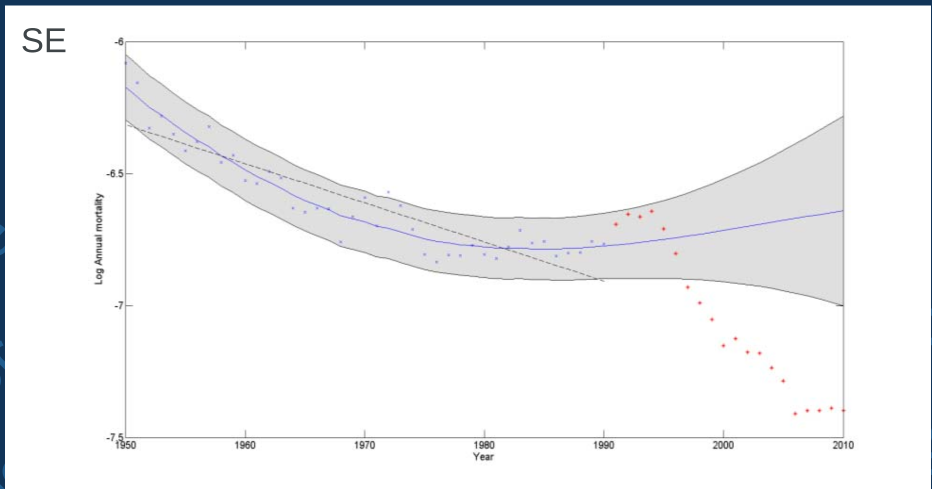
Applications of GPR models in forecasting French mortality rates

- Quote French total mortality data from Human Mortality Database (HMD)
- Data from 1950-2010: 1950-1990 as training data, 1991-2010 as testing data
- Basic GPR models, using SE, MA and RQ as kernels
- Pick out 20, 30, 40 and 50 years age group for analysis

Applications of GPR models in forecasting French mortality rates – Basic GPR models







Institute
and Faculty
of Actuaries

Basic GPR models & SM GPR models

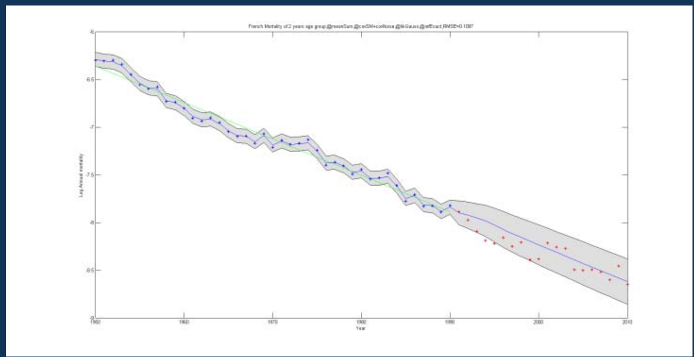
- Table 1.Record of RMSE of French log mortality of 20, 30, 40 and 50 years group using SE, MA, RQ and SM kernels respectively

Age group	Basic GPR			SM GPR
	SE	MA	RQ	
20	0.3516	0.4245	0.4808	0.5261
30	0.4752	0.2810	0.2445	0.1621
40	0.1017	0.1111	0.1201	0.1048
50	0.3026	0.0499	0.0540	0.0559

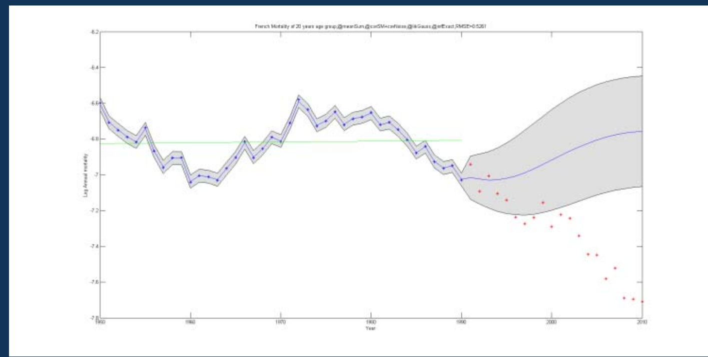
GPR model with spectral mixture (SM) kernel

- Pick out 17 specific age groups: 0, 1, 2, 5, 10, 12, 15, 18, 20, 30, 40, 50, 60, 70, 80, 90, 100 years group

