Two-dimensional smoothing of mortality surfaces with cohort and period ridges

Alexander Dokumentov, Rob J Hyndman, Leonie Tickle

Abstract

BACKGROUND

Mortality rates typically move smoothly over age and time. Exceptions occur, due to events such as wars and epidemics, which create, among other features, ridges in the mortality surface in a particular calendar year or for cohorts born in a particular year.

OBJECTIVES

We aim to develop and evaluate new methods that better model the smooth underlying age-period mortality surface and any cohort or period ridges.

METHODS

We propose two new practical methods for modelling the age-period surface of the logarithms of mortality rates. The first uses an approach similar to bivariate thin plate splines although with L_1 regularization. The second, which is our recommended method, also uses bivariate L_1 regularization but allows for smooth age-varying period and cohort effects.

RESULTS

Cross validation is used to compare these new methods with existing approaches. Evaluations on a multi-country dataset using four different age and period ranges as well as the simulation results indicate that the recommended method consistently gives a better estimate of the “true” age-period mortality surface.

CONCLUSIONS

Explicit modelling of cohort and period ridges in the mortality surface, using of L_1 norm for regularization and measuring errors as well as using two dimensions for smoothing, all improves accuracy and provides greater insight into the underlying mortality dynamics.

CONTRIBUTION

The new methods provide more accurate and nuanced models of the age-period mortality surface, informing mortality modelling, analysis and forecasting. Although designed for the modelling of mortality rates they can be applied to any bivariate data with occasional ridges and extend the statistical literature on quantile smoothing.

Keywords: Bivariate data, nonparametric smoothing, graduation, cohort effects, period effects.

JEL codes: J110

1 Introduction

Mortality rates are used to compute life tables, life expectancies, insurance premiums and reserves, and other items of interest to demographers and actuaries. However, observed mortality rates are noisy, and it is useful to smooth or "graduate" them in order to obtain estimates of the underlying rates that are more accurate and less variable. Exceptions to the underlying smooth mortality surface can occur due to features such as wars and epidemics, which manifest as either period effects (different mortality rates in a particular calendar year) or cohort effects (different mortality rates among those born in a particular year). These appear as ridges in the otherwise smooth mortality surface, and any effective smoothing methods applied to mortality data need to allow for such features.

Figure 1 shows log mortality rates $m_{x,t} = \log(M_{x,t})$ for females in France from 1950 to 1970. The observed mortality rate at age x in year t is $M_{x,t} = D_{x,t}/E_{x,t}$, where $D_{x,t}$ is the number of deaths aged x in year t, and $E_{x,t}$ is the total number of years lived at age x during year t, which can be approximated by the mid-year population. The data is sourced from the *Human Mortality Database* (2008).

The following features of the log mortality surface are evident:

- In the age dimension, the log mortality decreases rapidly at the early ages and reaches a minimum at about age 10. There is an "accident hump" around age 20 in some years (Heligman and Pollard 1980), after which log mortality increases almost linearly to the very old ages.
- At almost every age, mortality decreases over time during the period, more steeply (on the log scale) for younger ages than for older ages.
- There are diagonal ridges (along $t = x + k$ for a specific k) due to cohort effects (these are somewhat difficult to discern in Figure 1 but are analyzed in depth in later sections). Studies have indicated a relationship between mortality and

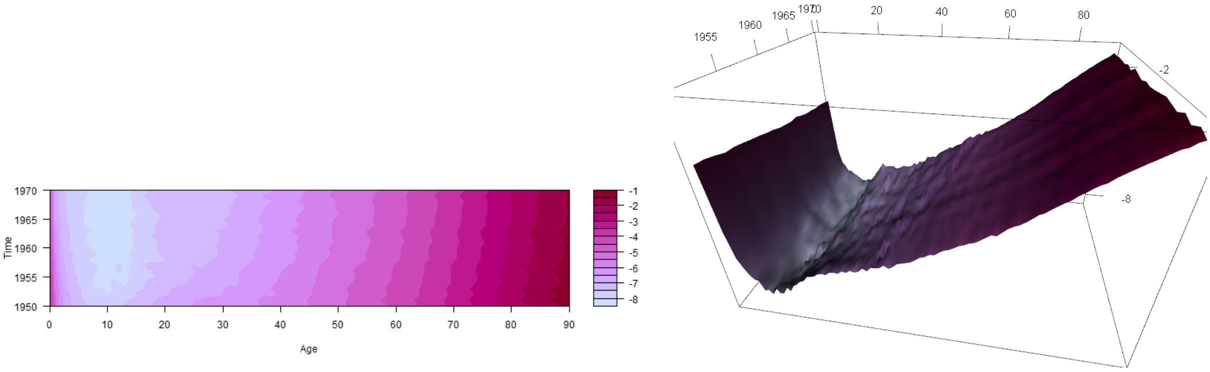

Figure 1: *Natural logarithms of French female mortality rates (years 1950-1970).*

year of birth in numerous countries including France and the UK (Andreev and Vaupel 2005; Janssen and Kunst 2005; O'Connell and Dunstan 2009; Willets 2004), particularly for males (Richards 2008). Cohort differences in smoking prevalence as well as pre-natal and early life factors are thought to play a role (Janssen and Kunst 2005; Richards, Kirkby and Currie 2006).

- There are horizontal ridges (for a fixed calendar year t) due to period effects. Such patterns are usually due to extreme environmental events such as wars and pandemics, which affect all people (with different magnitude) during a particular year. Figure 2, covering the period 1935 to 1955, illustrates marked period effects due to World War II. Major period effects in other periods are due to the Spanish influenza epidemic of 1918, which resulted in worldwide deaths of the order of 50 million (Johnson and Mueller 2002), and to World War I. Less extreme period effects are also evident in other years.

Ideally, any model of the mortality surface should appropriately distinguish and model all components: the smooth underlying bivariate age and time mortality surface; additional smooth functions of age for some cohorts k representing significant cohort effects; additional smooth functions of age in some calendar years t representing significant period effects; and random errors. Mortality analysis or forecasting may then proceed with a more comprehensive and nuanced understanding of the relevant patterns and their underlying causes.

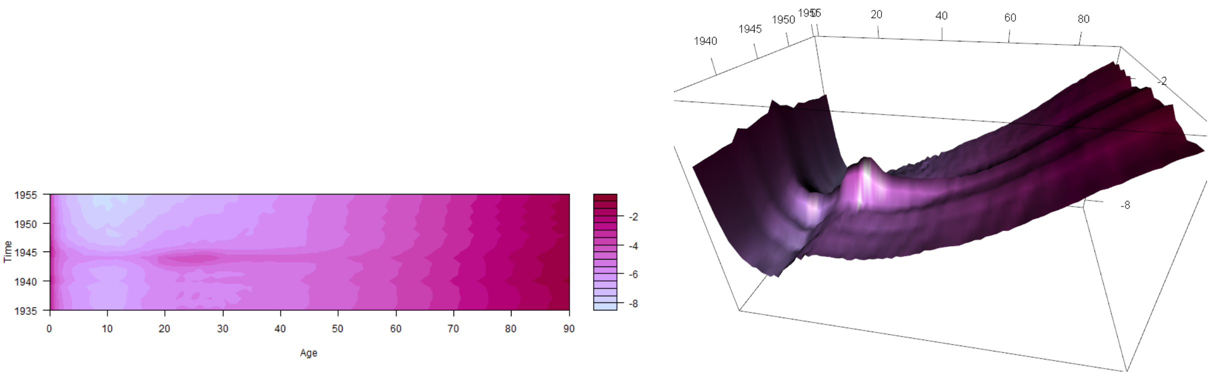

Figure 2: *Natural logarithms of French female mortality rates (years 1935-1955).*

Existing smoothing methods address this problem to some extent, but not completely.

A number of methods that have been developed are not tailored to the mortality surface. Early work in functional data analysis and nonparametric smoothing tended to involve general methods that could be applied to a wide range of problems (see, for example Ferraty and Vieu 2006; Horváth and Kokoszka 2012; Silverman and Ramsay 2005). More recently, the trend in functional data research has been to develop or adapt methods that take account of the specific features of the problem at hand – such as jump discontinuities in neural images (Zhu, Fan and Kong 2014) and amplitude and phase variation in nuclear magnetic resonance spectral data (Marron et al. 2015) – or that are applied to new types of data including trees (Shen et al. 2014) and graphs (Zhu, Strawn and Dunson 2014). Ideas from functional data analysis are also being applied in other areas of mathematics and statistics including delay differential equations (Brunel, Clairon and d'Alché-Buc 2014) and automated variable selection in functional linear models (Gertheiss, Maity and Staicu 2013). The methods that will be developed in this paper contribute to this broader literature and address particular issues that arise in the modelling of the mortality surface.

Other smoothing methods have been proposed that are tailored to the mortality surface but that smooth only in the age, not time, direction. Examples include the nonparametric functional data analysis approach of Hyndman and Ullah (2007) which finds a smooth function $f_t(x) = \mathrm{E}[m_{x,t}]$ separately for each t, the Whittaker-Henderson method and generalizations (Henderson 1924; Schuette 1978; Whittaker 1922), and the many approaches that fit parametric functions to mortality rates across age in a particular year (Forfar, McCutcheon and Wilkie 1988), including Heligman and Pollard (1980), Carriere (1992) and Hannerz (2001), among others. However, in the absence of wars and other extreme events, it is reasonable to expect that mortality rates will progress smoothly over time as well as age and can therefore be modelled using a smooth bivariate function $f(x,t) = \mathrm{E}[m_{x,t}]$. By allowing the assumption of smoothness in both dimensions, better performance should be possible due to the additional information included in the estimation. Hyndman and Ullah (2007) is among the methods evaluated in this paper, to provide a univariate comparison to the new proposed bivariate methods.

Methods exist that are tailored to the mortality surface and that smooth in both directions. Currie, Durban and Eilers (2004) use two-dimensional P-splines for smoothing and forecasting mortality rates, and Camarda (2012) implement this approach along with a one-dimensional version in the R package MortalitySmooth. However these methods, by using a global smoother across the whole age range, can fail to deal with the steep drop in mortality between birth and age 10 (Camarda 2012). Camarda, Eilers and Gampe (2012) use special bases for P-splines to better model the steep decline in mortality at the youngest ages for the univariate case, and Camarda, Eilers and Gampe 2010 for both the univariate and bivariate cases. Other approaches to bivariate smoothing of mortality rates include regularized singular value decomposition (Huang, Shen and Buja 2009) and repeated functional observations (Chen and Müller 2012).

Another issue with many bivariate smoothing approaches is potential sensitivity to single outlier observations due to the use of *squared* measures (i.e. an L_2 norm) rather than *absolute values* (i.e. an L_1 norm) in the estimation process. Bivariate smoothing estimation involves minimization of a measure incorporating both lack of fit and lack of smoothness, and the same or a different norm may be applied to each of these elements. When applied to the smoothness element, the L_2 norm penalizes large changes in the smoothed function much more heavily than smaller and more gradual changes and therefore is more likely to over-smooth abrupt features in the data. The L_1 norm does not possess such a feature; the "cost" of a single large change is exactly the same as sum of the "costs" of a number of smaller changes with the same combined magnitude. Likewise, when applied to the fit element, the L_2 norm will penalize large single errors more heavily and hence will be more sensitive to outliers than the L_1 norm. Most existing bivariate smoothing approaches for mortality use the L_2 norm, although the L_1 norm has been found in the univariate case to improve robustness when the data contains outliers (Portnoy 1997; Schuette 1978).

