Population Review

Volume 50, Number 2, 2011

Type: Article pp. 122-133

The Convergence of European Mortality in Both Sexes in the Near Future: A Spatio-Temporal Approach

Authors: Alejandro Álvaro-Meca, Ana Debón, Valentín Hernandez, Ruth Gil-Prieto and Ángel Gil de Miguel

Afffiliations: University Rey Juan Carlos (Alvaro-Meca); University of Valencia, Spain (Debón); University Rey Juan Carlos (Hernandez); University Rey Juan Carlos (R. Gil-Prieto); University Rey Juan Carlos (A. Gil de Miguel)

Corresponding author/address: Alejandro Alvaro-Meca, University Rey Juan Carlos, Department of Preventive Medicine and Public Health, Alcorcón, 28922 Spain;
email: alejandro.alvaro@urjc.es

Abstract

Mortality has decreased in all the countries of the European Union during the last century, presenting similar patterns within the change of mortality. Despite these similar trends, there are still considerable differences in the levels of mortality of these countries and between men and women. The aim of this article is to adjust and predict mortality and life expectancy at birth in both sexes within 16 countries of the European Union, modifying the Lee-Carter model by means of the inclusion of a spatial component. Mortality is decreasing within these countries, causing the historical difference between both sexes in the years of the projection to disappear. But differences among the countries of the study in mortality levels remain.

Keywords

Bayesian Lee-Carter model, Europe, mortality, spatial analysis

1 Introduction

Mortality has strongly decreased in all the countries of the European Union during the last century, presenting similar patterns within the change of mortality. Despite these similar trends, there are still considerable differences in the levels of mortality of these countries. During the first half of the 20th century, the drop in mortality was due to a strong drop in infectious diseases that caused a considerable drop in mortality at early ages. However, in spite of this reduction at young ages, there is an increase in mortality caused by cancer and cardiovascular diseases during the same period due to the ageing of the population. During the second half of the 20th century, these diseases became the most common causes of death. However, a common feature in Europe and in other parts of the developed world is the marked difference in mortality between men and women. This difference allows mortality to be studied separately for both sexes.

One of the most common models used for the representation of the evolution of mortality, and also one of the most used today by actuaries and demographers, is the Lee-Carter model (Lee and Carter 1992), This model and its different extensions have been applied by many authors. One outstanding example is the use of the model in the Bayesian framework work of Pedroza (2006). Li and Lee (2005) proposed that the Lee-Carter model can be used to predict mortality for countries that belong to a group, instead of considering them individually. Russolillo et al. (2011) proposed extending the Lee-Carter model to include a geographic rate that modifies mortality for each member of the group and permits its comparison.

Traditionally, disease mapping has been used to describe the geographical distribution of diseases and to identify high risk areas. This helps to draw up more suitable policies and increases scientific and public interest. Recent examples of this growing interest in epidemiology are the works of Deguen et al. 2010, Best and Hansell 2009, and Bacciniet et al. 2008. The presence of longitudinal and spatial information encourages the study and development of new spatio-temporal models that can address this important matter. In this framework, we suggest the extension to the Lee-Carter model to a spatio-temporal model.

The main aim of this article is to adjust and predict within the Bayesian framework mortality and life expectancy at birth in both sexes and in 16 countries in the European Union, with ages that range from 0 to 95 years old. To do this, we modified the Lee-Carter model by means of the inclusion of a spatial component, using WinBUGS (Lunnet et al. 2000) and R (R Development Core Team 2009).

2 Lee-Carter Model

Lee-Carter model, consists in fitting the following function to the mortality ratios

$$q_{xt} = (a_x + b_x k_t + \varepsilon_{xt}),$$

or equivalently to the matrix logarithm

$$log(q_{xt}) = (a_x + b_x k_t + \varepsilon_{xt}) \qquad (1)$$

In this expression q_{xt}, refers to death probability at age x in the year t, a_x and b_x are age-dependent parameters and k_t is a specific mortality rate for each year. The errors ε_{xt}, with mean 0 and variance σ_t^2, reflect the historical influence of each specific age which is not captured by the model.