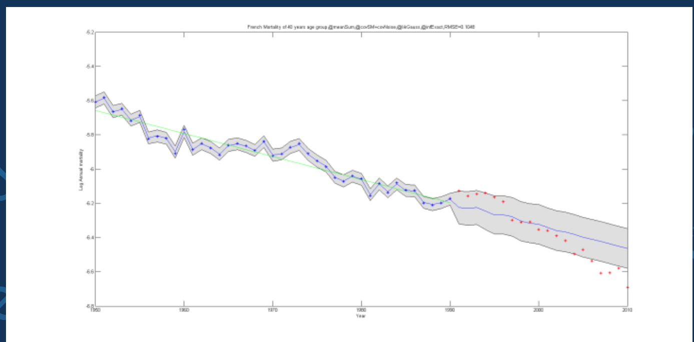
2



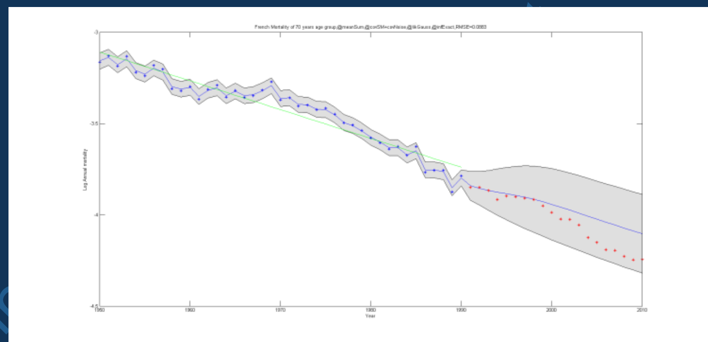
20



40



70



GPR model with spectral mixture (SM) kernel and weighted mean function

Age group	SM GPR	SM GPR with weighted mean
0	0.1558	0.1324
1	0.4157	0.0930
2	0.1097	0.0993
5	0.4204	0.2058
10	0.4566	0.3584
12	0.3182	0.2319
15	0.3625	0.2810
18	0.6495	0.3427
20	0.5261	0.3775
30	0.1621	0.3052
40	0.1043	0.0975
50	0.0559	0.0725
60	0.1211	0.0499
70	0.0883	0.0493
80	0.1453	0.0742
90	0.0619	0.0571
100	0.1200	0.0839

Average RMSE

SM GPR:0.2514

SM GPR with weighted mean function:0.1713

Institute and Faculty of Actuaries

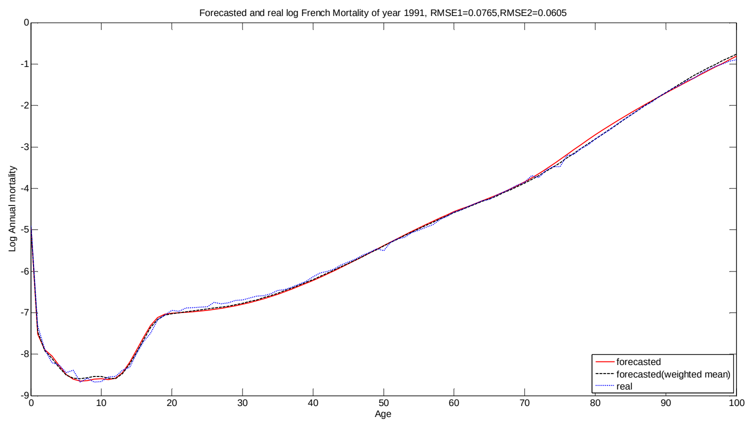
Generating the complete mortality picture

- 1991 (1-year horizon), 2000 (10-year horizon), 2010 (20-year horizon)
- Estimate log mortality of the 17 age groups for the three years, use splines for interpolation

