Université Claude Bernard Lyon 1
Institut de Science Financière et d'Assurances

A New Non-parametric approach to Smoothing and Forescating of Mortality

Julien Tomas - Frédéric Planchet
Institut de Science Financière et d'Assurances
Laboratoire de recherche de Sciences Actuarielle et Financière

IAA LIFE SECTION
SECTION VIE DE L'AAI

Contents

IAA LIFE SECTION
SECTION VIE DE L'AAI

Contents

IAA LIFE SECTION
SECTION VIE DE L'AAI

Contents

IAA LIFE SECTION
SECTION VIE DE L'AAI

Contents

IAA LIFE SECTION
SECTION VIE DE L'AAI

Contents

IAA LIFE SECTION
SECTION VIE DE L'AAI

Contents

IAA LIFE SECTION
SECTION VIE DE L'AAI

Contents

IAA LIFE SECTION
SECTION VIE DE L'AAI

Contents

IAA LIFE SECTION
SECTION VIE DE L'AAI

Contents

IAA LIFE SECTION
SECTION VIE DE L'AAI

Contents

IAA LIFE SECTION
SECTION VIE DE L'AAI

Contents

IAA LIFE SECTION
SECTION VIE DE L'AAI

Contents

IAA LIFE SECTION
SECTION VIE DE L'AAI

Contents

IAA LIFE SECTION
SECTION VIE DE L'AAI

Contents

IAA LIFE SECTION
SECTION VIE DE L'AAI

A a new robust and strictly non-parametric framework

- Our method is motivated by avoiding or at least minimizing the restrictive conditions as imposed by the structure of the Bongaarts (2005) or Cairns *et al.* (2006) models.
- The innovative aspect lies in
  - The articulation of the two steps : adjustment of the past mortality and extrapolation of the temporal trends of mortality
  - The choices made in technical terms about these two points.

A a new robust and strictly non-parametric framework

- Our method is motivated by avoiding or at least minimizing the restrictive conditions as imposed by the structure of the Bongaarts (2005) or Cairns *et al.* (2006) models.
- The innovative aspect lies in
  - The articulation of the two steps :
    - adjustment of the past mortality
    - and extrapolation of the tendencies of mortality
  - The choices made in technical terms about these two points.

A a new robust and strictly non-parametric framework

- Our method is motivated by avoiding or at least minimizing the restrictive conditions as imposed by the structure of the Bongaarts (2005) or Cairns *et al.* (2006) models.
- The innovative aspect lies in
  - The articulation of the two steps :
    - adjustment of the past mortality
    - and extrapolation of the tendencies of mortality
  - The choices made in technical terms about these two points.

IAA LIFE SECTION
SECTION VIE DE L'AAI

A a new robust and strictly non-parametric framework

- Our method is motivated by avoiding or at least minimizing the restrictive conditions as imposed by the structure of the Bongaarts (2005) or Cairns *et al.* (2006) models.
- The innovative aspect lies in
  - The articulation of the two steps :
    - adjustment of the past mortality
    - and extrapolation of the tendencies of mortality
  - The choices made in technical terms about these two points.

A a new robust and strictly non-parametric framework

- Our method is motivated by avoiding or at least minimizing the restrictive conditions as imposed by the structure of the Bongaarts (2005) or Cairns *et al.* (2006) models.
- The innovative aspect lies in
  - The articulation of the two steps :
    - adjustment of the past mortality
    - and extrapolation of the tendencies of mortality
  - The choices made in technical terms about these two points.

A a new robust and strictly non-parametric framework

- Our method is motivated by avoiding or at least minimizing the restrictive conditions as imposed by the structure of the Bongaarts (2005) or Cairns *et al.* (2006) models.
- The innovative aspect lies in
  - The articulation of the two steps :
    - adjustment of the past mortality
    - and extrapolation of the tendencies of mortality
  - The choices made in technical terms about these two points.

Two famillies of approaches

Two approaches can be proposed

- Endogenous approaches, which consist of exploiting the information contained in the crude forces of mortality
  - To obtain a smooth surface representing the data correctly and to yield a *realistic* projection.
  - In case of a small volume of data, these techniques could lead to biased estimations of the mortality trend.
- Exogenous approaches. Models using an external reference mortality table
  - A solution to overcome the difficulties associated with having a small volume of data.
  - The idea is to adjust a reference table to the experience of a given set of data.

Two famillies of approaches

Two approaches can be proposed

- Endogenous approaches, which consist of exploiting the information contained in the crude forces of mortality
  - To obtain a smooth surface representing the data correctly and to yield a *realistic* projection.
  - In case of a small volume of data, these techniques could lead to biased estimations of the mortality trend.
- Exogenous approaches. Models using an external reference mortality table
  - A solution to overcome the difficulties associated with having a small volume of data.
  - The idea is to adjust a reference table to the experience of a given set of data.

Two famillies of approaches

Two approaches can be proposed

- Endogenous approaches, which consist of exploiting the information contained in the crude forces of mortality
  - To obtain a smooth surface representing the data correctly and to yield a *realistic* projection.
  - In case of a small volume of data, these techniques could lead to biased estimations of the mortality trend.
- Exogenous approaches. Models using an external reference mortality table
  - A solution to overcome the difficulties associated with having a small volume of data.
  - The idea is to adjust a reference table to the experience of a given set of data.

Two famillies of approaches

Two approaches can be proposed

- Endogenous approaches, which consist of exploiting the information contained in the crude forces of mortality
  - To obtain a smooth surface representing the data correctly and to yield a *realistic* projection.
  - In case of a small volume of data, these techniques could lead to biased estimations of the mortality trend.
- Exogenous approaches. Models using an external reference mortality table
  - A solution to overcome the difficulties associated with having a small volume of data.
  - The idea is to adjust a reference table to the experience of a given set of data.

Two famillies of approaches

Two approaches can be proposed

- Endogenous approaches, which consist of exploiting the information contained in the crude forces of mortality
  - To obtain a smooth surface representing the data correctly and to yield a *realistic* projection.
  - In case of a small volume of data, these techniques could lead to biased estimations of the mortality trend.
- Exogenous approaches. Models using an external reference mortality table
  - A solution to overcome the difficulties associated with having a small volume of data.
  - The idea is to adjust a reference table to the experience of a given set of data.

Two famillies of approaches

Two approaches can be proposed

- Endogenous approaches, which consist of exploiting the information contained in the crude forces of mortality
  - To obtain a smooth surface representing the data correctly and to yield a *realistic* projection.
  - In case of a small volume of data, these techniques could lead to biased estimations of the mortality trend.
- Exogenous approaches. Models using an external reference mortality table
  - A solution to overcome the difficulties associated with having a small volume of data.
  - The idea is to adjust a reference table to the experience of a given set of data.

Contents

Notation & assumption

- We analyze the mortality as a function of both the attained age x and the calendar year t.
- $D_{x,t}$ denotes the number of deaths recorded at attained age x during calendar year t from an exposure-to-risk $E_{x,t}$.
- We note $p_x(t)$ is the probability that an individual aged x in calendar year t reaches age $x+1$, and by $q_x(t) = 1 - p_x(t)$ the corresponding probability of death. The force of mortality is denoted by $\varphi_x(t)$.
- Assume that the age-specific forces of mortality are constant within bands of time, but allowed to vary from one band to the next,

$$\varphi_{x+\tau}(t+\xi) = \varphi_x(t) \text{ for } 0 \leq \tau < 1 \text{ and } 0 \leq \xi < 1.$$

- Under this assumption

$$p_x(t) = \exp(-\varphi_x(t)) \Leftrightarrow \varphi_x(t) = -\log(p_x(t)).$$

Notation & assumption

- We analyze the mortality as a function of both the attained age x and the calendar year t.
- $D_{x,t}$ denotes the number of deaths recorded at attained age x during calendar year t from an exposure-to-risk $E_{x,t}$.
- We note $p_x(t)$ is the probability that an individual aged x in calendar year t reaches age $x+1$, and by $q_x(t) = 1 - p_x(t)$ the corresponding probability of death. The force of mortality is denoted by $\varphi_x(t)$.
- Assume that the age-specific forces of mortality are constant within bands of time, but allowed to vary from one band to the next,

$$\varphi_{x+\tau}(t+\xi) = \varphi_x(t) \text{ for } 0 \leq \tau < 1 \text{ and } 0 \leq \xi < 1.$$

- Under this assumption

$$p_x(t) = \exp(-\varphi_x(t)) \Leftrightarrow \varphi_x(t) = -\log(p_x(t)).$$

IAA LIFE SECTION
SECTION VIE DE L'AAI

Notation & assumption

- We analyze the mortality as a function of both the attained age x and the calendar year t.
- $D_{x,t}$ denotes the number of deaths recorded at attained age x during calendar year t from an exposure-to-risk $E_{x,t}$.
- We note $p_x(t)$ is the probability that an individual aged x in calendar year t reaches age $x+1$, and by $q_x(t) = 1 - p_x(t)$ the corresponding probability of death. The force of mortality is denoted by $\varphi_x(t)$.
- Assume that the age-specific forces of mortality are constant within bands of time, but allowed to vary from one band to the next,

$$\varphi_{x+\tau}(t+\xi) = \varphi_x(t) \text{ for } 0 \leq \tau < 1 \text{ and } 0 \leq \xi < 1.$$