In Lee (1995), the author points out that nothing ensures that death probabilities in equation 1 exceeded the unit. This may be a problem if we are modeling death probabilities that are solved by performing a logit transformation. We apply the model to the logit of death probabilities q_{xt},

$$log\left(\frac{q_{xt}}{1-q_{xt}}\right) = (a_x + b_x k_t + \varepsilon_{xt})$$

This model is indeterminate since given the solution (a_x,b_x,k_t) any transformation of the type $\left(a_x, \frac{b_x}{c}, cb_x\right)$ or $(a_x - cb_x, b_x, k_t + c)$, for every value of c is also the solution. In order to find out a single unique solution for (a_x,b_x,k_t), Lee and Carter (1992), imposed, two constraints

$$\sum_{t=t_1}^{t_n} k_t = 0 \text{ and } \sum_{x=x_1}^{x_n} b_x = 1$$

The last step of the Lee-Carter method r consists of finding a model for the mortality rates values, $\{\widehat{k_t}\}$. By using Box-Jenkins methodology a good model is obtained by means of the expression,

$$\hat{k}_t = p + \hat{k}_{t-1} + u_t.$$

where p is a constant and $u_t$is a white noise.

2.1 Bayesian Lee-Carter Model: A Spatio-Temporal Approach

The Lee-Carter model with a spatial component that we suggest can be reformulated as follows,

$$\log\left(\frac{q_{xtr}}{1-q_{xtr}}\right) = a_x + b_x k_t + S_r + \varepsilon_{xtr} \qquad (2)$$

where in equation (2), x refers to the number of age to adjust, t to the available year and, r to the number of regions to be included in the model. In addition, a_x ,b_x are vectors of unknown parameters, k_t is an unobserved time series process and S_r is the spatial random effect. ε_{xtr} errors are assumed to be independent and identically distributed according to a normal of mean 0 and common variance σ_t^2.

Under the Bayesian paradigm, the researcher can incorporate his knowledge about the matter he is dealing with as a priori information. Afterwards this information can be combined with the observed data in order to obtain the a posteriori distribution of the parameters about which inference is expected to be carried out. Moreover, the Bayesian estimation requires firstly the likelihood function to be provided. In our study, the Lee-Carter model is extended with the spatial component and the a priori distributions of the parameters of interest. Here, non-informative distributions have been chosen for the parameters a_x, b_x, k_ttogether with non-informative for the variances of parameter σ_t^2. To be precise, we have chosen a distribution $N(0, \sigma_t^2)$ with $\sigma_t^2 \sim Gamma(0, 0.001)$. The initial a priori distributions for the starting point b_0 and k_0 have been assumed to be 1 and 0 respectively. In order to study the spatial dependence

S_T , we chose a conditional autoregressive model $CAR(\sigma_T^2)$ (Besag 1974, Clayton et al. 1993) with $\sigma_T^2 \sim Gamma(0, 0.001)$ as a priori distribution. This approximation, the most common and the computationally simplest, approximates the spatial dependence as a mean of the spatial effect of its nearby areas.

In order to implement the Bayesian model, it is necessary to obtain the a posteriori distribution of all the parameters of the model. However, the posterior distribution inference is analytically intractable. Instead, several MCMC algorithms have been proposed to obtain the posterior distribution of the parameters. To be precise, we used Gibbs sampler (Geman and Geman 1984, Gelfand and Smith 1990) to draw samples from the joint posterior distribution. This algorithm consists in iteratively sampling from the conditional distribution of each of the parameters given assigned values to all the other parameters and the data. We use the software Winbugs to fit the model and perform all the posterior inference.

As we have already mentioned, the last step of the Lee-Carter model consists in fitting a temporal series in the index k_T . In order to carry out predictions of the Lee-Carter model within the Bayesian framework the work of Pedroza (2006) has to be pointed out. Pedroza transformed the original Lee-Carter model and carried out predictions for future years by means of Gibbs sample, following two steps:

First drawing the k_T from a normal random distribution with the correct parameters estimated from the data. Second, given the k_T , drawing the log mortality rates from a normal distribution with corresponding parameters.

Our proposal consists in fitting a temporal series in the index Kt by using the Box-Jenkins methodology. We decided to use the standard adjustment since it provides consistent predictions. However, our model has some disadvantages: 1) the predictions are based on the variation of only one parameter, and 2) the model assumes that the variabilities provided by the geographic component and age are constant.