Finally, most bivariate mortality smoothing methods fail to preserve and explicitly model cohort and period effects, instead removing them from the estimated mortality surface. Kirkby and Currie (2010) extend the approach of Currie, Durban and Eilers (2004) to model period (but not cohort) effects: the method is based on a Poisson model of deaths using a GLM for estimation, with the period effects estimated in a multi-step procedure. Barbi and Camarda (2011) extend Currie, Durban and Eilers (2004) to incorporate age-independent period and cohort effects. However, they acknowledge that use of age-independent effects may be too simplistic, and they call for the development of more sophisticated models that incorporate age-varying cohort and period effects.

We propose two new practical bivariate mortality smoothing methods that address shortcomings of existing approaches, and comprehensively evaluate these alongside two existing comparison methods. The second of the proposed methods – which is our preferred approach – incorporates age-varying cohort and period effects and can also be distinguished from most existing methods in its use of the L_1 norm to emphasize preservation of features in the data and robustness to outliers.

The evaluation, which is based on 12 countries and four different age and period subsets, indicates that the proposed method is the most accurate of those considered. In addition, the estimates of cohort and period effects that it provides are of use to demographers and others in understanding, modelling and forecasting mortality. The new methods are extensions and combinations of quantile smoothing splines (see, for example, He, Ng and Portnoy 1998; Koenker, Ng and Portnoy 1994; Portnoy 1997) with partial differential regularization (Sangalli 2014). They contribute to this literature and may be applied more generally to any bivariate data with ridges.

The next four sections describe the four smoothing algorithms, the last of which is our preferred procedure:

1. The Hyndman and Ullah (2007) method (Section 2), is implemented in the demography R package (Hyndman 2014), and smooths mortality rates only in the age dimension. This algorithm is presented for comparison only.

2. The Currie, Durban and Eilers (2004) method as implemented in the Mortality Smooth package of Camarda (2012) (Section 3) smooths mortality rates in both dimensions, although it is designed only to work for ages greater than 10. This algorithm is also presented for comparison only.

3. Section 4 describes a new algorithm that uses L_1 regularization and also uses both dimensions for smoothing. This algorithm copies thin plate splines in many ways, but it uses the L_1 norm instead of the L_2 norm.

4. The last algorithm (Section 5) also uses L_1 regularization and both dimensions for smoothing, but incorporates ridges to account for cohort and period effects. This improves performance and also provides greater insight into the structure of the mortality data.

The performance of the four methods is evaluated in Section 6 using a cross validation procedure [1]. Finally, we provide discussion and conclusions in Section 7.

[1]Cross validation is a technique allowing to validate the predictive ability of a model. In this paper we consider it to be equivalent to an ability to estimate the "true" mortality surface.

2 Hyndman-Ullah (2007) method

Hyndman and Ullah (2007) propose a method for smoothing mortality rates across ages in each year. The method is intentionally one-dimensional to allow for a forecasting procedure, applied after smoothing, that takes into account variation in the time dimension. An implementation of the method is provided in the demography package for R (Hyndman 2014).

This smoothing method uses constrained weighted penalized spline regression applied independently for each year. Weighted penalized spline regression involves calculating a vector β which minimizes the expression

$$||w(y - X\beta)||^2_{L_2} + \lambda^2 \beta^T D\beta,$$

where y is a vector of observations, X is a matrix representing linear spline bases, $D = \mathrm{diag}(0, 0, 1, 1, \ldots, 1)$ is a diagonal matrix, w is a vector of weights and $\lambda > 0$ is a "smoothing" parameter where larger values of λ impose greater smoothness (see, for example, Ruppert, Wand and Carroll 2003). Note that

$$||w(y - X\beta)||^2_{L_2} = \sum_i w_i^2 (y - X\beta)_i^2$$

is a measure of lack of fit, $\beta^T D\beta$ a measure of lack of smoothness as indicated by the size of elements of β, and λ^2 a measure of the weight given to smoothness versus fit in the estimation.

In the case of smoothing log mortality rates, observations in year t are given by $y_i = m_{x_i,t}$ for age group x_i years old. The weights w_i^2 are taken as the inverse of the estimated variances of y_i. Assuming deaths follow a Poisson distribution, and using a Taylor series approach, Hyndman and Booth (2008) estimate the variance of y_i as $\sigma_i^2 \approx (E_{x_i,t} M_{x_i,t})^{-1}$, where $E_{x_i,t}$ is the mid-year population and $M_{x_i,t}$ the mortality rate, at age x_i in year t.

Moreover such splines are constrained to ensure that the resulting function $f(x)$ is monotonically increasing for $x \geq c$ for some c (for example 50 years). Hyndman and Ullah (2007) use a modified version of the method described in Wood (1994) to implement this constraint.

The result of this approach is a surface which is smooth in the age dimension but still "wiggly" in the time dimension (Figure 3).

The residuals (Figure 4) show some vertical patterns at early ages indicating serial correlation, as well as diagonal patterns indicating cohort effects (effects related to people born in the same year). For example Figure 5 reveals some serial correlation of the residuals for ages 1 and 2. However, it is clear that the residuals do not show any horizontal patterns due to period effects. This is expected, because separate smoothing has been done independently for each year.

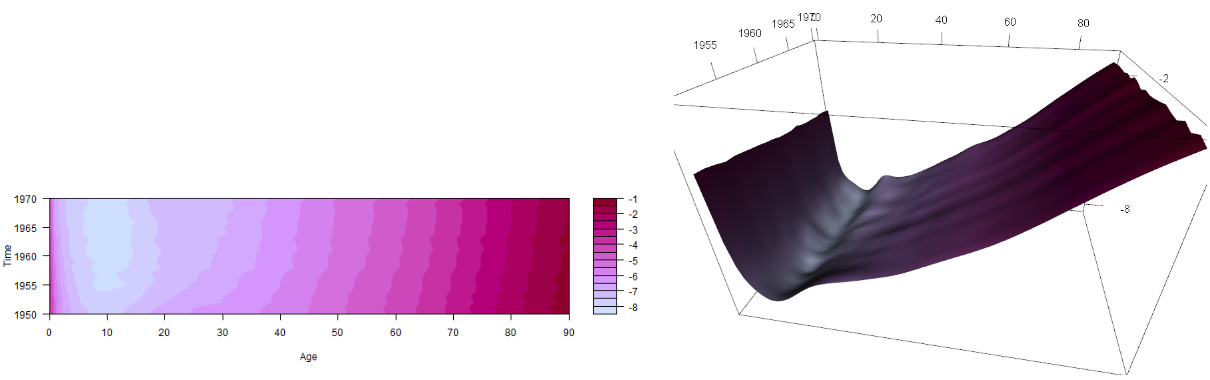

Figure 3: *French female mortality rates smoothed by Hyndman and Ullah (2007) method.*

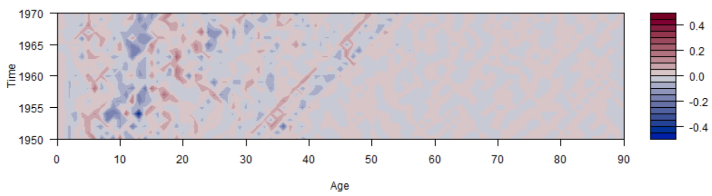

Figure 4: *French female mortality rates residuals after smoothing by Hyndman and Ullah (2007) method.*

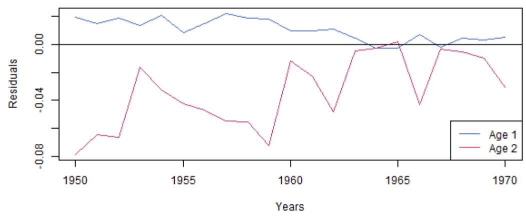

Figure 5: *French female mortality rates residuals for ages 1 and 2 after smoothing in age dimension.*

3 Currie, Durban and Eilers (2004) as implemented by Camarda (2012)

Camarda (2012) implements a two-dimensional method using P-splines for smoothing mortality rates (Currie, Durban and Eilers 2004) in the MortalitySmooth package for R. In simple terms, the method uses B-splines as a model for the bivariate log mortality surface, modified by incorporating a penalty on the regression coefficients to move them closer to zero ("shrinkage", a form of "regularization") to impose smoothness.

The resultant rates are smooth across both dimensions. For ages 0 to 10, the result of smoothing is notably biased (Figures 6 and 7). Such behavior does not depend on the number of knots used for smoothing. (By default knots are positioned as regular grid: one knot every 5 years in both dimensions. We tested up to one knot per every year, with no significant effect on results.) As we can see in Figure 7 the residuals are serially correlated for early ages. Also diagonal ridges due to cohort effects and horizontal ridges due to period effects are visible.

Camarda (2012) acknowledges that the method was not designed for smoothing of the youngest ages or data with outliers. Therefore, we do not include this method[2] in the comparisons which involve ages 0 to 10 or years with wars.

[2]A reader can perform such testing independently as the source code is provided.

4 L$_1$ Regularized Median Smoothing

A two dimensional thin plate regression spline is defined as the function $f(x,t)$ which minimizes

$$J\left(\{y_i\}_{i=1}^{n}, f\right) = \sum_{i=1}^{n}(y_i - f(x_i,t_i))^2 + \lambda \int \left[\left(\frac{\partial^2 f}{\partial x^2}\right)^2 + 2\left(\frac{\partial^2 f}{\partial x \partial t}\right)^2 + \left(\frac{\partial^2 f}{\partial t^2}\right)^2\right] \mathrm{d}x\,\mathrm{d}t \qquad (1)$$

for some smoothing parameter $\lambda > 0$, knots $\{(x_i,t_i)\}_{i=1}^{n}$ and values $\{y_i\}_{i=1}^{n}$ (see for example Wood 2006). Again, the first term measures lack of fit between the observed and smoothed log mortality rates, the second term measures lack of smoothness as indicated by second derivatives across age, time and their interaction, and λ indicates the weight given to smoothness relative to fit in the estimation. The L_2 norm (i.e. squared measures) is used for both the fit and smoothness elements.