- Under this assumption

$$p_x(t) = \exp(-\varphi_x(t)) \Leftrightarrow \varphi_x(t) = -\log(p_x(t)).$$

IAA LIFE SECTION
SECTION VIE DE L'AAI

Notation & assumption

- We analyze the mortality as a function of both the attained age x and the calendar year t.
- $D_{x,t}$ denotes the number of deaths recorded at attained age x during calendar year t from an exposure-to-risk $E_{x,t}$.
- We note $p_x(t)$ is the probability that an individual aged x in calendar year t reaches age $x+1$, and by $q_x(t) = 1 - p_x(t)$ the corresponding probability of death. The force of mortality is denoted by $\varphi_x(t)$.
- Assume that the age-specific forces of mortality are constant within bands of time, but allowed to vary from one band to the next,

$$\varphi_{x+\tau}(t+\xi) = \varphi_x(t) \text{ for } 0 \leq \tau < 1 \text{ and } 0 \leq \xi < 1.$$

- Under this assumption

$$p_x(t) = \exp(-\varphi_x(t)) \Leftrightarrow \varphi_x(t) = -\log(p_x(t)).$$

IAA LIFE SECTION
SECTION VIE DE L'AAI

Notation & assumption

- We analyze the mortality as a function of both the attained age x and the calendar year t.
- $D_{x,t}$ denotes the number of deaths recorded at attained age x during calendar year t from an exposure-to-risk $E_{x,t}$.
- We note $p_x(t)$ is the probability that an individual aged x in calendar year t reaches age $x+1$, and by $q_x(t) = 1 - p_x(t)$ the corresponding probability of death. The force of mortality is denoted by $\varphi_x(t)$.
- Assume that the age-specific forces of mortality are constant within bands of time, but allowed to vary from one band to the next,

$$\varphi_{x+\tau}(t+\xi) = \varphi_x(t) \text{ for } 0 \leq \tau < 1 \text{ and } 0 \leq \xi < 1.$$

- Under this assumption

$$p_x(t) = \exp(-\varphi_x(t)) \Leftrightarrow \varphi_x(t) = -\log(p_x(t)).$$

Summary of our methodology

i. Smooth the surface using non-parametric local kernel-weighted log-likelihood to estimate $\varphi_x(t)$ for $x \in [x_1, x_n]$ and $t = 1, \ldots, m$.

ii. Decompose the smoothed surfaces via a basis function expansion

$$y_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k}\, \phi_k(x) + \varepsilon_t(x), \varepsilon_t(x) \sim \mathcal{N}(0, v(x)), \quad (1)$$

where $y_t(x) = \log \widehat{\varphi}_x(t)$, $\mu(x)$ is the mean of $\log \widehat{\varphi}_x(t)$ across years and $\{\phi_k(x)\}$ is a set of orthonormal basis functions.

iii. ARIMA models are fitted to each of the $\{\beta_{t,k}\}$, $k = 1, \ldots, K$.

iv. Extrapolate the coefficients $\{\beta_{t,k}\}$, $k = 1, \ldots, K$, for $t = m+1, \ldots, m+h$ using the fitted time series models.

v. Finally, use the resulting forecast coefficients with (1) to obtain forecasts of $y_t(x)$, $t = m+1, \ldots, m+h$, and hence of $\varphi_x(t)$.

Summary of our methodology

i. Smooth the surface using non-parametric local kernel-weighted log-likelihood to estimate $\varphi_x(t)$ for $x \in [x_1, x_n]$ and $t = 1, \ldots, m$.

ii. Decompose the smoothed surfaces via a basis function expansion

$$y_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k}\, \phi_k(x) + \varepsilon_t(x), \varepsilon_t(x) \sim \mathcal{N}(0, \upsilon(x)), \quad (1)$$

where $y_t(x) = \log \widehat{\varphi}_x(t)$, $\mu(x)$ is the mean of $\log \widehat{\varphi}_x(t)$ across years and $\{\phi_k(x)\}$ is a set of orthonormal basis functions.

iii. ARIMA models are fitted to each of the $\{\beta_{t,k}\}$, $k = 1, \ldots, K$.

iv. Extrapolate the coefficients $\{\beta_{t,k}\}$, $k = 1, \ldots, K$, for $t = m+1, \ldots, m+h$ using the fitted time series models.

v. Finally, use the resulting forecast coefficients with (1) to obtain forecasts of $y_t(x)$, $t = m+1, \ldots, m+h$, and hence of $\varphi_x(t)$.

Summary of our methodology

i. Smooth the surface using non-parametric local kernel-weighted log-likelihood to estimate $\varphi_x(t)$ for $x \in [x_1, x_n]$ and $t = 1, \ldots, m$.

ii. Decompose the smoothed surfaces via a basis function expansion

$$y_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k}\, \phi_k(x) + \varepsilon_t(x), \varepsilon_t(x) \sim \mathcal{N}(0, v(x)), \quad (1)$$

where $y_t(x) = \log \widehat{\varphi}_x(t)$, $\mu(x)$ is the mean of $\log \widehat{\varphi}_x(t)$ across years and $\{\phi_k(x)\}$ is a set of orthonormal basis functions.

iii. ARIMA models are fitted to each of the $\{\beta_{t,k}\}$, $k = 1, \ldots, K$.

iv. Extrapolate the coefficients $\{\beta_{t,k}\}$, $k = 1, \ldots, K$, for $t = m+1, \ldots, m+h$ using the fitted time series models.

v. Finally, use the resulting forecast coefficients with (1) to obtain forecasts of $y_t(x)$, $t = m+1, \ldots, m+h$, and hence of $\varphi_x(t)$.

Summary of our methodology

i. Smooth the surface using non-parametric local kernel-weighted log-likelihood to estimate $\varphi_x(t)$ for $x \in [x_1, x_n]$ and $t = 1, \ldots, m$.

ii. Decompose the smoothed surfaces via a basis function expansion

$$y_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k}\, \phi_k(x) + \varepsilon_t(x), \varepsilon_t(x) \sim \mathcal{N}(0, v(x)), \quad (1)$$

where $y_t(x) = \log \widehat{\varphi}_x(t)$, $\mu(x)$ is the mean of $\log \widehat{\varphi}_x(t)$ across years and $\{\phi_k(x)\}$ is a set of orthonormal basis functions.

iii. ARIMA models are fitted to each of the $\{\beta_{t,k}\}$, $k = 1, \ldots, K$.

iv. Extrapolate the coefficients $\{\beta_{t,k}\}$, $k = 1, \ldots, K$, for $t = m+1, \ldots, m+h$ using the fitted time series models.

v. Finally, use the resulting forecast coefficients with (1) to obtain forecasts of $y_t(x)$, $t = m+1, \ldots, m+h$, and hence of $\varphi_x(t)$.

Summary of our methodology

i. Smooth the surface using non-parametric local kernel-weighted log-likelihood to estimate $\varphi_x(t)$ for $x \in [x_1, x_n]$ and $t = 1, \ldots, m$.

ii. Decompose the smoothed surfaces via a basis function expansion

$$y_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k}\, \phi_k(x) + \varepsilon_t(x), \varepsilon_t(x) \sim \mathcal{N}(0, \upsilon(x)), \quad (1)$$

where $y_t(x) = \log \widehat{\varphi}_x(t)$, $\mu(x)$ is the mean of $\log \widehat{\varphi}_x(t)$ across years and $\{\phi_k(x)\}$ is a set of orthonormal basis functions.

iii. ARIMA models are fitted to each of the $\{\beta_{t,k}\}$, $k = 1, \ldots, K$.

iv. Extrapolate the coefficients $\{\beta_{t,k}\}$, $k = 1, \ldots, K$, for $t = m+1, \ldots, m+h$ using the fitted time series models.

v. Finally, use the resulting forecast coefficients with (1) to obtain forecasts of $y_t(x)$, $t = m+1, \ldots, m+h$, and hence of $\varphi_x(t)$.