3 Application

We analyze mortality for any cause in both sexes, and for 16 countries of the European Union (Austria, Belgium, Denmark, , Finland, France, Germany, Holland, Italy, Ireland, Luxembourg, Portugal, Norway, the United Kingdom, Spain, Switzerland and Sweden), by using the Bayesian model of Lee-Carter, and adding a geographical component for the period 1989-2006. Subsequently, we predict mortality and life expectancy at birth for the years 2007 to 2011. Data were obtained from the Human Mortality Database (2009). These data have mortality details by individual age and year for each country. We collected data from 1989 up to 2006 in order to fit the model. We chose 1989 since it set a historic milestone caused by the fall of the Berlin Wall and the unification of Germany.

In order to analyze the Bayesian model, we run three different chains by using 2,500 iterations for the Gibbs sampler, with different over-dispersed starting values. We took the first 500 as burn-in and in the end we obtained a sample for each parameter by selecting the last 2,000 values of each one of the chains. The results, presented here, are based on the combined 6,000 draws from the posterior distribution. The convergence of the chains was checked by using the Gelman-Rubin statistic (Gelman and Rubin 1992) implemented in the R-CODA package (Plummer et al. 2009). Values lower than 1.1 suggest that convergence has been reached. We calculated the Gelman-Rubin statistic for the parameters of the model, and in all of them, values lower than 1.01 are reached, indicating convergence.

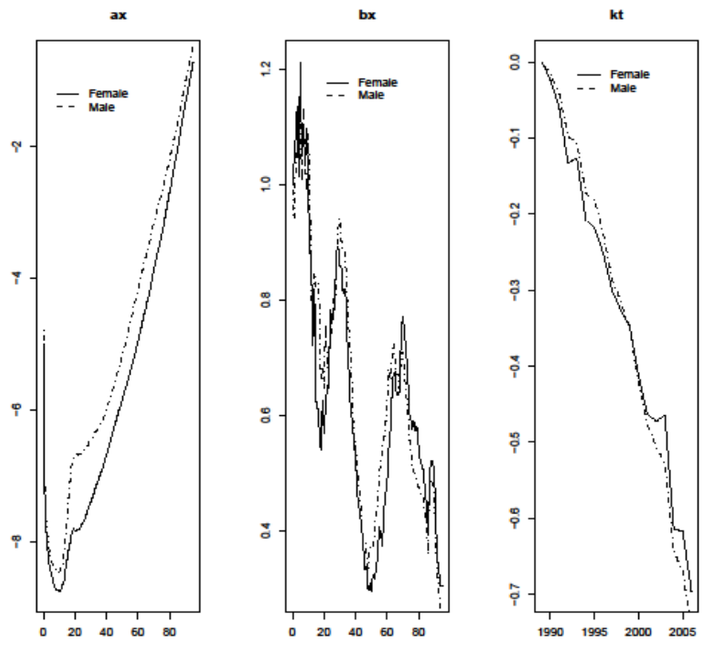

Figure 1: Lee-Carter Parameters

One of the greatest advantages of the Lee-Carter model is that it leads to a simple interpretation of all the parameters of the models by means of a graphical representation. Thus, in Figure 1, it can be noticed how there is a different profile in Europe in mortality for men and women (parameter a_x). Thus, mortality is higher in men than in women. Moreover, this difference is made even greater in ages between 20 and 40 years, appearing in men in what some authors call an ”accident hump“. Parameter b_x represents the decreasing speed of death probability at a certain age in response to changes in$k$ t positive values of b_x for all the age and sexes , indicating that mortality in the 16 countries that we are studying decreases across time. Nevertheless, some authors (Debón et al 2008) have found negative values of b_x for intermediate and old ages. This would indicate an increase of mortality with the passing of time, due to the HIV epidemic. Finally, k_t index shows how mortality is decreasing across time. The peaks that appear in the years 2003 and 2005 have to be pointed out. These increases in mortality can be explained by the heat wave that took place in Europe in 2003 and the virulence of the influenza virus in the year 2005. It is also noteworthy that after the year 2000 male mortality rate has decreased faster than female mortality. This is probably due to the increasing number of women involved in unhealthy habits such as tobacco and alcohol.