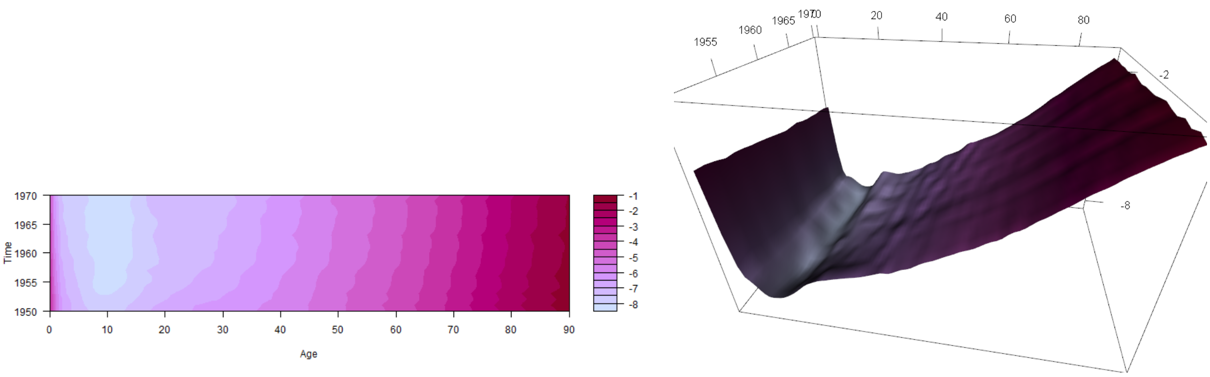

Figure 6: *French female mortality rates smoothed by Camarda (2012) method.*

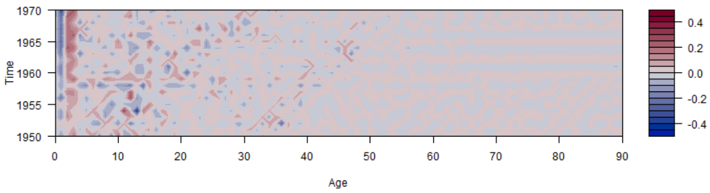

Figure 7: *French female mortality rates residuals after smoothing by Camarda (2012) method.*

Thin plate regression splines work well in many cases. However, in case of mortality data, direct application of thin plate splines (or adaptive thin plate splines where flexibility varies) to the logs of mortality rates does not lead to precise and unbiased results, especially for early ages. This is the same problem described for the method in Section 3.

We speculate that the reasons for the problems are twofold:

(a) The log mortality rate surface is very steep at early ages, and twisted along the time dimension due to a more rapid decrease in mortality for younger ages compared to older ages; and
(b) thin plate splines penalize big errors much more heavily than small errors.

This leads to the situation where abrupt jumps in the mortality data generate errors in the proximity of the jumps, causing unsatisfactory performance of thin plate splines over the abrupt surface.

In the rest of the section we show how these problems can be solved and present the resulting approach.

If the knots form a fine regular grid, then the integral in (1) can be approximated by a sum and so $J(\{y_i\}_{i=1}^n, f)$ can be approximated as

$$J\left(\{y_i\}_{i=1}^n, f\right) \approx \sum_{i=1}^n (y_i - f(x_i,t_i))^2 + \frac{\lambda}{n}\sum_{i=1}^n \left[\left(\frac{\partial^2 f}{\partial x^2}(x_i,t_i)\right)^2 + 2\left(\frac{\partial^2 f}{\partial x \partial y}(x_i,t_i)\right)^2 + \left(\frac{\partial^2 f}{\partial y^2}(x_i,t_i)\right)^2\right].$$

Also if the knots form a fine regular grid, then the second partial derivatives at knots can be approximated as linear combinations of function values at nearby knots.

Denoting $\{y_i\}_{i=1}^n$ as vector y and $\{f(x_i,y_i)\}_{i=1}^n$ as vector z, then $J(\{y_i\}_{i=1}^n, f)$ can be approximated as

$$J(y,z) \approx \|y-z\|_{L_2}^2 + \frac{\lambda}{n}\left(\|D_{xx}z\|_{L_2}^2 + 2\|D_{xt}z\|_{L_2}^2 + \|D_{tt}z\|_{L_2}^2\right)$$

where D_{xx}, D_{xt} and D_{tt} are linear operators (matrices) which calculate approximations of vectors $\left\{\frac{\partial^2 f}{\partial x^2}(x_i, t_i)\right\}_{i=1}^{n}$, $\left\{\frac{\partial^2 f}{\partial x \partial t}(x_i, t_i)\right\}_{i=1}^{n}$ and $\left\{\frac{\partial^2 f}{\partial t^2}(x_i, t_i)\right\}_{i=1}^{n}$.

Using the above expression, we can approximate a thin plate spline computed at its knots as

$$S(y) = \arg\min_{z} \left(\|y - z\|_{L_2}^2 + \frac{\lambda}{n} \left(\|D_{xx} z\|_{L_2}^2 + 2\|D_{xt} z\|_{L_2}^2 + \|D_{tt} z\|_{L_2}^2 \right) \right).$$

In the case of smoothing mortality rates, y becomes the data vector containing log mortality rates (two-dimensional data packed as vector). The order of packing affects only the representation of matrices D_{xx}, D_{xt} and D_{tt}.

The above model, along with those outlined previously, uses the L_2 norm for both the fit and smoothness components. The choice of the L_2 norm is optimal in the case of normality (of errors in the case of fit and second derivatives in the case of smoothness), because it gives a minimizing function that can be shown to be equivalent to a maximum likelihood approach.

In the case of fit, a normal approximation can be justified as follows. Deaths can be assumed to follow a Poisson distribution (Brillinger 1986), which can be approximated by a normal distribution when there are many deaths, and hence errors can also be assumed close to normal. However, it has been demonstrated that the alternative L_1 norm has the advantage of being robust in the face of outliers (see, for example, Portnoy 1997; Schuette 1978).

In the case of smoothness, a normal distribution of second derivatives is not always supported by data. For example, Figures 8 and 9 show that distributions of second derivatives of French female log mortality rates deviate considerably from normal, particularly over time and for the more variable 1935 to 1955 period. The distributions have longer tails (or outliers) and therefore assumptions of normality can lead to suboptimal smoothing performance. For an example of similar behavior in the case of linear models see Nyquist (1983) and Sposito, Hand and Skarpness (1983).

Hence there is justification for replacing the L_2 norm by the L_1 norm. We apply the new norm to both the fit and smoothness components of the minimizing equation. As outlined previously, the L_1 norm has advantages related to fit when the the data contain outliers. While this is a useful feature, the reason we adopt the L_1 norm here is mainly oriented to the smoothness component and in particular to avoid over-smoothing abrupt features in the mortality surface.

Replacing the L_2 norm with the L_1 norm now gives L_1 Regularized Median Smoothing. In addition, we use three different λ coefficients before every derivative to separately adjust the influence of each derivative on the smoothing. Therefore in this method we

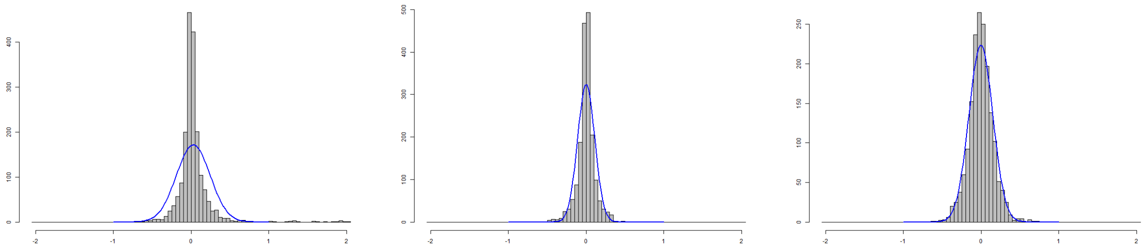

Figure 8: *Histograms of second differences* $\left(\frac{\partial^2}{\partial x^2}, \frac{\partial^2}{\partial x \partial t}, \frac{\partial^2}{\partial t^2}\right)$ *of natural logarithms of French female mortality rates (years 1950–1970) with curves for normal distribution, fitted to the data.*

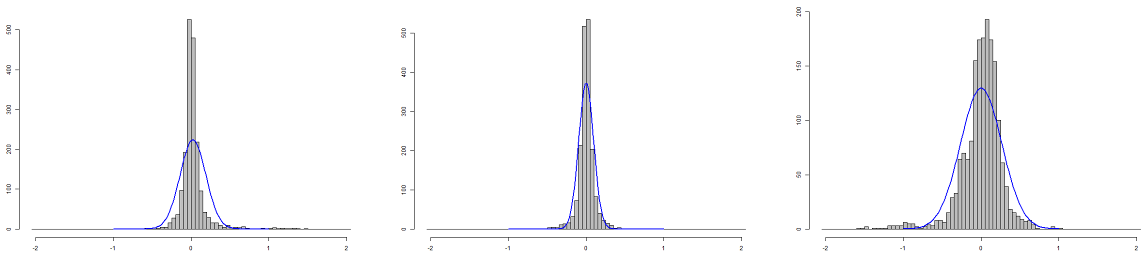

Figure 9: *Histograms of second differences* $\left(\frac{\partial^2}{\partial x^2}, \frac{\partial^2}{\partial x \partial t}, \frac{\partial^2}{\partial t^2}\right)$ *of natural logarithms of French female mortality rates (years 1935–1955) with curves for normal distribution, fitted to the data.*

define smoothing as $Q(y) = \arg\min_z(K(y,z))$ where

$$K(y,z) = ||y - z||_{L_1} + \lambda_{xx}||D_{xx}z||_{L_1} + \lambda_{xt}||D_{xt}z||_{L_1} + \lambda_{tt}||D_{tt}z||_{L_1}, \tag{2}$$

y, D_{xx}, D_{xt} and D_{tt} are as described above, and the L_1 norm of some vector v is defined as $||v||_{L_1} = \sum_i |v_i|$.

Minimization of $K(y,z)$ appears to be difficult, but due to a well known procedure (described for example in Wood 2006) the problem can be reduced to a quantile regression problem, which then can be solved with existing software (Koenker 2015). In this study we adopt the following reduction procedure:

- Matrices I, $\lambda_{xx}D_{xx}$, $\lambda_{xt}D_{xt}$, and $\lambda_{tt}D_{tt}$ are stacked on top of each other to give $R = [I,\ \lambda_{xx}D'_{xx},\ \lambda_{xt}D'_{xt},\ \lambda_{tt}D'_{tt}]'$.
- Vector y is extended by zeros until its length is equal to the number of rows in R: $y_{ext} = [y',\ 0']'$.
- $K(y,z) = ||y - z||_{L_1} + \lambda_{xx}||D_{xx}z||_{L_1} + \lambda_{xt}||D_{xt}z||_{L_1} + \lambda_{tt}||D_{tt}z||_{L_1}$ is replaced with the equivalent expression $K(y,z) = ||y_{ext} - Rz||_{L_1}$.

Then finding $Q(y) = \arg\min_z(K(y,z))$ is a quantile regression problem. The missed points in vector y, if they happen, are treated the natural way — by removing corresponding indexes in computation $||y - z||_{L_1} = \sum_i |y_i - z_i|$ (the first term of $K(y,z)$). The other terms of $K(y,z)$ stay intact in this case.

The smoothing method described above is defined for some fixed parameters λ_{xx}, λ_{xt} and λ_{tt}, which need to be optimized to get maximum performance. As a measure of performance we use the predictive ability of the procedure, estimated using the mean absolute error based on five-fold cross validation. The function measuring performance depends on parameters λ_{xx}, λ_{xt} and λ_{tt}. This function may have many local minima which makes the process of finding optimal parameters difficult. We optimize parameters λ_{xx}, λ_{xt} and λ_{tt} using the standard optimization procedure "optim

(Nelder-Mead)" (R Core Team 2015) which works well in practice and tends to avoid local minima and therefore has a great chance to find a global optimal solution.

Every subset of data has about 20% of missing values and they each have the same (but shifted) pattern (Figure 10). The points are missed in a regular pattern to ensure the distance between them is as large as possible. Assuming that distant points affect smoothed values less than close points, the result is a fair compromise between closeness to leave-one-out cross validation and good computational time. Clearly such pseudo leave-one-out cross validation requires about five times more resources (processor time) compared to a single smoothing task over the whole data set.