IAA LIFE SECTION
SECTION VIE DE L'AAI

The proposed models

Model	Formula	Ref. table	Estimation method: Local lik.	Estimation method: Min. dist.
M1	$D_{x,t} \sim \text{Poisson}\left(E_{x,t}\ \exp\left(f_1(x,t)\right)\right)$		M1	
M2	$D_{x,t} \sim \text{Poisson}\left(E_{x,t}\ \exp\left(f_2\left(\log\left(\varphi_x^{\text{ref}}(t)\right)\right)\right)\right)$	INSEE	M2.INSEE	
		TG05	M2.TG05	
M3	$D_{x,t} \sim \text{Poisson}\left(E_{x,t}\ \varphi_x^{\text{ref}}(t)\ \exp\left(f_1(x,t)\right)\right)$	INSEE	M3.INSEE	
		TG05	M3.TG05	
M4	$D_{x,t} \sim \text{Poisson}\left(E_{x,t}\ \exp\left(f_1(x,t) + f_2\left(\log\left(\varphi_x^{\text{ref}}(t)\right)\right)\right)\right)$	INSEE	M4.INSEE	
		TG05	M4.TG05	
M5	$\text{logit}\ \varphi_x(t) = \alpha + \beta\ \text{logit}\ \varphi_x^{\text{ref}}(t) + \epsilon_{x,t}$	INSEE		M5.INSEE
		TG05		M5.TG05

TABLE: Description of the models and estimation method used in the first step.

Estimation techniques for the preliminary smoothing

M1, M2, M3 and M4 are estimated by non-parametric methods

- We considered local kernel-weighted log-likelihood methods to estimate the smooth functions $f_1(x,t)$ and $f_2\left(\log\left(\varphi_x^{\text{ref}}(t)\right)\right)$ for $x \in [x_1, x_n]$ and $t = 1, \ldots, m$.
- Statistical aspects of local likelihood techniques have been discussed extensively in Loader (1996), Fan *et al.* (1998), Loader (1999) and Tomas (2013).
- These methods have been used in a mortality context by Delwarde *et al.* (2004), Debón *et al.* (2006), Tomas (2011) and Gschlössl *et al.* (2011) to graduate life tables with attained age. More recently, Tomas and Planchet (2013a) have covered smoothing in two dimension and introduced adaptive parameters choice with an application to long-term care insurance.

Estimation techniques for the preliminary smoothing

M1, M2, M3 and M4 are estimated by non-parametric methods

- We considered local kernel-weighted log-likelihood methods to estimate the smooth functions $f_1(x,t)$ and $f_2\left(\log\left(\varphi_x^{\text{ref}}(t)\right)\right)$ for $x \in [x_1, x_n]$ and $t = 1, \ldots, m$.
- Statistical aspects of local likelihood techniques have been discussed extensively in Loader (1996), Fan *et al.* (1998), Loader (1999) and Tomas (2013).
- These methods have been used in a mortality context by Delwarde *et al.* (2004), Debón *et al.* (2006), Tomas (2011) and Gschlössl *et al.* (2011) to graduate life tables with attained age. More recently, Tomas and Planchet (2013a) have covered smoothing in two dimension and introduced adaptive parameters choice with an application to long-term care insurance.

Estimation techniques for the preliminary smoothing

M1, M2, M3 and M4 are estimated by non-parametric methods

- We considered local kernel-weighted log-likelihood methods to estimate the smooth functions $f_1(x,t)$ and $f_2\left(\log\left(\varphi_x^{\text{ref}}(t)\right)\right)$ for $x \in [x_1, x_n]$ and $t = 1, \ldots, m$.
- Statistical aspects of local likelihood techniques have been discussed extensively in Loader (1996), Fan *et al.* (1998), Loader (1999) and Tomas (2013).
- These methods have been used in a mortality context by Delwarde *et al.* (2004), Debón *et al.* (2006), Tomas (2011) and Gschlössl *et al.* (2011) to graduate life tables with attained age. More recently, Tomas and Planchet (2013a) have covered smoothing in two dimension and introduced adaptive parameters choice with an application to long-term care insurance.

Contents

IAA LIFE SECTION
SECTION VIE DE L'AAI

Lee and Carter (1992) and its variants

- Future trends will essentially be a continuation of the past.
- Lee and Carter (1992) or its variants are now the dominant methods of mortality forecasting in actuarial sciences.
- Improvements have been proposed. A Poisson log-likelihood approach has been developed in Brouhns *et al.* (2002b), Brouhns *et al.* (2002a) and Renshaw and Haberman (2003). Cosette *et al.* (2007) use a binomial maximum likelihood, and a negative binomial version of the Lee-Carter model has been developed by Delwarde *et al.* (2007).
- The functional PCA proposed by Hyndman and Ullah (2007) and Hyndman and Booth (2008) can be considered as a successor.

Lee and Carter (1992) and its variants

- Future trends will essentially be a continuation of the past.
- Lee and Carter (1992) or its variants are now the dominant methods of mortality forecasting in actuarial sciences.
- Improvements have been proposed. A Poisson log-likelihood approach has been developed in Brouhns *et al.* (2002b), Brouhns *et al.* (2002a) and Renshaw and Haberman (2003). Cosette *et al.* (2007) use a binomial maximum likelihood, and a negative binomial version of the Lee-Carter model has been developed by Delwarde *et al.* (2007).
- The functional PCA proposed by Hyndman and Ullah (2007) and Hyndman and Booth (2008) can be considered as a successor.

Lee and Carter (1992) and its variants

- Future trends will essentially be a continuation of the past.
- Lee and Carter (1992) or its variants are now the dominant methods of mortality forecasting in actuarial sciences.
- Improvements have been proposed. A Poisson log-likelihood approach has been developed in Brouhns *et al.* (2002b), Brouhns *et al.* (2002a) and Renshaw and Haberman (2003). Cosette *et al.* (2007) use a binomial maximum likelihood, and a negative binomial version of the Lee-Carter model has been developed by Delwarde *et al.* (2007).
- The functional PCA proposed by Hyndman and Ullah (2007) and Hyndman and Booth (2008) can be considered as a successor.

Lee and Carter (1992) and its variants

- Future trends will essentially be a continuation of the past.
- Lee and Carter (1992) or its variants are now the dominant methods of mortality forecasting in actuarial sciences.
- Improvements have been proposed. A Poisson log-likelihood approach has been developed in Brouhns *et al.* (2002b), Brouhns *et al.* (2002a) and Renshaw and Haberman (2003). Cosette *et al.* (2007) use a binomial maximum likelihood, and a negative binomial version of the Lee-Carter model has been developed by Delwarde *et al.* (2007).
- The functional PCA proposed by Hyndman and Ullah (2007) and Hyndman and Booth (2008) can be considered as a successor.

Functional principal components analysis

- The approach differs by the use of the functional data paradigm, see Ramsay and Silverman (2005) and Ferraty and Vieu (2006) for an introduction to functional nonparametric statistics.
- The decomposition using an orthonormal basis (step **ii.**) is obtained via functional principal components analysis developed by Ramsay and Dalzell (1991)
- This basis set provides informative interpretation and coefficients $\{\beta_{t,k}\}$ that are uncorrelated.
- The number K of basis functions depends on many considerations :
  - On the number of discrete points m in the original data, whether some level of smoothing is imposed by using $K < m$.
  - On the efficiency of the basis functions in reproducing the behavior of the original functions.

Functional principal components analysis

- The approach differs by the use of the functional data paradigm, see Ramsay and Silverman (2005) and Ferraty and Vieu (2006) for an introduction to functional nonparametric statistics.
- The decomposition using an orthonormal basis (step **ii.**) is obtained via functional principal components analysis developed by Ramsay and Dalzell (1991)
- This basis set provides informative interpretation and coefficients $\{\beta_{t,k}\}$ that are uncorrelated.
- The number K of basis functions depends on many considerations :
  - On the number of discrete points m in the original data, whether some level of smoothing is imposed by using $K < m$.
  - On the efficiency of the basis functions in reproducing the behavior of the original functions.

Functional principal components analysis

- The approach differs by the use of the functional data paradigm, see Ramsay and Silverman (2005) and Ferraty and Vieu (2006) for an introduction to functional nonparametric statistics.
- The decomposition using an orthonormal basis (step **ii.**) is obtained via functional principal components analysis developed by Ramsay and Dalzell (1991)
- This basis set provides informative interpretation and coefficients $\{\beta_{t,k}\}$ that are uncorrelated.
- The number K of basis functions depends on many considerations :
  - On the number of discrete points m in the original data, whether some level of smoothing is imposed by using $K < m$.
  - On the efficiency of the basis functions in reproducing the behavior of the original functions.

Functional principal components analysis

- The approach differs by the use of the functional data paradigm, see Ramsay and Silverman (2005) and Ferraty and Vieu (2006) for an introduction to functional nonparametric statistics.
- The decomposition using an orthonormal basis (step **ii.**) is obtained via functional principal components analysis developed by Ramsay and Dalzell (1991)
- This basis set provides informative interpretation and coefficients $\{\beta_{t,k}\}$ that are uncorrelated.
- The number K of basis functions depends on many considerations :
  - On the number of discrete points m in the original data, whether some level of smoothing is imposed by using $K < m$.
  - On the efficiency of the basis functions in reproducing the behavior of the original functions.