In order to continue with the interpretation of the parameters of the fit of the model, in Figure 2 we present the exponentiating spatial effects during the years 1989-2006. Green colour range or values lower than 1 represent countries with a defect of mortality in comparison to its neighbouring countries. On the other hand, brown colour range or values higher than 1 represent those countries with an excess of mortality. Thus, for women, France, Italy, Spain, Sweden, Norway and Switzerland are the countries that represent lower mortality. In contrast, Ireland and Denmark are the countries with a higher excess of mortality. With regard to men, we find a similar pattern with the exception of Holland that shows a lower mortality in comparison to its neighbouring countries. The countries that show a higher excess in

mortality are Portugal and Ireland. Nowadays, historical differences in mortality between men and women seem to be disappearing, tending to be equal or even reversed in the sense that women´s mortality is higher than men´s, as is the case in Holland and Denmark. This is especially due to increased exposure of the former to nicotine poisoning, increasing the appearance of tumors and other diseases directly related to this factor.

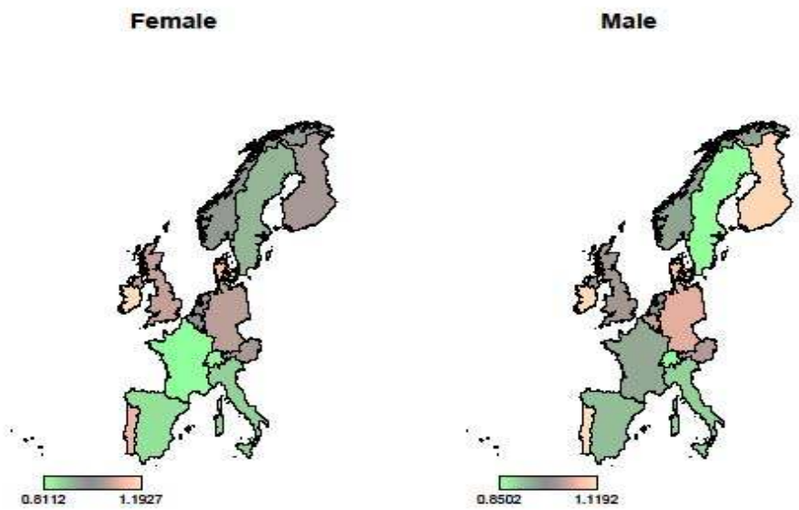

Figure 2: Spatial Effects 1989-2006

With regard to the study of mortality in a country, a common indicator is life expectancy. Life expectancy is an indicator of the overall mortality adjusted by age, which is easy to interpret. Moreover, it allows comparisons between countries. Life expectancy refers to the remaining years of life of one person who has reached a certain age, if these years of life are in the mortality conditions of the current period. However, if mortality rates are expected to decrease in the future, as it is happening, a newborn baby may expect to live longer than the years calculated at birth due to a cohort effect.

In Figure 3, life expectancy at birth for men and women for the year of adjustment 2006 is shown. Women´s life expectancy is 80.64 years for Ireland and 83.77 for France. Thus, France, Italy, Switzerland and Spain, are the countries with a higher life expectancy at birth for women. Moreover, a clear geographic pattern appears that uniquely characterizes Mediterranean, Northern European and Central European countries. However, there are two exceptions. On the one hand, although Portugal is a Mediterranean country, it has the lowest life expectancy. On the other hand, although Sweden is a Northern European country, it has high life expectancy in comparison to its neighbouring countries. With regard to men, life expectancy varies between 75.69 in Portugal and 78.42 in Switzerland, presenting a similar pattern to women´s but with some exceptions. For example, higher life expectancy at birth is found in Sweden and Switzerland. In addition, with regard to women, geographical differences appear among Mediterranean, Northern European and Central European countries.