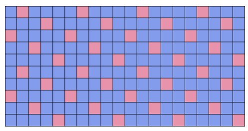

Figure 10: *Missed data pattern for pseudo leave-one-out cross validation.*

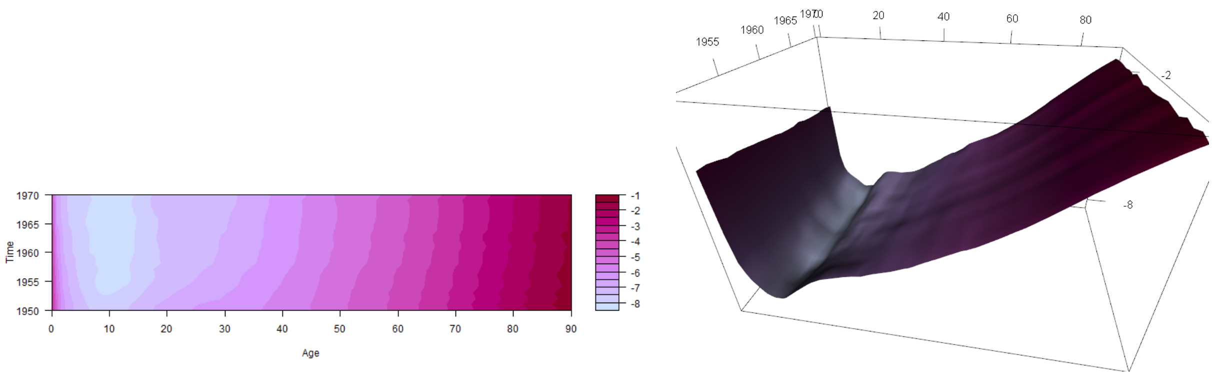

Figure 11: *French female mortality rates after applying the L_1 Regularized Median Smoothing.*

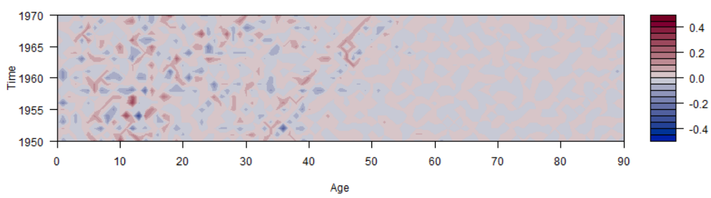

Figure 12: *French female mortality rates residuals after applying the L_1 Regularized Median Smoothing.*

The result of smoothing is shown in Figure 11. Visually, it is less "smooth" than the results of the previous two methods. Nevertheless we will see in Section 6 that this smoothing method reflects "features" of the data more precisely than two previous smoothing methods.

The residuals are shown in Figure 12. Visually it is difficult to find any serial correlation in the errors. However, cohort and possibly also period effects are visible.

If the knots do not form a regular grid, the method needs to be adjusted. On the other hand, the current version allows for missing data, therefore some irregularity in the grid is already permitted.

Finally, we can add that, similarly to the Hyndman and Ullah (2007) method described in Section 2, we can use weights to control for heteroscedasticity of the data. Thus, equation (2) becomes

$$K(y,z) = ||w(y-z)||_{L_1} + \lambda_{xx}||D_{xx}z||_{L_1} + \lambda_{xt}||D_{xt}z||_{L_1} + \lambda_{tt}||D_{tt}z||_{L_1}, \tag{3}$$

where weights w_i are taken as the inverse of the estimated standard deviations of y_i. In Section 6 we show that such modification has only a minimal effect on the method's performance.

The method is publicly available in the R package "smoothAPC" (Dokumentov and Hyndman 2017).

5 L_1 Regularized Median Smoothing with cohort and period effects

The cohort and period effects seen in Figure 12 suggest that the smoothing model can be improved by incorporating these features explicitly. This leads us to our new and final smoothing method in which we first apply the L_1 Regularized Median Smoothing algorithm of the previous section, and then identify and incorporate significant period and cohort effects. We call this the MARCIP1 method (read as “marzipan”): Median Regression with Cohort and Period effects, l One regularized.

To identify the period and cohort effects, we compute the residuals from the L_1 Regularized Median Smoothing algorithm described in the previous section. Then we split the matrix of the residuals into a set of vectors (of different length) representing diagonals, and carry out the following tests over each diagonal.

1. Perform two-sided t-tests of the residuals over all diagonals to find diagonals with residual mean values significantly different from zero.
2. Perform one-sided sample correlation tests for residual diagonals to find diagonals with positively correlated errors (every diagonal is tested for serial correlation with lag one).

To identify the period effects, we perform the same procedure over every column (representing the same year) of the residuals:

3. Perform two-sided t-tests of the residuals over all years to find years with residual mean values significantly different from zero.
4. Perform one-sided sample correlation tests to find years with positively correlated errors (residuals for every particular year are tested for serial correlation with lag one along age dimension).

Since we run multiple tests, there is a high probability that some of them will give false positive results. We accept such behavior since the vast majority of the tests will test data correctly and they will improve the performance more than the minority of false positive tests spoil it. Therefore overall we expect greater performance after such a procedure. Moreover the procedure reduces the sizes of the matrices in computations, which makes computations faster.

The new smoothing model involves summing four components: smooth mortality rates, cohort effects being non zero only along the diagonals identified in tests 1 and 2 above, period effects being non zero only along years identified in tests 3 and 4 above, and the noise. These four components are estimated using the following model:

$$Q(y) = \underset{z_{sm}, z_{coh}, z_{per}}{\arg\min} \left(K(y, z_{sm}, z_{coh}, z_{per}) \right),$$

where

$$\begin{aligned} K(y, z_{sm}, z_{coh}, z_{per}) = {} & \|y - (z_{sm} + z_{coh} + z_{per})\|_{L_1} + \lambda_{xx}\|D_{xx} z_{sm}\|_{L_1} + \lambda_{xt}\|D_{xt} z_{sm}\|_{L_1} + \lambda_{tt}\|D_{tt} z_{sm}\|_{L_1} \\ & + \lambda_{coh}\|D_{coh} z_{coh}\|_{L_1} + \theta_{coh}\|z_{coh}\|_{L_1} + \lambda_{per}\|D_{per} z_{per}\|_{L_1} + \theta_{per}\|z_{per}\|_{L_1}; \quad (4) \end{aligned}$$

- y, D_{xx}, D_{xt} and D_{tt} are as described above;
- z_{sm}, z_{coh} and z_{per} are estimated components representing respectively smooth mortality surface, cohort effects restricted to some diagonals and period effects restricted to some years;
- D_{coh} is a linear differentiation operator representing a discrete version of the second directional derivative in the diagonal cohort direction;
- $D_{per} = D_{tt}$ is a linear differentiation operator representing a discrete version of the second derivative along the years axis;
- $\lambda_{xx}, \lambda_{xt}$, and λ_{tt} are parameters controlling the smoothness of the mortality surface;
- λ_{coh} and λ_{per} are parameters controlling the smoothness of the cohort effects and the period effects respectively;
- θ_{coh} and θ_{per} are parameters controlling shrinking (respectively) the cohort effects and the period effects towards zero.

It may appear that components z_{sm}, z_{coh} and z_{per} duplicate each other. However, this is not the case because usually the best cross-validation performance is achieved when λ_{coh} and λ_{per} are much greater than values of parameters λ_{xx}, λ_{xt} and λ_{tt}. This is because when λ_{coh} and λ_{per} are much greater than λ_{xx}, λ_{xt} and λ_{tt}, then z_{coh} and z_{per} tend to reflect long "one dimensional" features and z_{sm} tend to reflect "two dimensional" features (which are in area "smaller" in diameter than "length" of "one dimensional" features). Thus, z_{coh} and z_{per} reflect cohort and period effects and z_{sm} reflects the smooth surface. We speculate that such separation of features is the reason for the improvement of cross-validation performance. Knowing such behavior, we restrict λ_{coh} and λ_{per} values to be much higher than values of parameters λ_{xx}, λ_{xt} and λ_{tt} to speed up the process of parameter estimation (technically it is done by setting ranges within which the parameters can vary).

We optimize five parameters, λ, and two parameters, θ, using the approach described in Section 4. First we run the L_1 Regularized Median Smoothing method to find starting values for the λ_{xx}, λ_{xt} and λ_{tt} parameters; these are then optimized along with the other parameters using L_1 Regularized Median Smoothing with cohort and period effects.

It is also worth mentioning that the above tests for cohort and period effects are done only for the purpose of reducing computational complexity of the minimization problem. The multiple testing that is carried out means that the selected cohort and period effects are not necessarily statistically significant overall. Some or all of these cohort and period effects will be dropped in the subsequent minimization.

As in the previous section, to minimize $K(y, z_{sm}, z_{coh}, z_{per})$, we use the corresponding quantile regression problem in which:

- vectors z_{sm}, z_{coh} and z_{per} are stacked on top of each other as a single vector, $z_{ext} = [z'_{sm}, z'_{coh}, z'_{per}]'$;

- matrices I, $\lambda_{xx}D_{xx}$, $\lambda_{xt}D_{xt}$, $\lambda_{tt}D_{tt}$, $\lambda_{coh}D_{coh}$, $\lambda_{per}D_{per}$, $\theta_{coh}I$, and $\theta_{per}I$ are combined in one matrix,

$$R = \begin{bmatrix} I & I & I \\ \lambda_{xx}D_{xx} & 0 & 0 \\ \lambda_{xt}D_{xt} & 0 & 0 \\ \lambda_{tt}D_{tt} & 0 & 0 \\ 0 & \lambda_{coh}D_{coh} & 0 \\ 0 & \theta_{coh}I & 0 \\ 0 & 0 & \lambda_{per}D_{per} \\ 0 & 0 & \theta_{per}I \end{bmatrix};$$

- vector y is extended by zeros to have its length equal to the number of rows in R: $y_{ext} = [y',\ 0']'$;
- and

$$K(y, z_{sm}, z_{coh}, z_{per}) = ||y-(z_{sm}+z_{coh}+z_{per})||_{L_1} + \lambda_{xx}||D_{xx}z_{sm}||_{L_1} + \lambda_{xt}||D_{xt}z_{sm}||_{L_1} + \lambda_{tt}||D_{tt}z_{sm}||_{L_1} + \lambda_{coh}||D_{coh}z_{coh}||_{L_1} + \theta_{coh}||z_{coh}||_{L_1} + \lambda_{per}||D_{per}z_{per}||_{L_1} + \theta_{per}||z_{per}||_{L_1}$$

  is replaced with the equivalent expression $K(y, z_{ext}) = ||y_{ext} - Rz_{ext}||_{L_1}$.

Then $Q(y) = \arg\min_{z_{ext}}(K(y, z_{ext}))$ is a quantile regression problem.

The resulting smoothed surface z_{sm} is shown in Figure 13. The cohort effects have been completely removed, and there are minimal remaining period effects.