Functional principal components analysis

- The approach differs by the use of the functional data paradigm, see Ramsay and Silverman (2005) and Ferraty and Vieu (2006) for an introduction to functional nonparametric statistics.
- The decomposition using an orthonormal basis (step **ii.**) is obtained via functional principal components analysis developed by Ramsay and Dalzell (1991)
- This basis set provides informative interpretation and coefficients $\{\beta_{t,k}\}$ that are uncorrelated.
- The number K of basis functions depends on many considerations :
  - On the number of discrete points m in the original data, whether some level of smoothing is imposed by using $K < m$.
  - On the efficiency of the basis functions in reproducing the behavior of the original functions.

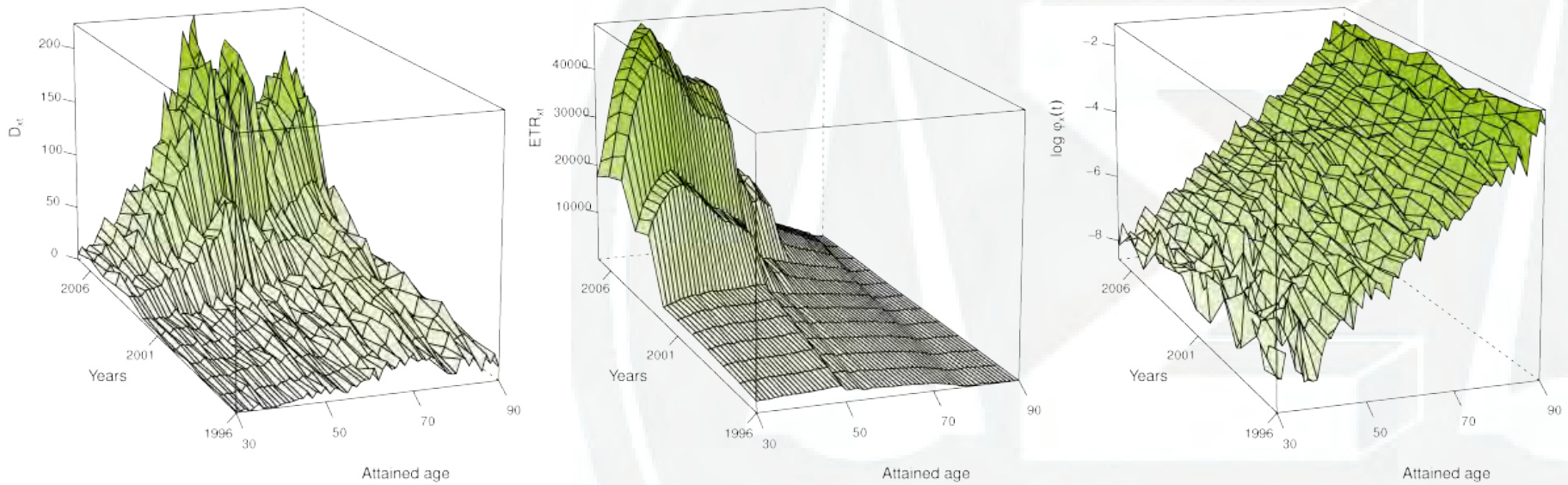
Functional principal components analysis

- The approach differs by the use of the functional data paradigm, see Ramsay and Silverman (2005) and Ferraty and Vieu (2006) for an introduction to functional nonparametric statistics.
- The decomposition using an orthonormal basis (step **ii.**) is obtained via functional principal components analysis developed by Ramsay and Dalzell (1991)
- This basis set provides informative interpretation and coefficients $\{\beta_{t,k}\}$ that are uncorrelated.
- The number K of basis functions depends on many considerations :
  - On the number of discrete points m in the original data, whether some level of smoothing is imposed by using $K < m$.
  - On the efficiency of the basis functions in reproducing the behavior of the original functions.

Contents

IAA LIFE SECTION
SECTION VIE DE L'AAI

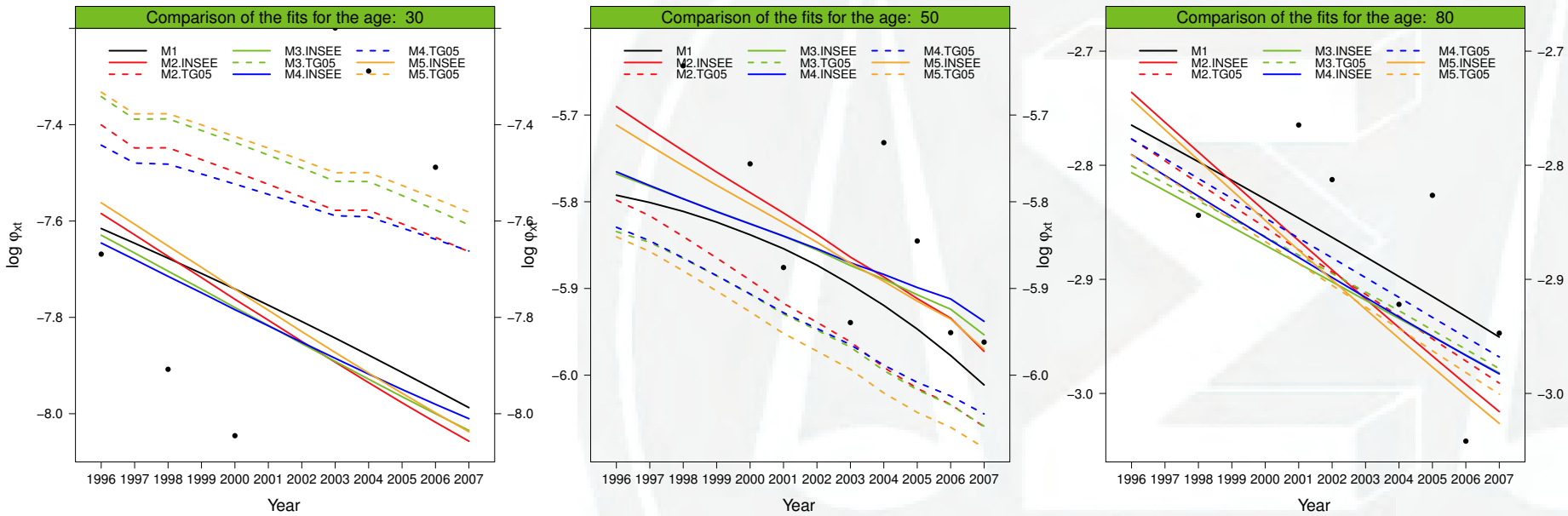
The data



(a) $D_{x,t}$ (b) $E_{x,t}$ (c) $\log \varphi_x(t)$

FIGURE: Observed surfaces of the aggregated datasets, male population.

Graphical comparison of the fits (step i.)



(a) Attained age 30 (b) Attained age 50 (c) Attained age 80

FIGURE: Comparisons of the Fits for several attained ages, log scale, step i.

Tests and quantities to compare graduations (step i.)

		M1	M2.INSEE	M2.TG05	M3.INSEE	M3.TG05	M4.INSEE	M4.TG05	M5.INSEE	M5.TG05
Fitted DF		7.56	4.02	3.94	4.88	4.88	4	4	NA	NA
Deviance		1302.53	1302.83	1328.04	1296.58	1351.66	1313.93	1366.80	1355.92	1417.08
Standardised residuals	> 2 > 3	86 24	84 22	92 24	86 23	93 24	87 23	94 26	86 26	100 31
Signs test	$+(-)$ p-value	327(405) 0.0043	319(413) 0.0005	336(396) 0.0291	345(387) 0.1296	343(389) 0.0962	334(398) 0.0198	337(395) 0.0350	333(399) 0.0162	368(364) 0.9117
Runs test	Nb of runs Value p-value	346 -1.55 0.1188	326 -3.00 0.0026	322 -3.46 $5.28e-4$	328 -3.17 0.0014	302 -5.02 $5.25e-7$	332 -2.77 0.0056	306 -4.82 $1.46e-6$	334 -2.61 0.0089	319 -3.91 $8.86e-5$
Kolmogorov Smirnov test	Value p-value	0.0327 0.8262	0.0286 0.9239	0.0601 0.1419	.0286 0.9239	0.0642 0.0978	0.0286 0.9239	0.0614 0.1257	0.0300 0.8955	0.0683 0.0657
χ^2		1400.19	1402.53	1418.89	1405.81	1445.53	1421.85	1466.93	1473.63	1545.42
R^2		0.9326	0.9256	0.9325	0.9312	0.9325	0.9302	0.9306	0.9221	0.9306
MAPE (%)		25.86	26.81	26.41	25.86	26.69	26.01	26.51	26.84	26.63
SMR $-$ 1 (%)		-0.79	0.29	0.21	0.11	0.30	$1.22e-12$	$-7.77e-13$	1.87	2.87

TABLE: Comparisons between the smoothing approaches, step i.

IAA LIFE SECTION
SECTION VIE DE L'AAI

Average log-mortality at attained ages, $\mu(x)$

Decompose the smoothed surfaces via a basis function expansion using the model (1) :

$$y_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k}\, \phi_k(x) + \varepsilon_t(x),$$

where $\varepsilon_t(x) \sim \mathcal{N}(0, \upsilon(x))$.