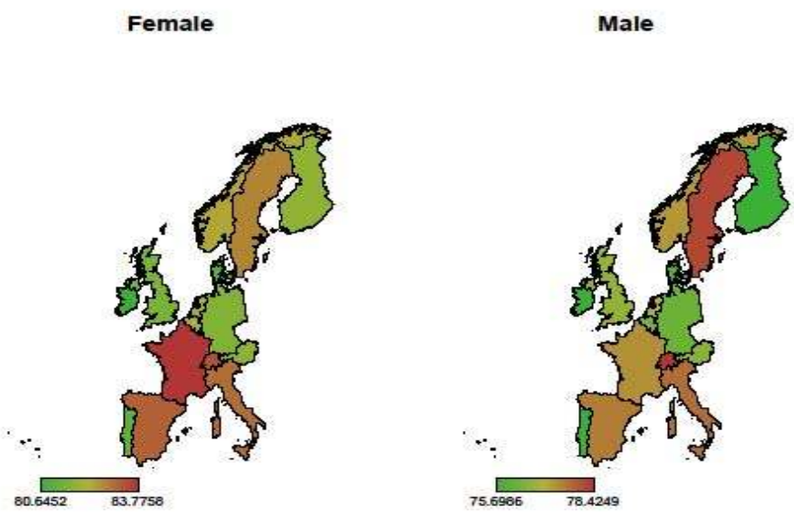

Figure 3: Life expectancy at birth in 2006

3.1 Forecast

As mentioned in Section 2, the last step of the Lee-Carter model consists in predicting mortality for future years, in our case 10 years. In order to carry out predictions, we opted for the classic approach adjusting a time series to k_t index by using Box-Jenkins methodology. Thus, an $ARIMA\,(1,1,0)$, turned out to be the best model for both women and men. Projection of index kt as well as its confidence intervals are shown in Figure 4, thus, in a continuous line we show the projection for women whereas men´s prediction is shown in a discontinuous line, where it can clearly be noticed how mortality decreases in a widespread way in the countries of our study during the years of the adjustment and the years of the projection. However, since year 2000, the trend in mortality decrease has been higher in men than in women.

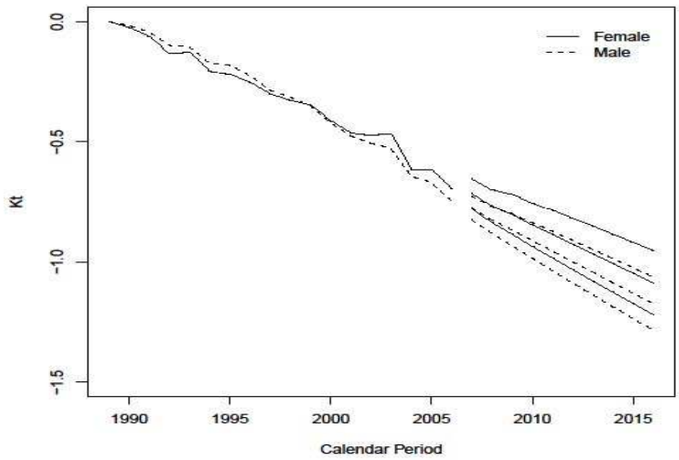

Figure 4: Projection of K_t index

Among the different countries and sexes of our study, mortality tends to approach a joint mortality in the future. This is observed, for instance, in Figure 5, where infant mortality (for men at the right of the figure and for women at the left) seems to be starting to converge towards a joint mortality for all the countries and all the ages of the study. Figure 5 shows that mortality is decreasing faster for men than for women. However, for the last year of the forecast, the model provides very similar estimations for men and women. This is due to the faster decrease in mortality that men are experiencing.