The estimated cohort effects z_{coh} are shown in Figure 14. The strongest effects, in order, are for cohorts born in 1920, 1916, 1919, 1915 and 1926. Caselli et al. (1987) found evidence that Italian males and females born during the war years 1914 to 1918 had higher mortality than adjacent cohorts for at least 30 years. They did not, however, observe a similar effect in France. Our findings are consistent with the lack of an overall

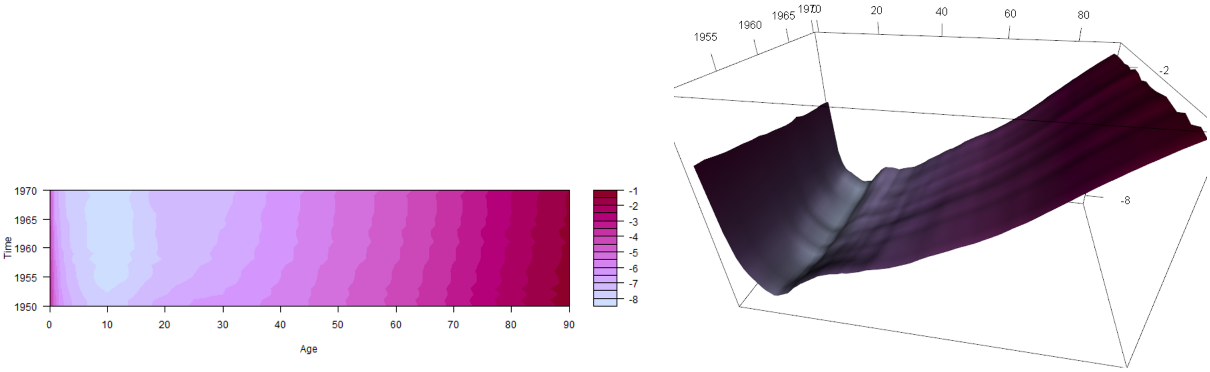

Figure 13: *Logarithms of French female mortality rates smoothed with the MARCIP1 method.*

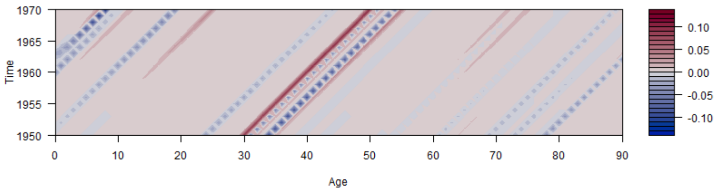

Figure 14: *Cohort effects of logarithms of French female mortality rates.*

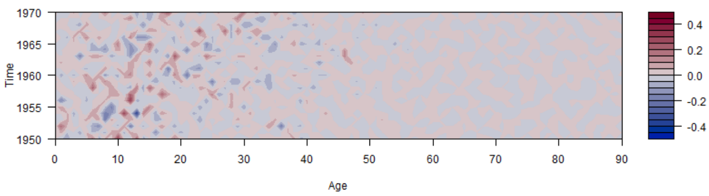

Figure 15: *Residuals of logarithms of French female mortality rates after smoothing with the MARCIP1 method.*

cohort effect, because the 1915 and 1920 cohorts show higher mortality and the 1916 and 1919 cohorts lower mortality, with only a very small net increase for the combined cohort. They are also consistent with the findings of Richards (2008) of lower mortality for the 1919 cohort in England and Wales. Richards (2008) traces this pattern, in part, to a surge of births in late 1919 and early 1920 that resulted in the 1919 cohort being young on average and the 1920 cohort old on average; similar surges were observed in France and other European countries (Vandenbroucke 2014). Higher perinatal mortality for the 1919 cohort due to the Spanish influenza epidemic (in a study by Harris 1919, 26% of pregnant women who survived the epidemic lost the child) potentially also resulted in lower ongoing mortality among the select, hardy survivors. Finally, data artefacts due to difficulties estimating population size because of wild swings in the pattern of births may have contributed to observed cohort patterns (Richards 2008).

The MARCIP1 method does not find any significant period effects in this data. In Appendix 2 we show that the method estimates small period and cohort effects rather roughly. Nevertheless, it allows the method to estimate the "true" mortality surface more precisely.

The residuals from the model are shown in Figure 15. The period and cohort effects are no longer visible. Figure 16 depicts the complete surface with the cohort and period effects added to the smooth surface. It is the "signal" which we have separated from the "noise" (represented by the residuals).

Finally, similarly to Section 4, we can use weights to control for heteroscedasticity of the data. Thus, equation (4) becomes

$$K(y, z_{sm}, z_{coh}, z_{per}) = \|w(y-(z_{sm}+z_{coh}+z_{per}))\|_{L_1} + \lambda_{xx}\|D_{xx}z_{sm}\|_{L_1} + \lambda_{xt}\|D_{xt}z_{sm}\|_{L_1} + \lambda_{tt}\|D_{tt}z_{sm}\|_{L_1} + \lambda_{coh}\|D_{coh}z_{coh}\|_{L_1} + \theta_{coh}\|z_{coh}\|_{L_1} + \lambda_{per}\|D_{per}z_{per}\|_{L_1} + \theta_{per}\|z_{per}\|_{L_1}, \quad (5)$$

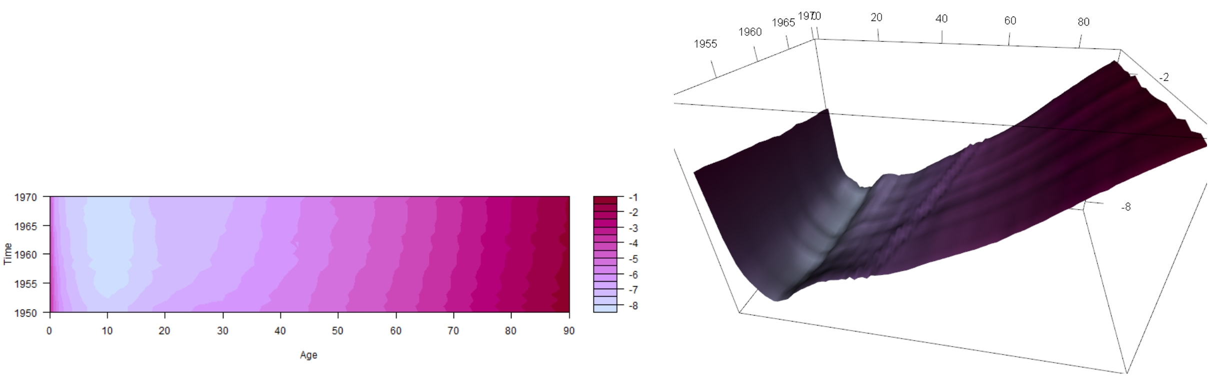

Figure 16: *Logarithms of French female mortality rates smoothed with the MARCIP1 method.*

where weights w_i are taken as the inverse of the estimated standard deviations of y_i. In Section 6 we show that such modification has only a minimal effect on the method's performance.

The method is publicly available in R package "smoothAPC" (Dokumentov and Hyndman 2017).

6 Comparison

We use a cross validation procedure for comparing the different smoothing methods we have discussed. Cross validation is a technique that can be used to validate the predictive ability of a model. In a smoothing paradigm, "clean" data is contaminated with "noise", and smoothing is used to remove the noise, without distorting the data. Since the "clean" data cannot be observed (except in simulation), cross validation provides a valuable tool to evaluate the quality of a smoothing method.

In this paper we consider the predictive ability of a smoothing model to be equivalent to an ability to estimate the "true" mortality surface. The predictive ability of a model should not be confused with a better fit to the data. For example, more complex and, therefore, more flexible models do not always have better predictive abilities, although they usually give a better fit to the data. An extreme example of a flexible model is the model which fits the original data perfectly and, therefore, provides no generalization, has no predictive ability and does not smooth the data.

We use a 20-fold cross validation procedure where every subset of data used in the procedure randomly misses 5% of the original data points. The resulting cross validation error is the average of errors over those 20 subsets.

We compare six methods. The first is the Hyndman and Ullah (2007) method. The second is the Camarda (2012) method (which is tested only for ages 10 to 60 and years 1950 to 1970 due to its limitations). The third is L_1 Regularized Median Smoothing method without taking heteroscedasticity into account. The fourth is MARCIP1 without accounting for heteroscedasticity. The last two methods repeat methods three and four but they take heteroscedasticity into account. To do that, weights w_i in formulas (3) and (5) are set to the inverse of the estimated standard deviations of the corresponding data points.

We use four subsets of the available French female mortality data for our comparisons.

1. Data for years 1950–1970 and ages 10–60 represent a relatively smooth surface. This comparison is useful to ensure that the most "responsive" algorithms using the L_1 norm do not perform any worse than the more "stable" algorithms based on the L_2 norm. This data set is also important because it is the only comparison satisfying the requirements for Method 2 (Camarda 2012) which is designed to work for ages starting from 10 years and where there are no outliers.

2. Data for years 1950–1970 and ages 0–60 represent a period when no outliers happened — there were no global wars or large pandemics. The younger ages from 0 to 9 have abrupt changes in mortality rates which are more challenging to smooth.

3. Data for years 1935–1955 and ages 10–60 represent a period including outliers due to World War II. It is important to mention that such an outlier should be considered as an outlier only along the time dimension and not over age. Therefore in our case, when only uncorrelated errors are considered as noise, such a one-dimensional outlier should be preserved during smoothing as signal. These data are important for testing the smoothing abilities of the algorithms in the presence of one-dimensional outliers.

4. Data for years 1935–1955 and ages 0–60 represents the most complex dataset containing the one-dimensional outliers (WWII) and also a period of abrupt mortality changes for ages 0 to 9.

All calculations are done using R (R Core Team 2015). To ensure an automatic and fair selection of the smoothing parameters, we simply used all methods with default arguments as set by the package authors (the second method was also tested on more granular grid than suggested by the default settings and the modifications led to superficial difference in performance).

Years	Ages	SE1	SE2	SE3	SE4	SE5	SE6	AE1	AE2	AE3	AE4	AE5	AE6
1950-1970	10-60	1.15	0.72	0.57	0.46	0.55	0.51	6.39	5.86	5.12	4.64	4.94	4.73
1950-1970	0-60	0.78		0.43	0.41	0.42	0.40	5.61		4.81	4.73	4.71	4.60
1935-1955	10-60	1.05		0.45	0.39	0.44	0.38	4.91		4.59	4.11	4.52	4.04
1935-1955	0-60	0.85		0.40	0.39	0.37	0.38	5.09		4.38	4.13	4.12	4.15

Table 1: *Cross validation performance of different smoothing methods against French female mortality data (SE1 is MSE for cross validation of Method 1 mutiplied by 100, AE1 is MAE for cross validation of Method 1 mutiplied by 100, ... , AE6 is MAE for cross validation of Method 6 mutiplied by 100)*

Table 1 shows Mean Square Error (MSE) and Mean Absolute Error (MAE) multiplied by 100 for the five methods. In all cases the error in a particular age and time cell is the difference between the observed log mortality rate and the modelled / smoothed log mortality rate incorporating the smooth bivariate surface as well as any cohort and period effects. Results indicate that the MARCIP1 method (with and without taking heteroscedasticity into account) shows better or similar performance compared to the other methods. Amongst the other methods, the L_1 Regularized Median Smoothing performs best in all cases. Overall, the MARCIP1 method is the best performing method for French female mortality.

In Appendix 1 we provide comparable cross validation performance tests for French males and for males and females in eleven other countries. Table 2 shows the average MSE and MAE across all 24 populations for each of the four period and age subsets. Table 3 shows the number of populations for which each method was ranked best (i.e. the lowest error, with each method given one half in the case of a tie to two decimal places). The results clearly indicate that the previous conclusions based on French females also apply for this more comprehensive dataset.