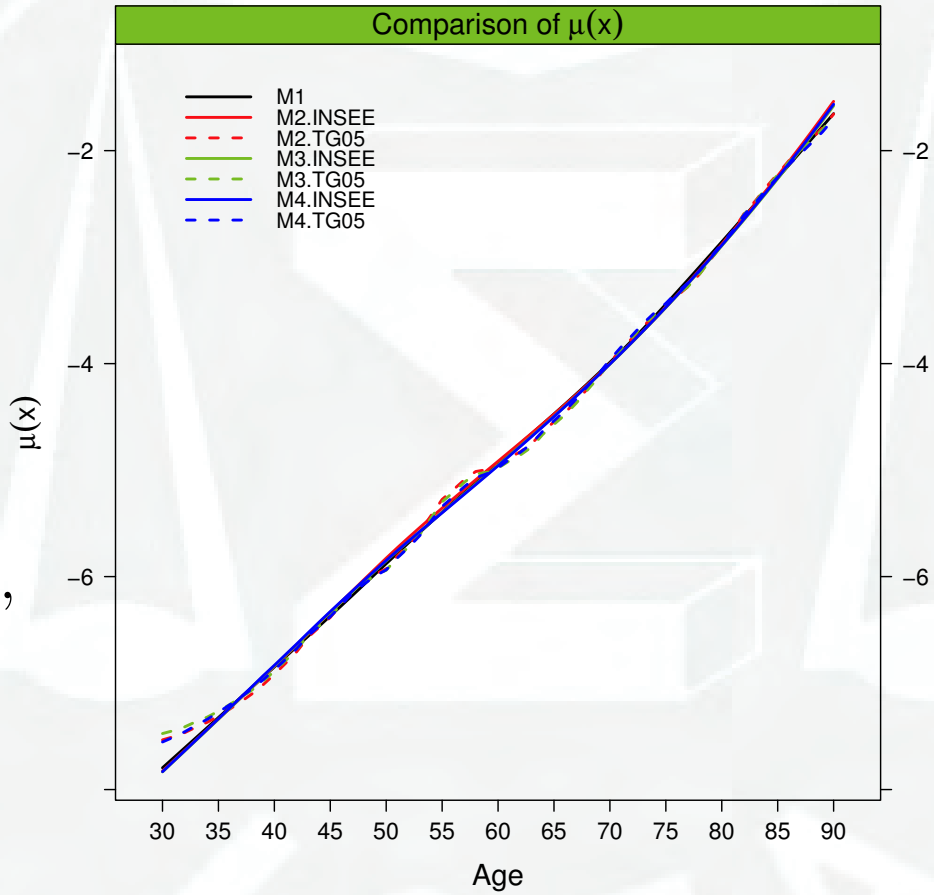
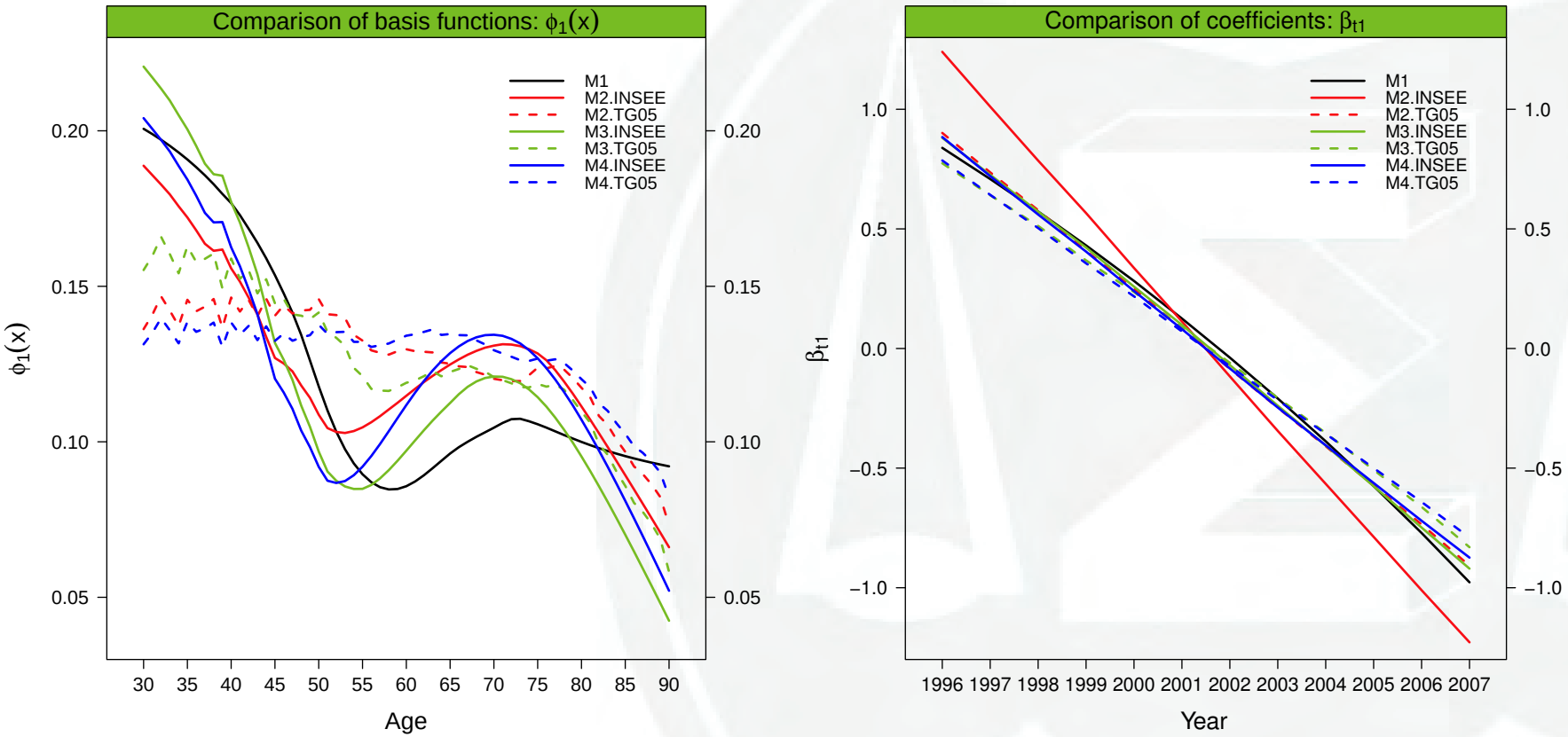


FIGURE: Average log-mortality $\mu(x)$, with $K = 3$, step ii.

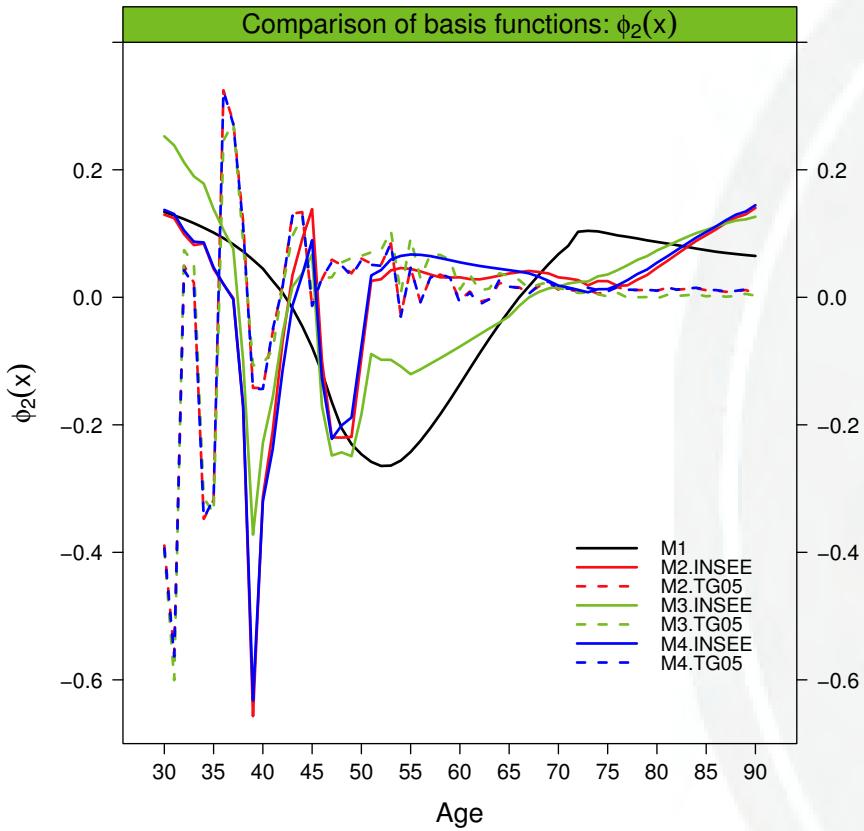
First basis function $\phi_1(x)$ and $\beta(t,1)$