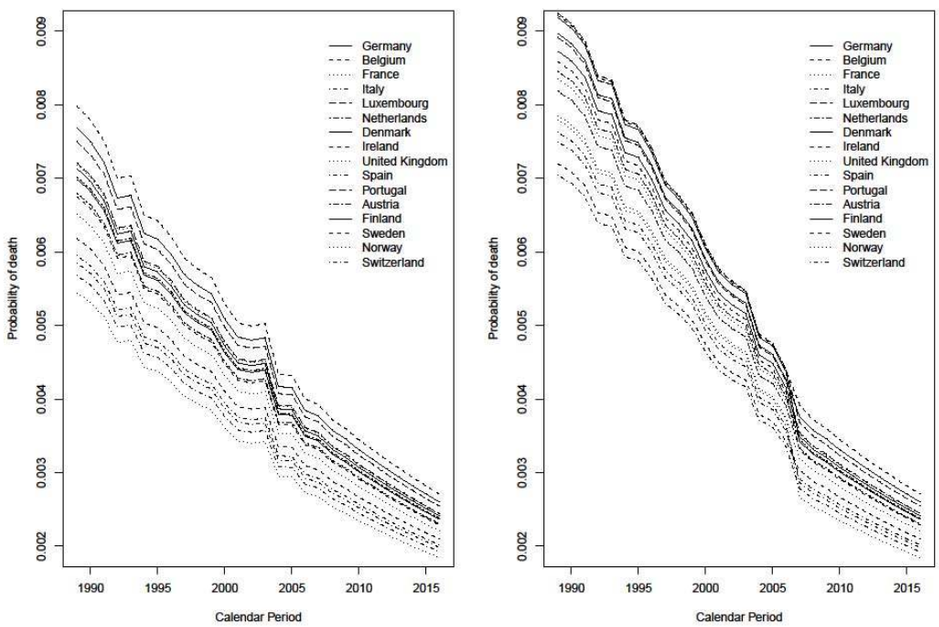

Figure 5: Death probability at birth for women (left) and men (right)

The last predictions that we carried out correspond with the projection of life expectancy at birth for both men and women. In Figure 6 life expectancy at birth for men and women in the year 2006 is shown. Thus, for women (as it happened when we carried out the fit of the model) France is the country with a higher life expectancy at birth, with 85.33 years, followed by Switzerland, Spain and Italy. On the other end of the scale, we find Portugal, Ireland and Denmark which are the countries with a lower life expectancy.

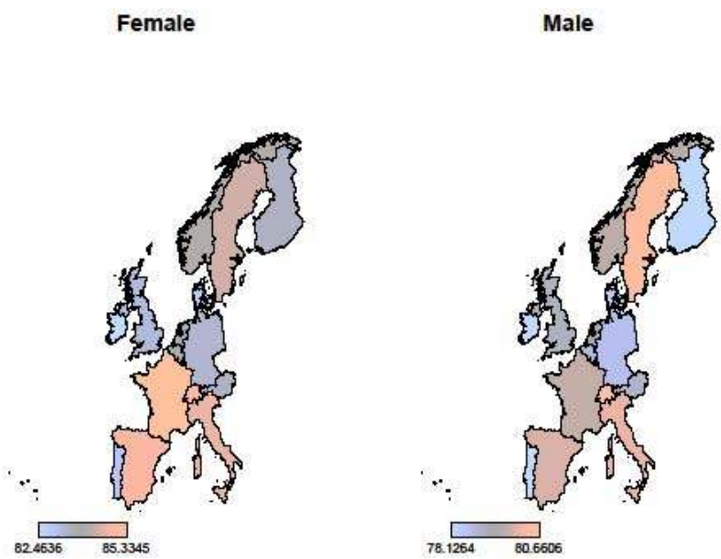

Figure 6: Life expectancy at birth in 2016

With regard to men, Switzerland, Sweden and Italy are the countries with a life expectancy above 80 years, very closely followed by Spain. On the other hand, as with women, Portugal, Ireland and Finland are the countries with a lower life expectancy for men. Thus, a geographic pattern in life expectancy at birth for the 16 countries surfaces.. The Mediterranean countries, France, Switzerland, Spain and Italy together with Sweden, appear with high life expectancy at birth. Moreover, Portugal and Ireland are the countries with lower life expectancy. Finally, Central European and Northern European countries have an intermediate life expectancy between these two groups.

Life expectancy increased in all the studied countries. Moreover, differences between men and women decreased. Thus, in the year 2007, there was a mean difference of 5.21 years in life expectancy at birth between men and women, whereas in 2016, those differences between sexes are estimated to be 4.63 in mean. Another point to stress the greater increase in life expectancy, is that men will increase their life expectancy to a mean of 2.18 years, whereas women will increase 1.61 years from 2007 to 2016.

4 Conclusion

In this study we show, within the Bayesian framework, an adaptation of the Lee-Carter model adding a geographic component in order to predict mortality and life expectancy at birth in 16 countries of the European Union. The main finding is that although mortality is decreasing in both sexes this decrease becomes more apparent in men. This is especially due to the exposure of women to unhealthy habits. Therefore, the historical difference in mortality between both sexes is disappearing in the years of our prediction. Although mortality is decreasing with a similar trend in the different countries of our study, there are still marked differences between them. This may be due to increased access to better and more varied health services.