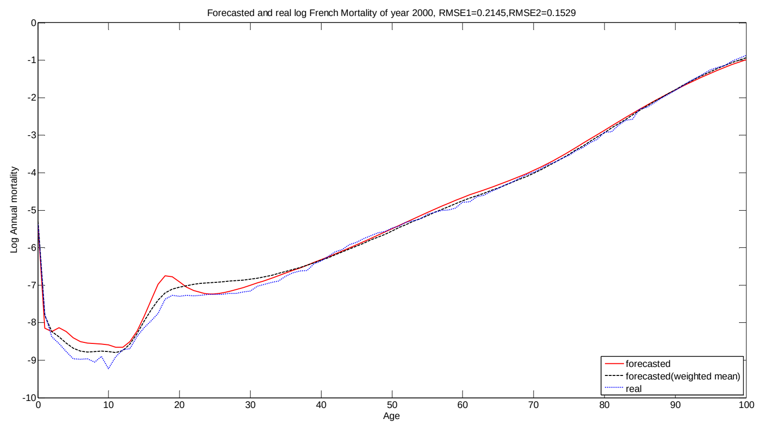
Institute and Faculty of Actuaries

Generating the complete mortality picture

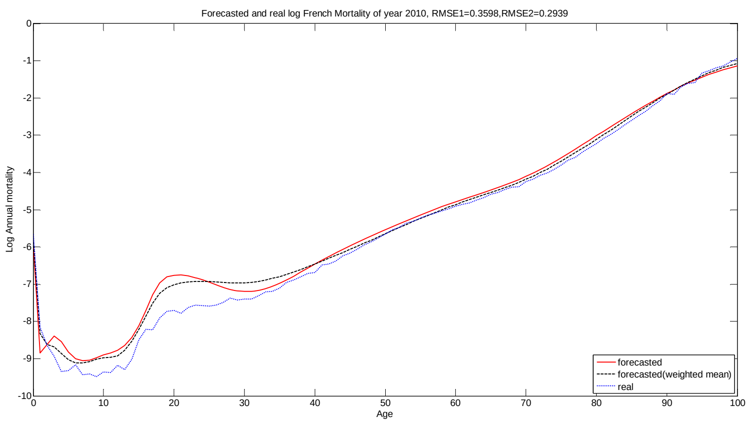
1991



2000



2010



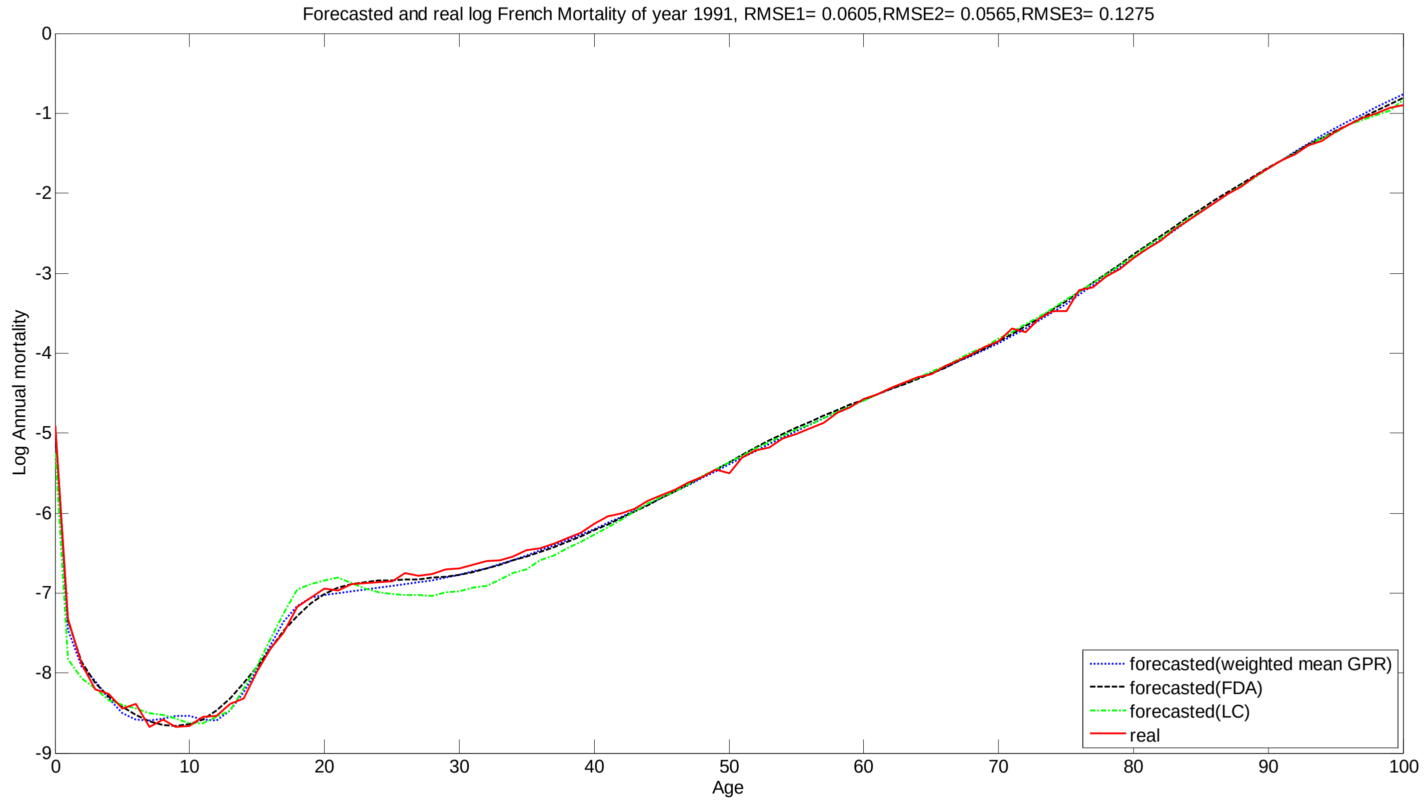
RMSE

SM GPR: 0.0765, 0.2145, 0.3598

SM GPR with weighted mean function: 0.0605, 0.1529, 0.2939