Years	Ages	SE1	SE2	SE3	SE4	SE5	SE6	AE1	AE2	AE3	AE4	AE5	AE6
1950-1970	10-60	2.83	1.51	1.36	1.31	1.35	1.30	9.07	7.97	7.44	7.25	7.48	7.20
1950-1970	0-60	3.57		1.37	1.35	1.39	1.38	9.65		7.62	7.49	7.67	7.53
1935-1955	10-60	2.01		0.94	0.88	0.93	0.88	7.82		6.46	6.11	6.45	6.09
1935-1955	0-60	2.45		1.05	0.98	1.03	0.98	8.48		6.88	6.57	6.91	6.62

Table 2: *Average cross validation results across all 24 populations.*

Years	Ages	SE1	SE2	SE3	SE4	SE5	SE6	AE1	AE2	AE3	AE4	AE5	AE6
1950-1970	10-60	0	0	3	11	4	6	0	1	5	12	2	4
1950-1970	0-60	0		8	10	0	6	0		3	11	2	7
1935-1955	10-60	0		6	9	1	8	1		5	9	0	9
1935-1955	0-60	0		1	10	4	9	0		2	10	2	10

Table 3: *Number of times among 24 populations that each method achieves lowest error.*

7 Conclusion

In this paper we propose two methods to smooth mortality data in two dimensions, and compare them with existing one- and two-dimensional methods. The recommended MARCIP1 method gives the best estimate of the "true" mortality surface[3] , in addition to providing insights into the existence of cohort and period effects that might otherwise be overlooked.

We conclude that use of two-dimensional mortality data and thin plate splines for smoothing generally leads to improvements compared to a one-dimensional approach. On the other hand, using the L_1 instead of the L_2 norm can lead to further performance improvements. Another performance improvements can be achieved by building into the model explicit allowance for cohort and period effects.

Additional improvements in these proposed methods are possible. While only partially adaptive splines were used for the L_1 Regularized Median Smoothing methods, a fully adaptive approach applied to the L_1 Regularized Median Smoothing methods may provide further improvements and, therefore, requires further investigation.

[3]As shown by comparing cross validation performance of MARCIP1 with the performance of other methods.

The λ coefficients used in the L_1 Regularized Median Smoothing methods were estimated using lengthy numerical methods. A simpler procedure, similar to that used by Camarda (2012), would improve their practical usefulness. We leave this to a later paper.

Another potential improvement is to model several countries simultaneously, taking account of the existence of similar features in related countries. However, this would substantially increase the complexity of the optimization routines, and it is not clear that there is sufficient information available in the cross-country correlations that is not already exploited by smoothing within countries.

Code for the article can be accessed from the on-line repository at https://goo.gl/ZuDXkW.

References

Andreev, K. and Vaupel, J. (2005). 'Patterns of Mortality Improvement over Age and Time: Estimation, Presentation and Implications'. In: *Population Association of America 2005 Annual Meeting*. Vol. 31.

Barbi, E. and Camarda, C. G. (2011). Period and cohort effects on elderly mortality: a new relational model for smoothing mortality surfaces. *Statistica* 71(1): 51–69.

Brillinger, D. R. (1986). The Natural variability of vital rates and associated statistics. *Biometrics* 42: 693–734.

Brunel, N. J. B., Clairon, Q. and d'Alché-Buc, F. (2014). Parametric estimation of ordinary differential equations with orthogonality conditions. *Journal of the American Statistical Association* 109(505): 173–185.

Camarda, C. G. (2012). MortalitySmooth: An R package for smoothing Poisson counts with P-splines. *Journal of Statistical Software* 50(1): 1–24.

Camarda, C. G., Eilers, P. H. C. and Gampe, J. (2010). 'Sums of smooth exponentials'. In: *Proceedings of the 25th International Workshop on Statistical Modelling*: 113–118.

Camarda, C. G., Eilers, P. H. C. and Gampe, J. (2012). 'Additive decomposition of vital rates from grouped data'. In: *Proceedings of the 27th International Workshop on Statistical Modelling*: 57–62.

Carriere, J. F. (1992). Parametric models for life tables. *Transactions of the Society of Actuaries* 44: 77–99.

Caselli, G., Vallin, J., Vaupel, J. W. and Yashin, A. (1987). Age-specific mortality trends in France and Italy since 1900: Period and cohort effects. *European Journal of Population / Revue européenne de Démographie* 3(1): 33–60.

Chen, K. and Müller, H.-G. (2012). Modeling repeated functional observations. *Journal of the American Statistical Association* 107(500): 1599–1609.

Currie, I. D., Durban, M. and Eilers, P. H. (Dec. 2004). Smoothing and forecasting mortality rates. *Statistical Modelling* 4(4): 279–298.

Dokumentov, A. and Hyndman, R. J. (2017). *Smoothing of Two-Dimensional Demographic Data, Optionally Taking into Account Period and Cohort Effects*. R package version 0.2. URL: http://cran.r-project.org/package=smoothAPC.

Ferraty, F. and Vieu, P. (2006). *Nonparametric functional data analysis: theory and practice*. Springer.

Forfar, D. O., McCutcheon, J. J. and Wilkie, A. D. (1988). On graduation by mathematical formula. *Journal of the Institute of Actuaries* 115(1): 1–149.

Gertheiss, J., Maity, A. and Staicu, A.-M. (2013). Variable selection in generalized functional linear models. *Stat* 2(1): 86–101.

Hannerz, H. (2001). Presentation and Derivation of a Five-Parameter Survival Function Intended to Model Mortality in Modern Female Populations. *Scandinavian Actuarial Journal* 2001(2): 176–187.

Harris, J. (1919). Influenza occurring in pregnant women: A statistical study of thirteen hundred and fifty cases. *Journal of the American Medical Association* 72(14): 978–980.

He, X., Ng, P. and Portnoy, S. (1998). Bivariate quantile smoothing splines. *Journal of the Royal Statistical Society: Series B* 60(3): 537–550.

Heligman, L. and Pollard, J. H. (1980). The age pattern of mortality. *Journal of the Institute of Actuaries (1886-1994)* 107(1): 49–80.

Henderson, R. (1924). A new method of graduation. *Transactions of the Actuarial Society of America* 25: 29–40.

Horváth, L. and Kokoszka, P. (2012). *Inference for functional data with applications*. Springer.

Huang, J. Z., Shen, H. and Buja, A. (2009). The Analysis of Two-Way Functional Data Using Two-Way Regularized Singular Value Decompositions. *Journal of the American Statistical Association* 104(488): 1609–1620.

Human Mortality Database (2008). University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Data downloaded on 20 Feb 2008. URL: http://www.mortality.org.

Hyndman, R. J. (2014). *demography: Forecasting mortality, fertility, migration and population data*. R package version 1.16. With contributions from Heather Booth and Leonie Tickle and John Maindonald.

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.

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

Janssen, F. and Kunst, A. E. (2005). Cohort patterns in mortality trends among the elderly in seven European countries, 1950-99. *International Journal Of Epidemiology* 34(5): 1149–1159.

Johnson, N. P. and Mueller, J. (2002). Updating the accounts: global mortality of the 1918-1920 "Spanish" influenza pandemic. *Bulletin of the History of Medicine* 76(1): 105–115.

Kirkby, J. and Currie, I. (2010). Smooth models of mortality with period shocks. *Statistical Modelling* 10(2): 177–196.

Koenker, R. (2015). *quantreg: Quantile Regression*. R package version 5.11.

Koenker, R., Ng, P. and Portnoy, S. (1994). Quantile smoothing splines. *Biometrika* 81(4): 673–680.

Marron, J., Ramsay, J. O., Sangalli, L. M. and Srivastava, A. (2015). *Functional Data Analysis of Amplitude and Phase Variation*. MOX Report 27/2015. Department of Mathematics, Politecnico di Milano, Italy.

Nyquist, H. (1983). The optimal L_p norm estimator in linear regression models. *Communications in Statistics-Theory and Methods* 12(21): 2511–2524.

O'Connell, A. and Dunstan, K. (Jan. 2009). Do cohort mortality trends emigrate? Insights on the U.K.'s Golden Cohort from a comparison with a British settler country. *British Actuarial Journal* 15(S1): 91–121.

Portnoy, E. (1997). Regression-quantile graduation of Australian life tables, 1946–1992. *Insurance: Mathematics and Economics* 21(2): 163–172.

R Core Team (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria.

Richards, S. J. (2008). Detecting year-of-birth mortality patterns with limited data. *Journal of the Royal Statistical Society. Series A (Statistics in Society)* 171(1): 279–298.

Richards, S. J., Kirkby, J. G. and Currie, I. D. (2006). The importance of year of birth in two-dimensional mortality data. *British Actuarial Journal* 12(1): 5–38.

Ruppert, D., Wand, M. P. and Carroll, R. J. (2003). *Semiparametric Regression*. New York: Cambridge University Press.

Sangalli, L. M. (2014). 'Statistical and numerical techniques for spatial functional data analysis'. In: *Contributions in infinite-dimensional statistics and related topics*. Ed. by E. G. Bongiorno, E. Salinelli, A. Goia and P. Vieu. Società Editrice Esculapio.

Schuette, D. R. (1978). A linear programming approach to graduation. *Transactions of Society of Actuaries* 30: 407–431.

Shen, D., Shen, H., Bhamidi, S., Muñoz Maldonado, Y., Kim, Y. and Marron, J. (2014). Functional data analysis of tree data objects. *Journal of Computational and Graphical Statistics* 23(2): 418–438.

Silverman, B. and Ramsay, J. (2005). *Functional Data Analysis*. Springer.

Sposito, V., Hand, M. and Skarpness, B. (1983). On the efficiency of using the sample kurtosis in selecting optimal lpestimators. *Communications in Statistics-Simulation and Computation* 12(3): 265–272.

Vandenbroucke, G. (2014). Fertility and wars: The case of World War I in France. *American Economic Journal: Macroeconomics* 6(2): 108–136.

Whittaker, E. T. (1922). On a new method of graduation. *Proceedings of the Edinburgh Mathematical Society* 41: 63–75.

Willets, R. (2004). The cohort effect: insights and explanations. *British Actuarial Journal* 10(04): 833–877.

Wood, S. N. (1994). Monotonic smoothing splines fitted by cross validation. *SIAM Journal on Scientific Computing* 15(5): 1126–1133.

Wood, S. N. (2006). *Generalized Additive Models: an introduction with R*. CRC Press.

Zhu, H., Fan, J. and Kong, L. (2014). Spatially varying coefficient model for neuroimaging data with jump discontinuities. *Journal of the American Statistical Association* 109(507): 1084–1098.

Zhu, H., Strawn, N. and Dunson, D. B. (2014). *Bayesian graphical models for multivariate functional data*. arXiv preprint arXiv:1411.4158.

Appendix 1

In this appendix, we repeat the cross-validation performance tests for the French male data, and eleven other countries (male and female). All data is taken from *Human Mortality Database* (2008). The countries selected are all those for which the Human Mortality Database has data available for the period 1935–1970. Overall 96 tests representing 12 (countries) by 2 (male, female) by 4 (age and period subsets) were performed.