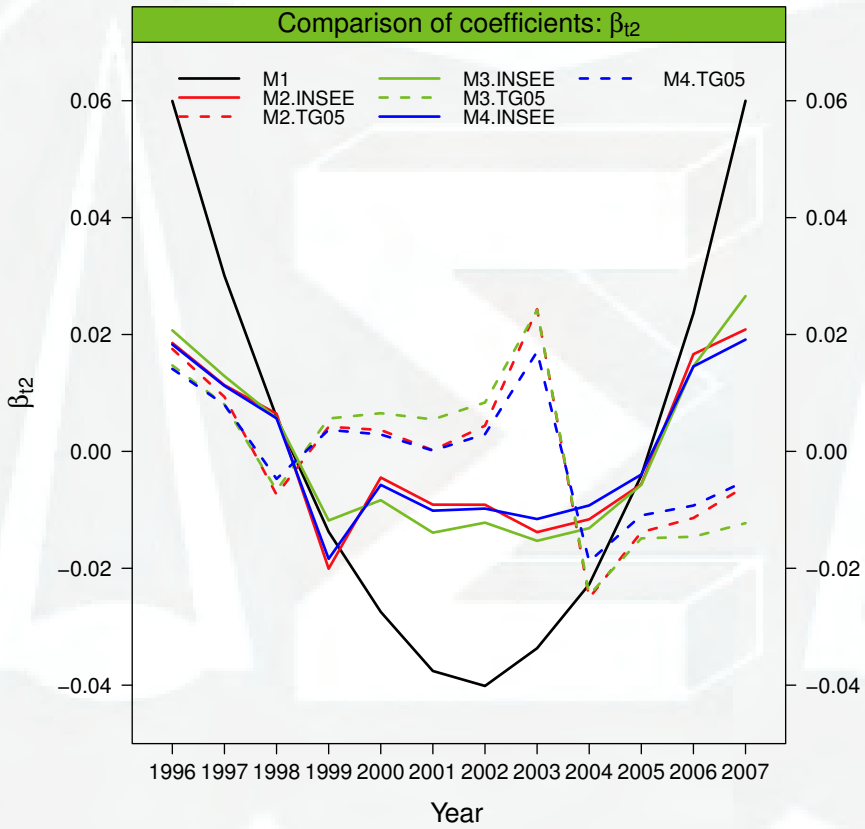
(a) $\phi_1(x)$ (b) $\beta(t,1)$

FIGURE: 1st Basis functions and associated coefficients with $K = 3$, step ii.

Second basis function $\phi_2(x)$ and $\beta(t,2)$



(a) $\phi_2(x)$

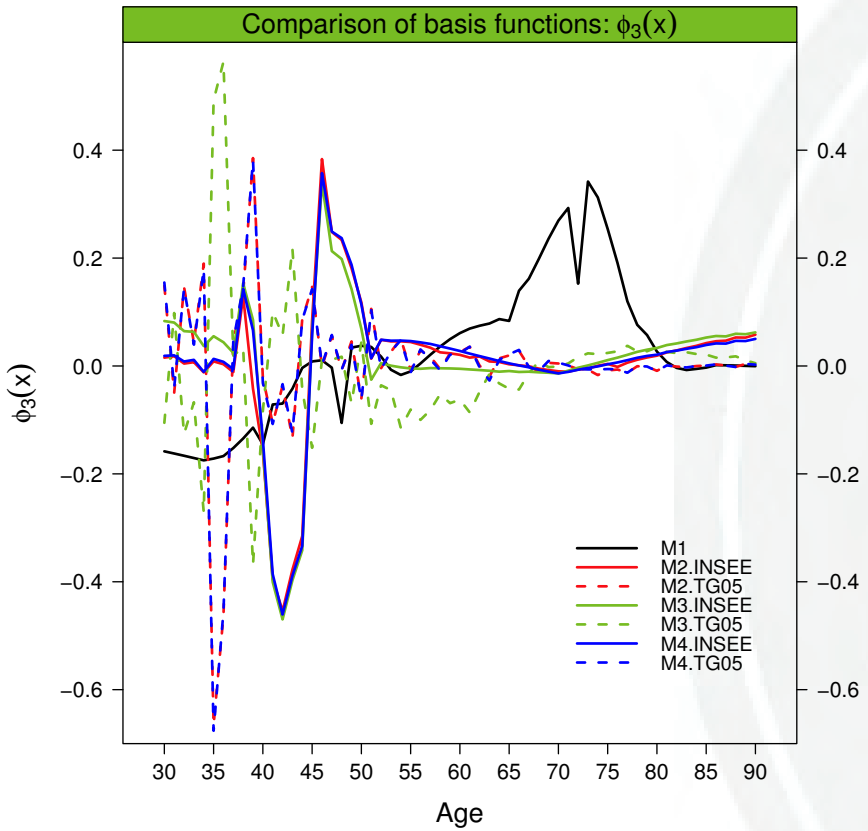


(b) $\beta(t,2)$

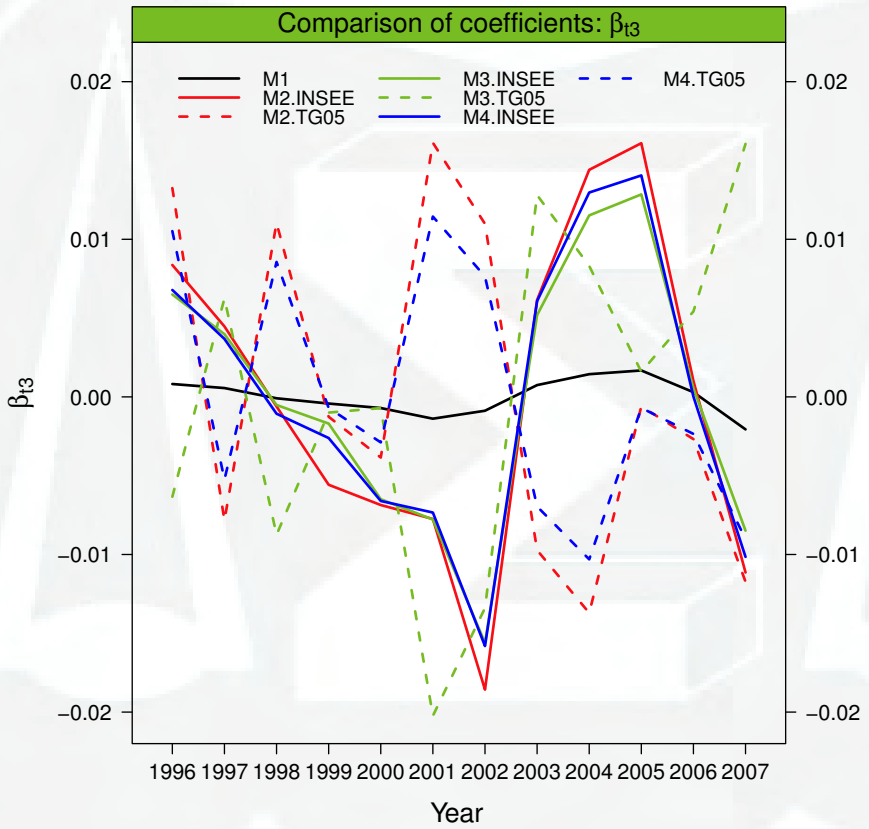
FIGURE: 2nd Basis functions and associated coefficients with $K = 3$, step ii.

IAA LIFE SECTION
SECTION VIE DE L'AAI

Third basis function $\phi_3(x)$ and $\beta(t,3)$



(a) $\phi_3(x)$



(b) $\beta(t,3)$

FIGURE: 3rd Basis functions and associated coefficients with $K = 3$, step ii.

IAA LIFE SECTION
SECTION VIE DE L'AAI

Description of the models for the time-varying coefficients (step iii.)

- Full range of ARIMA(p,d,q) with $d = 0, 1, 2$ and $p, q = 0, 1, 2, 3, 4$.
- Parameters p, d and q, selected on the basis of the BIC criterion.