By working with the Bayesian formulation, we realise that the Lee-Carter model´s formulation is more complex than with the classic approach of the model. Nonetheless, we think that it is important for adjusting and predicting mortality, since it permits the incorporation of all of the sources of variation to the model. A second problem of the Bayesian model is the *a priori* specification of information. If we know the actual information *a priori*, it should be included in the model. However, this information is not always known, therefore, it is necessary to study the *a priori* distributions carefully and carry out a sensitivity analysis to see how strong are the results of the adjustment and the prediction to the choice of the prior distributions.

Another problem of our model is that we assume that the 16 countries of the study share the same pattern. This would not have seemed very realistic if we had considered adjusting dates before 1970, since in Europe great differences existed between some countries. For instance, Spain experienced a dictatorship that resulted in a very high emigration rate, with marked differences between rural and urban areas. And in Germany, Berlin´s partition caused great differences between the oriental and occidental areas. Fortunately, since 1989, all of these differences disappeared completely. Thus, we can consider that the mortality of all the countries that we studied has evolved in a homogeneous way among them. For this reason, we did not create a separate model for each country. Instead, we created a common model for all of the 16 countries that differentiated each country from the others using a spatial random effect.

Our estimated life expectancy for the 16 countries of the study can be compared to the reference study on mortality carried out in Europe by Eurostat (2009). Our model provides very similar estimations to those that Eurostat proposed for Europe. The difference is that the study by Eurostat can work with divisions within the countries, but we were not able to work at that level because the Human Mortality Database does not provide the necessary, detailed information. However, it is encouraging that both the Eurostat (2009) study and our study produced similar results. Our estimations, as a whole, are very plausible given their similarity to those provided by Eurostat. On advantage of our model is that it provides researchers with the ability to forecast in the short, medium and long term.

Future research will adjust mortality by cause of death using the proposed model and expand the findings by incorporating a cohort effect. The model that we have put forth in this study predicts, in the near future, the convergence of mortality in both sexes of the 16 countries of the European Union that we studied.

Acknowledgement

Research by Ana Debón was supported by a grant from Generalitat Valenciana (grant No. GVPRE/2008/103).

References

Baccini, Michelaet al. (2008). Heat effects on mortality in 15 European cities. *Epidemiology*. 19, 711–719.

Besag, Julian (1974). Spatial interaction and the statistical analysis of laticesystemns (with discussion). *Journal Royal Statistical Society*, Series B 36, 192–236.

Best, N. and Anna Louise Hansell (2009). Adjusting for Unmeasured Confounders Through Joint Modeling of Multiple Diseases. *Epidemiology.* 20, 400–410.

Clayton, David G et al (1993). Spatial correlation in ecological analysis. *International Journal Epidemiology* 22, 193–202.

Debón, Ana et al (2008). Modelling and forecasting mortality in Spain. *European Journal of Operation Research* 189, 624–637.

Deguen, Séverineet al. (2010). A Small-area Ecologic Study of Myocardial Infarction, Neighborhood Deprivation, and Sex: A Bayesian Modeling Approach. *Epidemiology* 21, 459–466.

Eurostat (2009). Health statistics – Atlas on mortality in the European Union. ISBN: 978-92-79-08763-9)

Gelfand, A.E. and Adrian F.M. Smith (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association* 85, 398–409.

Gelman, A. and Donald B Rubin (1992). Inference form iterative simulations using muiltiple sequences (with discussion). *Statistical Science* 7, 457–472.

Geman, S. and Donald Geman (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 6, 721–741.

Human Mortality Database (2009). University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at www.mortality.org or www.humanmortality.de (data downloaded on June 1, 2010).

Lee, Ronald D. (2000). The Lee-Carter method for forecasting mortality, with various extensions and applications. *North American Actuarial Journal* 4, 80–91.

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

Li, N. and Ronald D Lee (2005). Coherent mortality forecast for a group of populations: an extension of the Lee-Carter method. *Demography* 42, 575–593.

Lunn, David J et al (2000). WinBUGS – a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing* 10, 325–337.

Pedroza, Claudia (2006). A Bayesian forecasting model: predicting U.S. male mortality. *Biostatistics* 7, 530–550.

Plummer, Martynet al (2009). coda: Output analysis and diagnostics for MCMC. R package version 0.13-4. 9

R Development Core Team (2009). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Russolillo, Maria et al (2011). Extending the Lee-Carter model: a three-way decomposition. *Scandinavian Actuarial Journal*. 2, 96-117