The test results provided in Tables 2 to 24 show that the MARCIP1 method, either without (Method 4) or with (Method 6) allowance for heteroscedasticity, show similar or better performance for ages from 10 to 60 and outperforms other tested methods in the majority of tests for ages from 0 to 60. In general the MARCIP1 is the best performing method among those tested.

The most difficult data sets for Methods 3, 4, 5 and 6 are the data where log mortality rates are almost flat (for example for ages 10–60), where there are few features to extract, and which contain much noise (for example when country population and mortality are low at the same time). The combination of such features leads to modest results of the new methods. Although disappointing, this is to be expected since for completely flat and very noisy data, it is very difficult to outperform a simple linear regression.

Similarly to Table 1, Tables 2 to 24 contain figures for cross validation performance of tested smoothing methods against mortality data for a particular country and gender. SE1 column contains MSE for cross validation of Method 1 multiplied by 100, AE1 column contains MAE for cross validation of Method 1 multiplied by 100,. . . , AE5 column contains MAE for cross validation of Method 5 multiplied by 100.

Years	Ages	SE1	SE2	SE3	SE4	SE5	SE6	AE1	AE2	AE3	AE4	AE5	AE6
1950-1970	10-60	2.80	2.25	2.08	1.86	2.12	1.88	11.78	11.37	11.00	10.19	10.95	10.23
1950-1970	0-60	3.94		1.93	1.71	1.79	1.73	12.66		10.23	9.53	9.99	9.66
1935-1955	10-60	2.14		1.31	1.18	1.25	1.12	10.43		9.14	8.48	8.82	8.35
1935-1955	0-60	2.02		1.18	1.16	1.17	1.16	10.39		8.45	8.08	8.43	8.27

Table 4: *Australian female*

Years	Ages	SE1	SE2	SE3	SE4	SE5	SE6	AE1	AE2	AE3	AE4	AE5	AE6
1950-1970	10-60	4.05	0.95	0.92	0.89	0.96	0.93	10.10	7.62	7.32	7.18	7.66	7.24
1950-1970	0-60	1.79		1.10	1.02	1.14	1.08	9.45		7.84	7.58	7.95	7.54
1935-1955	10-60	2.88		1.08	0.81	0.98	0.95	8.42		7.70	6.80	7.53	7.15
1935-1955	0-60	2.97		1.26	1.12	1.24	1.21	10.16		8.29	7.74	8.38	7.89

Table 5: *Australian male*

Years	Ages	SE1	SE2	SE3	SE4	SE5	SE6	AE1	AE2	AE3	AE4	AE5	AE6
1950-1970	10-60	1.34	0.93	0.80	0.78	0.83	0.83	7.80	7.23	6.69	6.63	6.85	6.75
1950-1970	0-60	3.39		1.20	1.03	1.09	1.00	9.56		7.93	7.34	7.62	7.15
1935-1955	10-60	2.24		0.79	0.76	0.79	0.79	8.58		6.73	6.48	6.76	6.54
1935-1955	0-60	2.00		0.94	0.81	0.98	0.79	8.47		7.55	7.09	7.71	6.97

Table 6: *Canadian female*

Years	Ages	SE1	SE2	SE3	SE4	SE5	SE6	AE1	AE2	AE3	AE4	AE5	AE6
1950-1970	10-60	2.53	0.64	0.57	0.55	0.60	0.58	8.40	6.12	6.00	5.82	6.26	5.95
1950-1970	0-60	3.84		0.49	0.51	0.53	0.52	7.89		5.17	5.36	5.49	5.47
1935-1955	10-60	1.03		0.70	0.54	0.66	0.54	7.32		6.47	5.66	6.30	5.66
1935-1955	0-60	2.39		0.54	0.46	0.54	0.47	7.32		5.72	5.35	5.84	5.36

Table 7: *Canadian male*

Years	Ages	SE1	SE2	SE3	SE4	SE5	SE6	AE1	AE2	AE3	AE4	AE5	AE6
1950-1970	10-60	7.02	4.82	4.45	4.50	4.49	4.50	17.57	15.58	15.17	15.15	15.00	15.06
1950-1970	0-60	7.12		4.83	4.96	5.01	5.01	17.15		14.81	14.92	15.10	15.14
1935-1955	10-60	4.30		2.34	2.45	2.51	2.37	13.56		11.38	11.64	11.85	11.65
1935-1955	0-60	5.41		2.84	2.40	2.45	2.51	13.55		12.17	11.37	11.67	11.76

Table 8: *Danish female*

Years	Ages	SE1	SE2	SE3	SE4	SE5	SE6	AE1	AE2	AE3	AE4	AE5	AE6
1950-1970	10-60	8.26	2.74	2.79	2.58	2.47	2.47	14.64	10.82	11.45	10.96	10.72	10.75
1950-1970	0-60	6.60		2.60	2.98	2.60	3.04	15.08		12.06	12.58	12.01	12.61
1935-1955	10-60	3.79		2.39	2.41	2.45	2.69	11.32		10.49	10.71	10.92	10.76
1935-1955	0-60	5.75		2.82	2.71	2.71	2.61	13.86		12.20	12.08	12.08	11.82

Table 9: *Danish male*

Years	Ages	SE1	SE2	SE3	SE4	SE5	SE6	AE1	AE2	AE3	AE4	AE5	AE6
1950-1970	10-60	1.22	1.29	0.64	0.41	0.79	0.42	6.65	7.98	5.90	4.76	6.29	4.86
1950-1970	0-60	0.96		0.61	0.47	0.66	0.46	6.17		5.68	5.03	5.83	4.96
1935-1955	10-60	1.96		0.96	0.91	0.95	0.94	7.79		6.81	6.64	6.82	6.79
1935-1955	0-60	0.94		0.88	0.85	0.94	0.79	7.09		7.10	7.11	7.48	6.74

Table 10: *Dutch female*

Years	Ages	SE1	SE2	SE3	SE4	SE5	SE6	AE1	AE2	AE3	AE4	AE5	AE6
1950-1970	10-60	0.97	0.93	0.32	0.31	0.30	0.33	4.79	6.39	4.07	3.68	3.95	3.80
1950-1970	0-60	1.93		0.40	0.35	0.46	0.36	6.02		4.69	4.38	4.84	4.43
1935-1955	10-60	1.75		0.54	0.59	0.60	0.61	7.00		5.55	5.64	5.82	5.78
1935-1955	0-60	0.88		0.58	0.61	0.60	0.58	6.81		5.91	6.03	6.02	6.01

Table 11: *Dutch male*

Years	Ages	SE1	SE2	SE3	SE4	SE5	SE6	AE1	AE2	AE3	AE4	AE5	AE6
1950-1970	10-60	0.69	0.47	0.56	0.41	0.47	0.44	5.45	4.92	5.07	4.52	5.08	4.55
1950-1970	0-60	2.32		0.59	0.57	0.67	0.62	7.62		5.64	5.58	6.06	5.76
1935-1955	10-60	1.46		0.40	0.27	0.47	0.28	6.51		4.66	3.64	5.13	3.74
1935-1955	0-60	2.19		0.49	0.35	0.50	0.35	6.38		5.01	4.16	5.01	4.07

Table 12: *English female*

Years	Ages	SE1	SE2	SE3	SE4	SE5	SE6	AE1	AE2	AE3	AE4	AE5	AE6
1950-1970	10-60	1.56	0.33	0.27	0.22	0.31	0.21	5.96	4.44	3.96	3.51	4.28	3.48
1950-1970	0-60	3.37		0.47	0.37	0.45	0.38	7.49		5.03	4.50	5.06	4.66
1935-1955	10-60	0.71		0.48	0.39	0.47	0.31	5.50		5.09	4.46	4.94	4.06
1935-1955	0-60	3.06		0.57	0.48	0.53	0.44	7.82		5.42	4.94	5.31	4.85

Table 13: *English male*

Years	Ages	SE1	SE2	SE3	SE4	SE5	SE6	AE1	AE2	AE3	AE4	AE5	AE6
1950-1970	10-60	6.11	3.62	3.59	3.54	3.40	3.39	15.78	13.67	13.79	13.66	13.45	13.41
1950-1970	0-60	6.68		3.28	3.41	3.35	3.46	16.27		13.74	13.96	13.67	13.88
1935-1955	10-60	3.04		1.54	1.65	1.73	1.63	10.81		8.71	9.04	9.28	8.87
1935-1955	0-60	2.67		1.85	1.79	1.79	1.77	10.79		9.33	9.27	9.49	9.36

Table 14: *Finnish female*

Years	Ages	SE1	SE2	SE3	SE4	SE5	SE6	AE1	AE2	AE3	AE4	AE5	AE6
1950-1970	10-60	3.37	1.96	1.85	1.88	1.84	1.88	11.29	10.49	10.24	10.47	10.29	10.27
1950-1970	0-60	8.11		2.33	2.35	2.41	2.45	13.70		10.83	10.81	10.90	10.95
1935-1955	10-60	3.58		1.52	1.43	1.21	1.25	9.42		8.82	8.53	8.22	8.18
1935-1955	0-60	3.15		1.41	1.42	1.32	1.35	10.03		8.78	8.71	8.98	8.70

Table 15: *Finnish male*

Years	Ages	SE1	SE2	SE3	SE4	SE5	SE6	AE1	AE2	AE3	AE4	AE5	AE6
1950-1970	10-60	0.83	0.40	0.34	0.29	0.32	0.29	5.43	4.71	4.49	4.00	4.29	3.97
1950-1970	0-60	0.85		0.34	0.26	0.29	0.27	5.23		4.35	3.70	4.01	3.77
1935-1955	10-60	0.43		0.29	0.25	0.27	0.25	3.74		3.55	3.29	3.55	3.30
1935-1955	0-60	1.77		0.35	0.31	0.33	0.32	5.53		4.05	3.87	3.94	3.78

Table 16: *Italian female*

Years	Ages	SE1	SE2	SE3	SE4	SE5	SE6	AE1	AE2	AE3	AE4	AE5	AE6
1950-1970	10-60	0.97	0.37	0.33	0.24	0.32	0.23	5.36	4.38	4.14	3.46	4.10	3.34
1950-1970	0-60	1.38		0.28	0.27	0.28	0.26	5.76		4.01	3.96	4.14	4.00
1935-1955	10-60	0.31		0.31	0.22	0.30	0.28	3.39		3.69	3.40	3.69	3.54
1935-1955	0-60	1.29		0.34	0.39	0.31	0.38	5.15		3.90	4.02	3.77	3.90

Table 17: *Italian male*

Years	Ages	SE1	SE2	SE3	SE4	SE5	SE6	AE1	AE2	AE3	AE4	AE5	AE6
1950-1970	10-60	1.29	0.90	0.59	0.73	0.59	0.64	8.34	7.33	5.50	6.62	5.76	5.87
1950-1970	0-60	2.40		0.74	0.70	0.78	0.71	8.61		6.42	6.12	6.69	6.25
1935-1955	10-60	3.64		0.62	0.37	0.56	0.39	12.82		5.93	4.46	5.55	4.42
1935-1955	0-60	2.88		0.67	0.43	0.79	0.43	11.25		5.89	4.76	6.35	4.88