Model & component		Model for the $\beta_{t,k}$	
M1	$k = 1$	ARIMA(1,2,1)	$\beta_t = 2\,\beta_{t-1} - \beta_{t-2} + \mu + \phi(\beta_{t-1} - 2\,\beta_{t-2} + \beta_{t-3} - \mu) + Z_t + \theta\,Z_{t-1}$
M1	$k = 2$	ARIMA(2,0,0) with zero mean	$\beta_t = \mu + \phi_1(\beta_{t-1} - \mu) + \phi_2(\beta_{t-2} - \mu) + Z_t$
M1	$k = 3$	ARIMA(2,0,1) with zero mean	$\beta_t = \mu + \phi_1(\beta_{t-1} - \mu) + \phi_2(\beta_{t-2} - \mu) + Z_t + \theta\,Z_{t-1}$
M2.INSEE & M4.INSEE	$k = 1$	ARIMA(0,1,0) with drift	$\beta_t = \beta_{t-1} + d + Z_t$
M2.INSEE & M4.INSEE	$k = 2$	ARIMA(1,0,0) with zero mean	$\beta_t = \mu\phi(\beta_{t-1} - \mu) + Z_t$
M2.INSEE & M4.INSEE	$k = 3$	ARIMA(0,0,1) with zero mean	$\beta_t = Z_t + \theta\,Z_{t-1}$
M2.TG05 & M4.TG05	$k = 1$	ARIMA(1,1,0) with drift	$\beta_t = \beta_{t-1} + \mu + d + \phi(\beta_{t-1} - \beta_{t-2} - \mu) + Z_t$
M2.TG05 & M4.TG05	$k = 2, 3$	ARIMA(0,0,0) with zero mean	$\beta_t = Z_t$
M3.INSEE	$k = 1$	ARIMA(0,2,0)	$\beta_t = 2\,\beta_{t-1} - \beta_{t-2} + Z_t$
M3.INSEE	$k = 2$	ARIMA(2,0,0) with zero mean	$\beta_t = \mu + \phi_1(\beta_{t-1} - \mu) + \phi_2(\beta_{t-2} - \mu) + Z_t$
M3.INSEE	$k = 3$	ARIMA(0,0,2) with zero mean	$\beta_t = Z_t + \theta_1\,Z_{t-1} + \theta_2\,Z_{t-2}$
M3.TG05	$k = 1$	ARIMA(1,2,0)	$\beta_t = 2\,\beta_{t-1} - \beta_{t-2} + \mu + \phi(\beta_{t-1} - 2\,\beta_{t-2} + \beta_{t-3} - \mu) + Z_t$
M3.TG05	$k = 2, 3$	ARIMA(0,0,0) with zero mean	$\beta_t = Z_t$

TABLE: Description of the models for the time-varying coefficients, step iii.

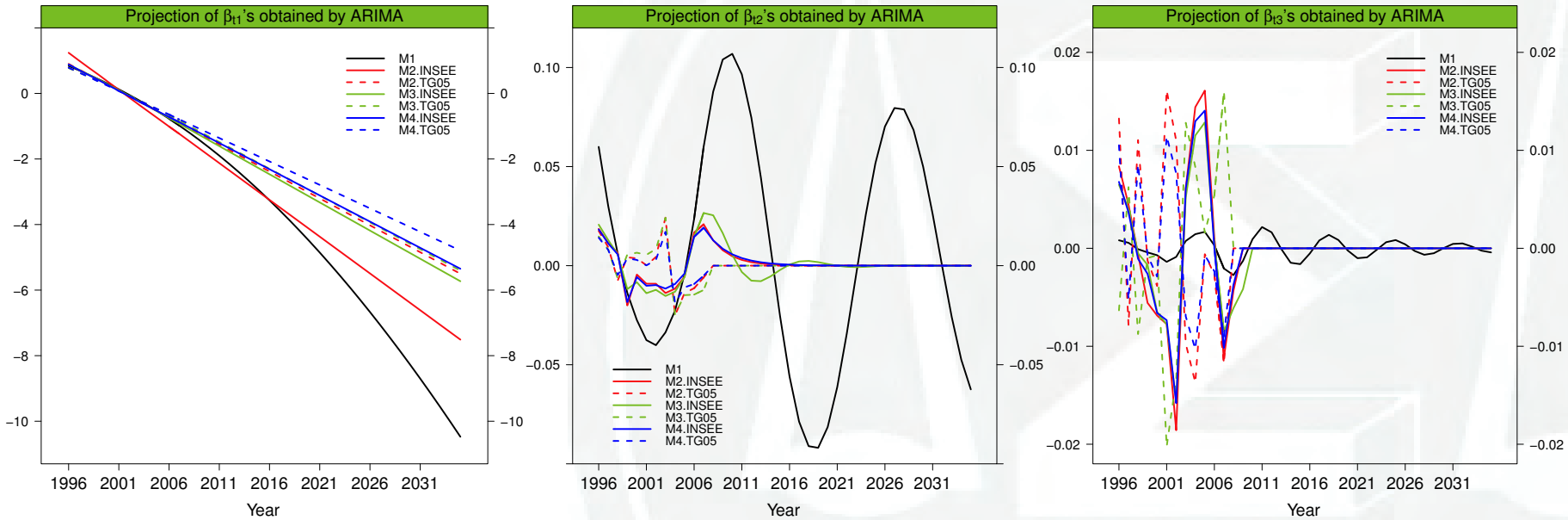
Description of the models for the time-varying coefficients (step iii.)

- Full range of ARIMA(p,d,q) with $d = 0, 1, 2$ and $p, q = 0, 1, 2, 3, 4$.
- Parameters p, d and q, selected on the basis of the BIC criterion.

Model & component		Model for the $\beta_{t,k}$	
M1	$k = 1$	ARIMA(1,2,1)	$\beta_t = 2\,\beta_{t-1} - \beta_{t-2} + \mu + \phi(\beta_{t-1} - 2\,\beta_{t-2} + \beta_{t-3} - \mu) + Z_t + \theta\,Z_{t-1}$
M1	$k = 2$	ARIMA(2,0,0) with zero mean	$\beta_t = \mu + \phi_1(\beta_{t-1} - \mu) + \phi_2(\beta_{t-2} - \mu) + Z_t$
M1	$k = 3$	ARIMA(2,0,1) with zero mean	$\beta_t = \mu + \phi_1(\beta_{t-1} - \mu) + \phi_2(\beta_{t-2} - \mu) + Z_t + \theta\,Z_{t-1}$
M2.INSEE & M4.INSEE	$k = 1$	ARIMA(0,1,0) with drift	$\beta_t = \beta_{t-1} + d + Z_t$
M2.INSEE & M4.INSEE	$k = 2$	ARIMA(1,0,0) with zero mean	$\beta_t = \mu\phi(\beta_{t-1} - \mu) + Z_t$
M2.INSEE & M4.INSEE	$k = 3$	ARIMA(0,0,1) with zero mean	$\beta_t = Z_t + \theta\,Z_{t-1}$
M2.TG05 & M4.TG05	$k = 1$	ARIMA(1,1,0) with drift	$\beta_t = \beta_{t-1} + \mu + d + \phi(\beta_{t-1} - \beta_{t-2} - \mu) + Z_t$
M2.TG05 & M4.TG05	$k = 2, 3$	ARIMA(0,0,0) with zero mean	$\beta_t = Z_t$
M3.INSEE	$k = 1$	ARIMA(0,2,0)	$\beta_t = 2\,\beta_{t-1} - \beta_{t-2} + Z_t$
M3.INSEE	$k = 2$	ARIMA(2,0,0) with zero mean	$\beta_t = \mu + \phi_1(\beta_{t-1} - \mu) + \phi_2(\beta_{t-2} - \mu) + Z_t$
M3.INSEE	$k = 3$	ARIMA(0,0,2) with zero mean	$\beta_t = Z_t + \theta_1\,Z_{t-1} + \theta_2\,Z_{t-2}$
M3.TG05	$k = 1$	ARIMA(1,2,0)	$\beta_t = 2\,\beta_{t-1} - \beta_{t-2} + \mu + \phi(\beta_{t-1} - 2\,\beta_{t-2} + \beta_{t-3} - \mu) + Z_t$
M3.TG05	$k = 2, 3$	ARIMA(0,0,0) with zero mean	$\beta_t = Z_t$

TABLE: Description of the models for the time-varying coefficients, step iii.

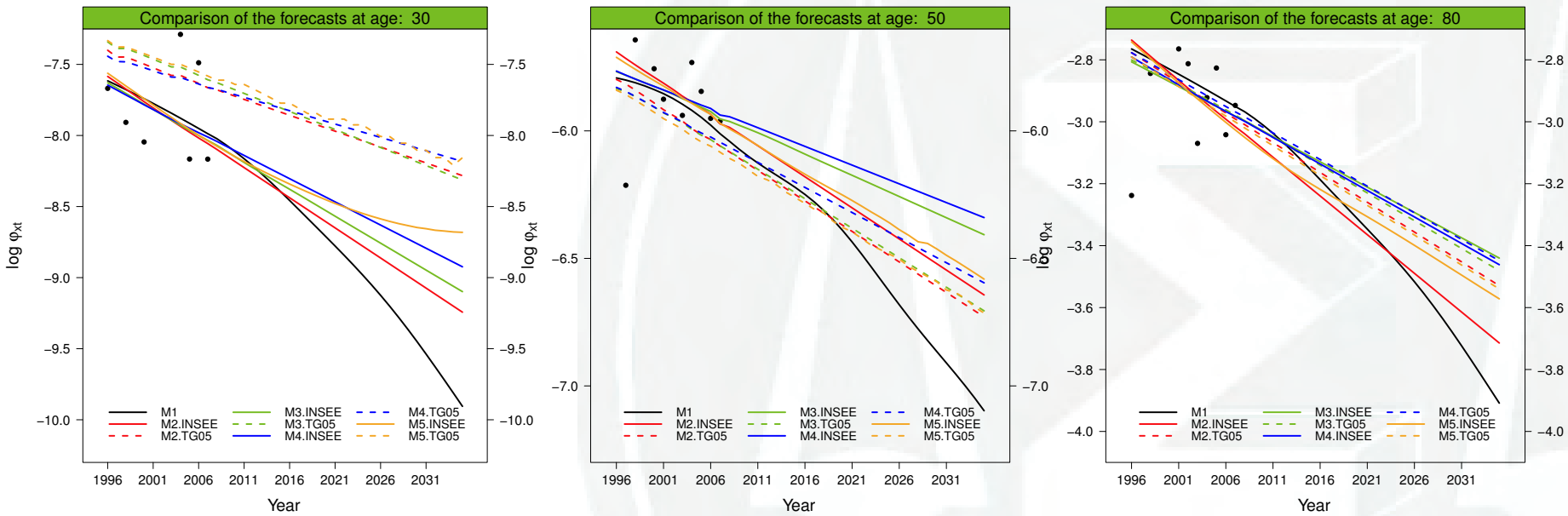
Projections of the time-varying coefficients (step iv.)