Compare with LC model and FDA model

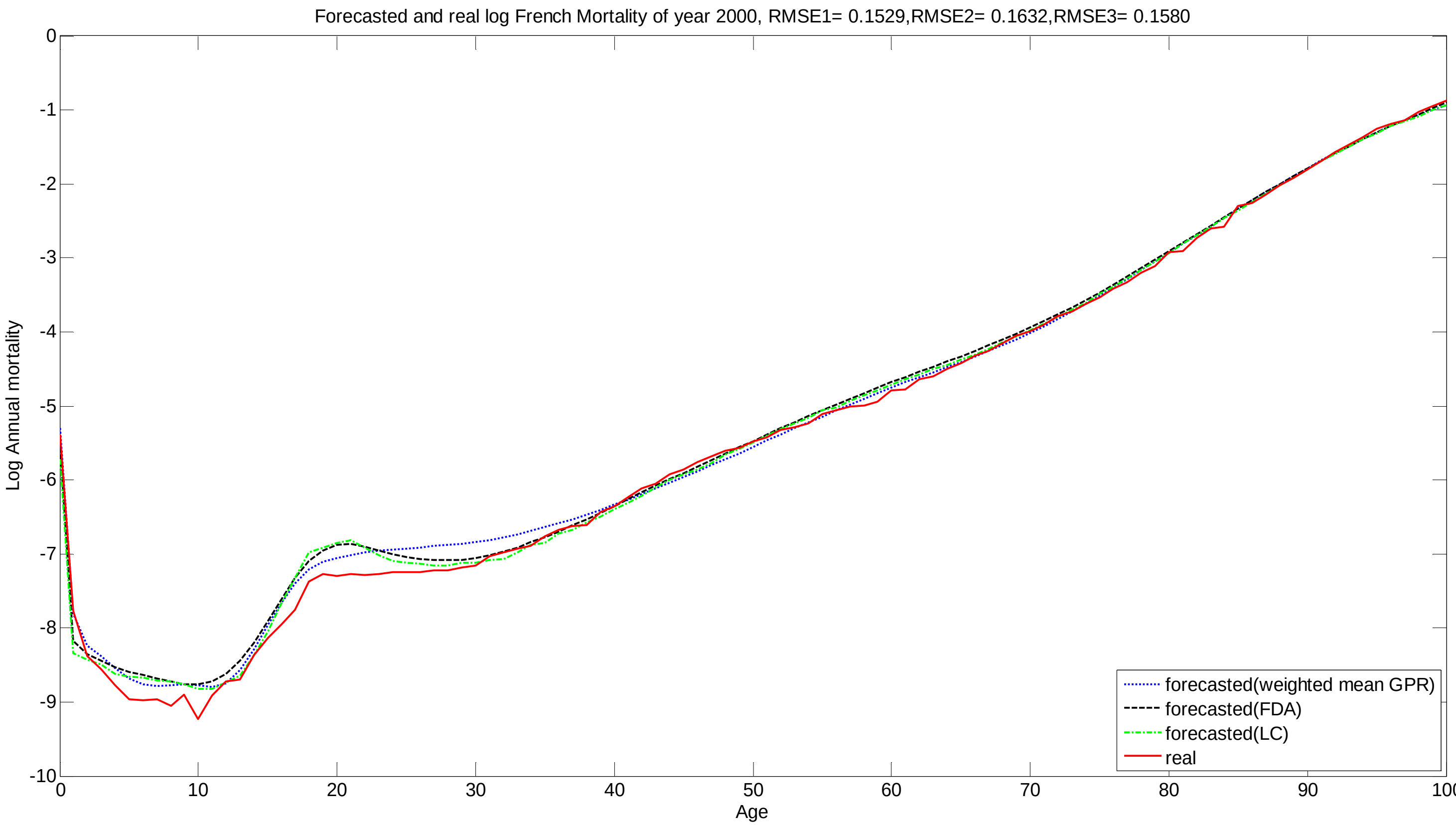
1991



Institute and Faculty of Actuaries

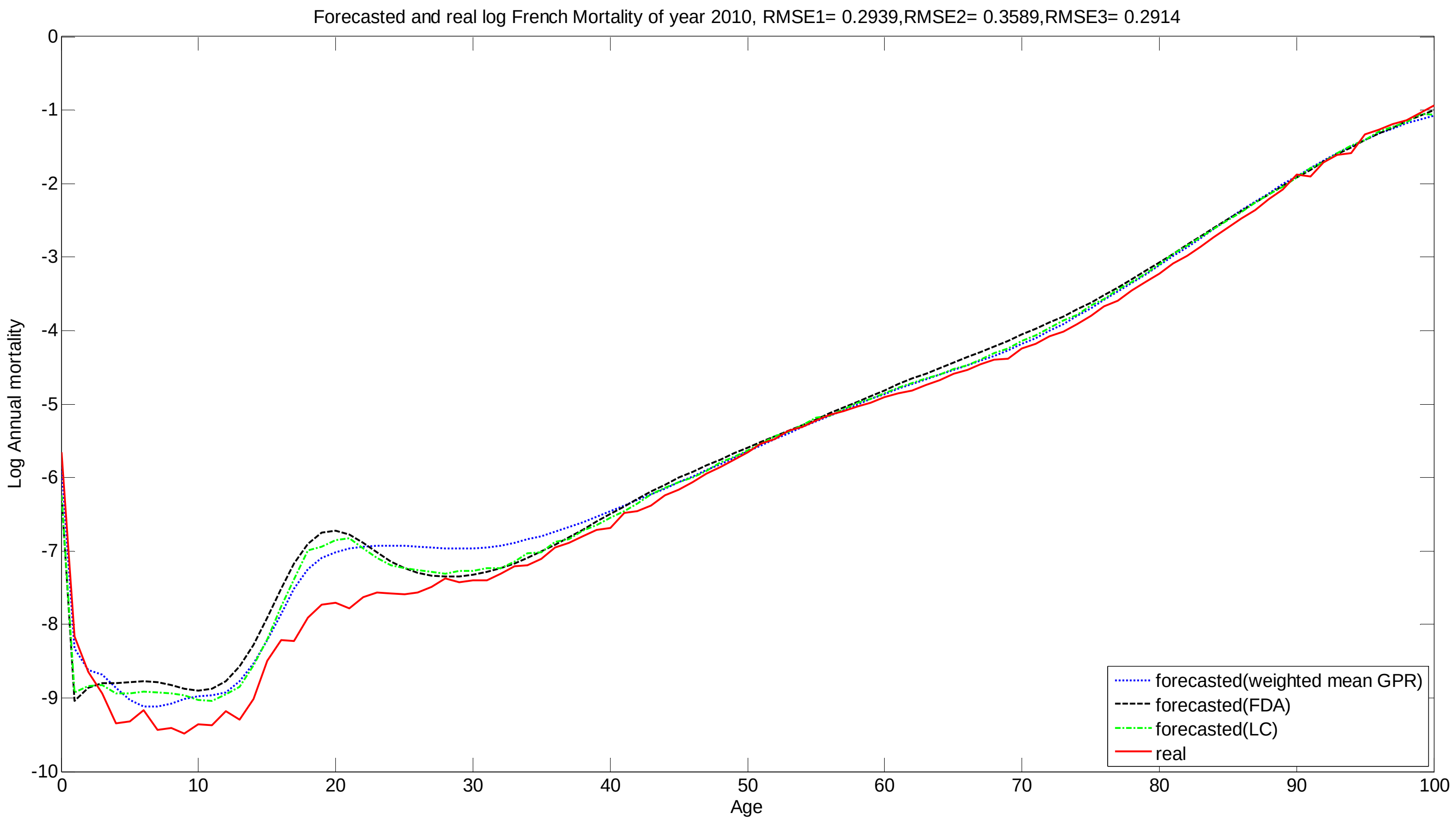
Compare with LC model and FDA model

2000



Compare with LC model and FDA model

2010



Compare with LC model and FDA model

RMSE

	1991	2000	2010
SM GPR WM	0.0605	0.1529	0.2939
FDA	0.0565	0.1632	0.3589
LC	0.1275	0.1580	0.2914

Institute and Faculty of Actuaries

The End

The project is jointly funded by the IFoA and the College of Science and Engineering, University of Leicester. Thanks IFoA for the support to attend this conference!

Q&A