Table 18: *Spanish female*

Years	Ages	SE1	SE2	SE3	SE4	SE5	SE6	AE1	AE2	AE3	AE4	AE5	AE6
1950-1970	10-60	1.78	0.72	0.48	0.48	0.47	0.45	7.62	6.40	5.13	5.54	5.14	5.28
1950-1970	0-60	2.34		0.47	0.46	0.47	0.43	7.90		5.25	5.48	5.25	5.20
1935-1955	10-60	1.65		0.57	0.35	0.46	0.30	9.42		5.59	4.39	4.99	4.24
1935-1955	0-60	2.40		0.86	0.61	0.83	0.90	9.92		6.04	4.64	5.97	6.04

Table 19: *Spanish male*

Years	Ages	SE1	SE2	SE3	SE4	SE5	SE6	AE1	AE2	AE3	AE4	AE5	AE6
1950-1970	10-60	6.29	4.03	3.76	3.74	3.69	3.72	14.39	13.26	13.02	12.89	12.98	12.95
1950-1970	0-60	8.17		3.20	3.33	3.35	3.35	16.40		12.30	12.65	12.57	12.65
1935-1955	10-60	3.71		1.75	1.68	1.59	1.55	10.02		8.83	8.78	8.89	8.72
1935-1955	0-60	4.55		1.96	1.86	2.03	2.01	11.68		9.98	9.61	10.13	10.00

Table 20: *Swedish female*

Years	Ages	SE1	SE2	SE3	SE4	SE5	SE6	AE1	AE2	AE3	AE4	AE5	AE6
1950-1970	10-60	3.28	1.91	1.83	1.80	1.74	1.67	10.57	9.63	9.07	9.35	9.09	9.08
1950-1970	0-60	6.19		1.96	1.97	1.97	1.99	12.84		10.22	10.43	10.17	10.50
1935-1955	10-60	2.51		1.35	1.29	1.29	1.21	9.62		8.62	8.22	8.38	8.08
1935-1955	0-60	2.78		1.34	1.37	1.34	1.37	10.18		8.54	8.74	8.60	8.70

Table 21: *Swedish male*

Years	Ages	SE1	SE2	SE3	SE4	SE5	SE6	AE1	AE2	AE3	AE4	AE5	AE6
1950-1970	10-60	6.45	4.02	4.08	3.98	4.13	4.12	16.91	14.60	14.94	14.85	15.01	14.93
1950-1970	0-60	4.29		3.31	3.21	3.30	3.30	14.75		13.56	13.15	13.52	13.50
1935-1955	10-60	1.95		1.68	1.60	1.78	1.75	10.34		9.78	9.59	9.89	9.88
1935-1955	0-60	3.18		2.04	2.02	2.09	1.97	12.12		10.95	10.91	11.12	10.76

Table 22: *Swiss female*

Years	Ages	SE1	SE2	SE3	SE4	SE5	SE6	AE1	AE2	AE3	AE4	AE5	AE6
1950-1970	10-60	3.44	1.51	1.30	1.40	1.45	1.34	10.77	8.93	8.42	8.85	9.02	8.77
1950-1970	0-60	2.93		1.70	1.65	1.69	1.73	11.85		9.77	9.69	9.73	9.89
1935-1955	10-60	2.13		1.22	1.23	1.24	1.27	9.75		8.15	8.08	8.16	8.13
1935-1955	0-60	2.12		1.48	1.46	1.47	1.47	10.15		9.15	9.03	9.15	9.03

Table 23: *Swiss male*

Years	Ages	SE1	SE2	SE3	SE4	SE5	SE6	AE1	AE2	AE3	AE4	AE5	AE6
1950-1970	10-60	0.40	0.11	0.10	0.09	0.11	0.10	3.07	2.50	2.15	1.95	2.29	2.14
1950-1970	0-60	2.00		0.10	0.09	0.11	0.11	3.79		2.21	2.11	2.32	2.28
1935-1955	10-60	0.16		0.00	0.00	0.02	0.00	0.88		0.34	0.36	0.52	0.36
1935-1955	0-60	1.08		0.03	0.03	0.04	0.03	2.05		0.88	0.90	1.04	0.86

Table 24: *USA female*

Years	Ages	SE1	SE2	SE3	SE4	SE5	SE6	AE1	AE2	AE3	AE4	AE5	AE6
1950-1970	10-60	0.45	0.07	0.05	0.04	0.06	0.04	2.77	1.97	1.67	1.41	1.74	1.41
1950-1970	0-60	3.02		0.08	0.09	0.08	0.08	3.95		1.99	2.02	2.03	1.87
1935-1955	10-60	0.16		0.01	0.01	0.02	0.01	1.11		0.48	0.45	0.58	0.46
1935-1955	0-60	1.43		0.03	0.03	0.04	0.03	2.45		0.94	0.84	0.99	0.88

Table 25: *USA male*

Years	Ages	SE1	SE2	SE3	SE4	SE5	SE6	AE1	AE2	AE3	AE4	AE5	AE6
1950-1970	10-60	1.75	0.44	0.35	0.30	0.39	0.33	5.94	4.98	4.17	3.79	4.43	3.97
1950-1970	0-60	1.27		0.36	0.36	0.39	0.35	5.91		4.27	4.19	4.38	4.11
1935-1955	10-60	1.61		0.34	0.34	0.33	0.29	5.00		3.87	3.81	3.77	3.59
1935-1955	0-60	0.93		0.41	0.35	0.36	0.33	5.24		4.43	4.24	4.20	4.09

Table 26: *French male*

Appendix 2

In this appendix, we perform tests for simulated mortality data.

To simulate logarithms of mortality rates, we used the formula proposed by Heligman and Pollard (1980):

$$\frac{q_x}{p_x} = A^{(x+B)^C} + De^{-E(\ln x - \ln F)^2} + GH^x$$

where q_x is the probability of dying at age x, $p_x = 1 - q_x$, and A, B, C, D, E, F, G and H are the parameters.

We used this equation to generate sequences of 21 curves of mortality rates for ages from 0 to 90. The coefficients $A,\ldots,H$ followed simulated $I(2)$ processes. The variance parameter of each $I(2)$ process was estimated from the coefficients for male mortality data provided by the same article (Heligman and Pollard 1980, Table 1).

We also introduced errors, cohort and period effects to the simulated data. The variance of the errors was allowed to change with age according to the population at that age and assuming that deaths are Poisson distributed with the parameter proportional to the population size. The initial cohort population was assumed to be 250,000 people for the first 16 simulations, 500,000 for the next 16 and 1,000,000 for the last 16 simulations. Thus, one set of simulations contains 48 simulations. Cohort effects were assumed to be constant in time. The magnitude of a particular cohort effect was taken randomly and independently from the array of the simulated errors. Period effects were assumed to be piece-wise linear starting with 0 at age 0 and then magnitude changed randomly and independently at ages 24, 46, 68, and 91 years. The magnitude of the effect was taken randomly and independently from the array of the simulated errors for every period effect. All effects were assumed to be positive in sign (so that they increased the mortality). All tests were run with 24 simulated cohort and 3 simulated period effects.

Two types of errors were calculated for every simulation. The first type is calculated as differences between simulated mortality rates including cohort and period effects and smoothed (estimated) mortality rates with period and cohort effects (where available). Mean Squared Errors (columns SE1, SE2, SE3, SE4, SE5, SE6) and Mean Absolute Errors (columns AE1, AE2, AE3, AE4, AE5, AE6) are reported in Table 27.

The second type of error reported is the Mean Absolute Percentage Error between estimated and simulated period and cohort effects (in Table 28 columns PER4 and COH4 for MARCIP1 without accounting for heteroscedasticity and PER6 and COH6 for MARCIP1 with heteroscedasticity taken into account). We used the following formula to calculate MAPE:

$$MAPE = \frac{\sum_i |E_i - A_i|}{\sum_i |A_i|} \times 100\%,$$

where E_i are the estimated and A_i are the actual values.

The simulations were also done for two greater levels of magnitude of period and cohort effects: when the effects were twice bigger (set 2) and four times bigger (set 3) than in the first set of simulations. The aggregated results fir these two sets are also reported in Tables 27 and 28.

Set	Population	SE1	SE2	SE3	SE4	SE5	SE6	AE1	AE2	AE3	AE4	AE5	AE6
1	250000	0.10	0.07	0.07	0.03	0.06	0.04	2.07	1.81	1.59	1.12	1.56	1.22
1	500000	0.06	0.04	0.04	0.02	0.03	0.02	1.49	1.33	1.15	0.86	1.14	0.91
1	1000000	0.03	0.02	0.02	0.01	0.02	0.01	1.12	1.03	0.87	0.65	0.86	0.70
2	250000	0.22	0.16	0.16	0.04	0.15	0.10	2.98	2.67	2.17	1.25	2.15	1.41
2	500000	0.15	0.09	0.10	0.02	0.10	0.02	2.36	2.00	1.62	0.93	1.61	1.00
2	1000000	0.08	0.04	0.05	0.01	0.05	0.01	1.64	1.37	1.13	0.69	1.15	0.75
3	250000	0.69	0.34	0.56	0.05	0.52	0.07	5.14	3.92	3.39	1.37	3.33	1.57
3	500000	0.50	0.22	0.43	0.03	0.41	0.03	4.37	3.19	2.71	1.01	2.68	1.10
3	1000000	0.16	0.09	0.13	0.02	0.13	0.02	2.40	1.98	1.62	0.72	1.62	0.77

Table 27: *Cross validation errors of different smoothing methods (SE1 is MSE for cross validation of Method 1 mutiplied by 100, AE1 is MAE for cross validation of Method 1 mutiplied by 100, ... , AE6 is MAE for cross validation of Method 6 mutiplied by 100).*

Overall, 144 tests were performed. The results of these tests show that the MARCIP1 method performs better than other methods for the simulated data (see columns SE4 and AE4 of Table 27). Notably, the Section 3 method performed better, in terms of Mean Squared Errors (column SE2), than L_1 Regularized Median Smoothing method

Set	Population	PER4	COH4	PER6	COH6
1	250000	53.99	52.64	60.73	57.94
1	500000	65.65	53.18	67.47	56.21
1	1000000	70.12	54.53	74.52	58.68
2	250000	48.42	40.98	38.04	42.02
2	500000	31.62	35.27	38.09	39.63
2	1000000	36.95	35.81	54.87	40.43
3	250000	27.02	25.75	49.84	31.72
3	500000	37.89	19.64	39.78	22.77
3	1000000	46.79	25.17	56.36	27.90

Table 28: *Cohort (COH4, COH6 columns) and period errors (PER4, PER6 columns) of MARCIP1 smoothing. Method 4 (columns PER4 and COH4) does not take heteroscedasticity into account and method 6 (columns PER6 and COH6) accounts for it.*

(column SE3) for sets 2 and 3, where cohort and period effects are twice and four times larger than in set 1.

The tests also show that MARCIP1 method found cohort and period effects rather imprecisely (Table 28). The lowest precision is observed for set 1, where the period and cohort effects are the smallest compared to sets 2 and 3. The lowest error is observed for the largest cohort effects which are in set 3 where the error is as low as 19.64%. Thus, according to the simulations, the MARCIP1 method finds cohort effects better than period effects and for both type of effects it finds more pronounced effects more precisely.

Despite low precision in finding period and cohort effects in some cases, the MARCIP1 method exploits the ability to find cohort and period effects remarkably well, which allows the method to estimate the "true" mortality surface significantly better than other methods.