(a) $\tilde{\beta}_{m,h,1}$. (b) $\tilde{\beta}_{m,h,2}$. (c) $\tilde{\beta}_{m,h,3}$.

FIGURE: Projections of the estimated coefficients $\beta_{t,k}$, step iv.

Graphical comparison of the fits (step v.)



(a) Attained age 30 (b) Attained age 50 (c) Attained age 80

FIGURE: Comparisons of the fits and forecasts for several attained ages, log scale, step v.

Contents

IAA LIFE SECTION
SECTION VIE DE L'AAI

Conclusion

- We have presented a new robust and strictly non-parametric framework to smooth and forecast mortality.
- We used local kernel-weighted log-likelihood techniques to graduate the observed mortality.
- The extrapolation is then performed by identifying the mortality components and their importance over time using functional principal components analysis. Time series methods are used to extrapolate the time-varying coefficients.
- We investigated the divergences in the mortality surfaces generated by a number of proposed models.
- Selecting models purely on the basis of how well they fit historical data is dangerous, because the model may lead to a good fit to the historical data, and still give inadequate forecasts.

Conclusion

- We have presented a new robust and strictly non-parametric framework to smooth and forecast mortality.
- We used local kernel-weighted log-likelihood techniques to graduate the observed mortality.
- The extrapolation is then performed by identifying the mortality components and their importance over time using functional principal components analysis. Time series methods are used to extrapolate the time-varying coefficients.
- We investigated the divergences in the mortality surfaces generated by a number of proposed models.
- Selecting models purely on the basis of how well they fit historical data is dangerous, because the model may lead to a good fit to the historical data, and still give inadequate forecasts.

IAA LIFE SECTION
SECTION VIE DE L'AAI

Conclusion

- We have presented a new robust and strictly non-parametric framework to smooth and forecast mortality.
- We used local kernel-weighted log-likelihood techniques to graduate the observed mortality.
- The extrapolation is then performed by identifying the mortality components and their importance over time using functional principal components analysis. Time series methods are used to extrapolate the time-varying coefficients.
- We investigated the divergences in the mortality surfaces generated by a number of proposed models.
- Selecting models purely on the basis of how well they fit historical data is dangerous, because the model may lead to a good fit to the historical data, and still give inadequate forecasts.

Conclusion

- We have presented a new robust and strictly non-parametric framework to smooth and forecast mortality.
- We used local kernel-weighted log-likelihood techniques to graduate the observed mortality.
- The extrapolation is then performed by identifying the mortality components and their importance over time using functional principal components analysis. Time series methods are used to extrapolate the time-varying coefficients.
- We investigated the divergences in the mortality surfaces generated by a number of proposed models.
- Selecting models purely on the basis of how well they fit historical data is dangerous, because the model may lead to a good fit to the historical data, and still give inadequate forecasts.

IAA LIFE SECTION
SECTION VIE DE L'AAI

Conclusion

- We have presented a new robust and strictly non-parametric framework to smooth and forecast mortality.
- We used local kernel-weighted log-likelihood techniques to graduate the observed mortality.
- The extrapolation is then performed by identifying the mortality components and their importance over time using functional principal components analysis. Time series methods are used to extrapolate the time-varying coefficients.
- We investigated the divergences in the mortality surfaces generated by a number of proposed models.
- Selecting models purely on the basis of how well they fit historical data is dangerous, because the model may lead to a good fit to the historical data, and still give inadequate forecasts.

References I

Bongaarts, J. (2005). Long-range trends in adult mortality : Models and projection methods. *Demography*, **42**(1), 23–49.

Brouhns, N., Denuit, M., and Vermunt, J. K. (2002a). Measuring the longevity risk in mortality projections. *Bulletin of the Swiss Association of Actuaries*, **2**, 105–130.

Brouhns, N., Denuit, M., and Vermunt, J. K. (2002b). A poisson log-bilinear regression approach to the construction of projected lifetables. *Insurance : Mathematics & Economics*, **31**, 373–393.

Cairns, A. J. G., Blake, D., and Dowd, K. (2006). A two-factor model for stochastic mortality with parameter uncertainty : theory and calibration. *The Journal of Risk and Insurance*, **73**(4), 687–718.

Cosette, H., Delwarde, A., Denuit, M., Guillot, F., and Marceau, E. (2007). Pension plan valuation and dynamic mortality tables. *North American Actuarial Journal*, **11**(2), 1–34.

Debón, A., Montes, F., and Sala, R. (2006). A comparison of nonparametric methods in the graduation of mortality : Application to data from the Valencia region (Spain). *International statistical Review*, **74**(2), 215–233.

IAA LIFE SECTION
SECTION VIE DE L'AAI

References II

Delwarde, A., Kachkhdze, D., Olie, L., and Denuit, M. (2004). Modèles linéaires et additifs géneralisés, maximum de vraisemblance local et méthodes relationelles en assurance sur la vie. *Bulletin Français d'Actuariat*, **6**(12), 77–102.

Delwarde, A., Denuit, M., and Partrat, C. (2007). Negative binomial version of the Lee-Carter model for mortality forecasting. *Applied Stochastic Models in Buisiness and Industry*, **23**(5), 385–401.

Fan, J., Farmen, M., and Gijbels, I. (1998). Local maximum likelihood estimation and inference. *Journal of the Royal Statistical Society*, **60**(3), 591–608.

Ferraty, F. and Vieu, P. (2006). *Nonparametric funcitonal data analysis - Theory and Practice*. Series in Statistics. Springer-Verlag New-York, Inc.

Gschlössl, S., Schoenmaekers, P., and Denuit, M. (2011). Risk classification in life insurance : methodology and case study. *European Actuarial Journal*, **1**(1), 23–41.

Hyndman, R. J. and Booth, H. (2008). Stochastic population forecasts using functional data models for mortality, fertility and migration. *International Journal of Forecasting*, **24**(3), 323–342.

IAA LIFE SECTION
SECTION VIE DE L'AAI

References III

Hyndman, R. J. and Ullah, M. (2007). Robust forecasting or mortality and fertility rates : A functional data approach. *Computational Statistics & Data Analysis*, **51**(10), 4942–4956.

Lee, R. D. and Carter, L. R. (1992). Modeling and forecasting U.S. mortality. *Journal of the American Statistical Association*, **87**(419), 659–671.

Loader, C. R. (1996). Local likelihood density estimation. *The Annals of Statistics*, **24**(4), 1602–1618.

Loader, C. R. (1999). *Local Regression and Likelihood*. Statistics and Computing Series. New York : Springer Verlag.

Ramsay, J. O. and Dalzell, C. J. (1991). Some tools for functional data analysis. *Journal of the Royal Statistical Society*, **53**(3), 539–572.

Ramsay, J. O. and Silverman, B. W. (2005). *Functional data analysis*. Springer-Verlag New-York, Inc., second edition.

Renshaw, A. E. and Haberman, S. (2003). Lee-Carter mortality forecasting with age-specific enhancement. *Insurance : Mathematics & Economics*, **33**(2), 255–272.

IAA LIFE SECTION
SECTION VIE DE L'AAI

References IV

Tomas, J. (2011). A local likelihood approach to univariate graduation of mortality. *Bulletin Français d'Actuariat*, **11**(22), 105–153.

Tomas, J. (2013). *Quantifying biometric life insurance risks with non-parametric methods*. Ph.D. thesis, Amsterdam School of Economics Research Institute.

Tomas, J. and Planchet, F. (2013a). Multidimensional smoothing by adaptive local kernel-weighted log-likelihood with application to long-term care insurance. *Insurance : Mathematics & Economics*, **52**(3), 573–589.

Tomas, J. and Planchet, F. (2013b). A new non-parametric approach to smoothing and forescating of mortality. *Cahiers de Recherche de l'ISFA*, **2013**(5), 1–28.