Convergent mortality levels?
Coherent mortality forecasts among industrialized countries

Extended abstract for the 2014 Annual Meeting, Population Association of America

Marie-Pier Bergeron Boucher
University of Southern Denmark, Institute of Public Health
Max Planck Odense Center on the Biodemography of Aging

Vladimir Canudas-Romo
University of Southern Denmark, Institute of Public Health
Max Planck Odense Center on the Biodemography of Aging

James W. Vaupel
Max Planck Institute for Demographic Research
Max Planck Odense Center on the Biodemography of Aging

Odense, Denmark, September 26, 2013

Abstract

Convergence of mortality levels across industrialized countries is observed since the middle of the XXth century. Considering this convergence, forecasting mortality by single-country becomes less proper and a forecasting component common for all countries is necessary. We compare existing forecasting models, and adapt new models, that include common regional trends and asses which model best describes the past and future pattern of mortality in industrialized countries. The forecasting proposal by Li and Lee model and the Compositional Data Analysis (CoDa) approach introduced by Oeppen in the context of forecasting causes of death are compared and combined. Although these two methods have the same goal, their approaches tap different aspect of the "coherent forecasting" problem and we make the hypothesis that a combination of the two might be the optimal methodology.

Key words: Forecast; mortality; industrialized countries; Compositional Data Analysis; Li and Lee model; multiple-decrement

1 Context

Convergence of mortality levels across industrialized countries was observed since the middle of the XXth century (Li and Lee, 2005; White, 2002; Wilson, 2001, 2011). This occurred as a general process, where populations got closer via communication, transportation, trade, technology as well as propensity to disease, without however totally eliminating regional specificities. Considering this convergence, forecasting mortality by single-country becomes then less proper and a forecasting component common for all countries is necessary (Li and Lee, 2005).

Forecasting for a subgroup of a population, as sex, causes of death and even country, has been a challenge for demographers. The reason for questioning this procedures is that subgroup's independent forecasts failed behaving in a coherent way. The different subgroups are often projected separately and then sum up, which 1) tend to increase the divergence between subgroups on the long run, even when using a similar extrapolative procedures (Li and Lee, 2005) and 2), as it has been demonstrated for forecasts by cause of death, the total population forecasts tends to be dominated by an increase or slower decrease of certain subgroups, leading to more pessimistic forecasts (Wilmoth, 1995). Thus, the coherence problem came mainly from an inability to consider the correlation between the subgroups (Oeppen, 2008).

In this perspective, some authors offer solutions. Among the solutions offered for the coherence problem, Li and Lee (2005) suggest modifying the Lee-Carter method, by using a factor for central tendencies for the whole group and a factor for individual-country trends. More recently, Oeppen (2008) suggest abandoning the conventional way to forecast and using Compositional Data Analysis (CoDa), pioneered by Aitchison (1986), in a forecasting context. The CoDa methodology constrains the components of the forecast to vary between limits and as such it addresses the coherence problem between trends of subgroups (Oeppen, 2008).

While the Li and Lee (2005) method has been used to forecast mortality by country in a coherent way, the Oeppen (2008) method has been used to forecast causes of death and has not been applied in the context of mortality by country. The combination of both methods has also never been explored. Even if those two methods have a same goal, their approaches tap different aspect of the problem. Li and Lee (2005) method is using a common factor, representing a general process and the commonalities of historical experience through industrialized countries, while the Oeppen (2008) method suggests to use constrained variable to "force" the variables to vary between two limits and to behave in a coherent way.

The main purpose of this study is to explore the possibilities for coherent forecasts among countries. Both Li and Lee (2005) and Oeppen (2008) methods will be compared and then combined.

2 Data

The data come from the *Human Mortality Database* (HMD, 2013). The HMD offers historical data on mortality for 37 countries. The data series are constructed according to a common protocol, making the HMD an excellent comparative tool. The study will mainly focus on industrialized countries and most specifically on Western European (Austria, Belgium, Denmark, Finland, France, Germany, Iceland, Ireland, Italy, Luxembourg, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland and United Kingdom) and North American (Canada and United States) countries, as well as Japan and Australia.

The number of years with available mortality data differs for each country in the HMD. However, with an exception for Germany, the HMD cover the period 1950-2009 for all the selected countries. This period will be the reference period to forecast mortality.

3 Methods

3.1 Oeppen method: Compositional data analysis

Compositional data (CoDa) is useful to analyze multivariate data, in which the components represent part of a whole. Properties of compositional data is that the components always sum to a constant and carry only relative information. Values for components in compositional data are not free to vary independently, aspect that is manifested in their dependency structure (Pawlowsky-Glahn and Egozcue, 2006). This property renders it an excellent tool for forecasting.

Oeppen (2008) suggested to use the life-table distributions of deaths (d_x) to realize the forecasts, as their sum is always equal to the life-table radix (l_0) and thus are compositional data. To have this life table radix-sum constrain for forecasts by country, a multiple-decrement life table by country has to be constructed. The risk of dying for country i, at age x and for year t for the population of high longevity countries, defined as $m_x^i(t)$, is calculated as:

$$m_x^i(t) = \frac{D(x,t,i)}{\sum_i P(x,t,i)} \tag{1}$$

where D is the number of death and P is the person-years. Here it should be noted that, in the denominator, the person-years are added over all countries, making the death rate m(x,t,i) different from the country specific death rate. From the $m_x^i(t)$, d_x^i can be calculated from a multiple-decrement life table methodology.

As suggested by Oeppen (2008), the forecast steps go as follow:

1. A matrix A of size $T * X * I$ of the multiple-decrement life table deaths (d_x^i), with T rows representing the number of years and $X * I$ columns representing the ages for each country, is then created (Table 1). The sum of each row is summing to the life-table radix.

Table 1: **Example of a matrix A**

	d_0^1	d_1^1	...	d_{110+}^1	d_0^2	d_1^2	...	d_{110+}^2	...	d_0^{21}	d_1^{21}	...	d_{110+}^{21}
1950	-	-	-	-	-	-	-	-	-	-	-	-	-
1951	-	-	-	-	-	-	-	-	-	-	-	-	-
...													
2009	-	-	-	-	-	-	-	-	-	-	-	-	-

2. We obtain a second matrix, B, by subtracting the column geometric means (over years) from the matrix A, by using a CoDa operator. This step centers the matrix, which allows a better visualization of the structure.

3. As most of methodologies are adapted for unconstrained variables, it could be useful to represent our matrix B in the real space, where the variables can vary freely from -∞ to ∞. Aitchison (1986) defined the sample space of compositional data as "simplex". The simplex is restricted space where the variables can only vary from 0 to a given constant (life table radix in our case). To pass from the simplex to what Aitchison defined as the real space, Oeppen (2008) is using the centered-log ratio

(clr):

$$clr(X_{i,j}) = [log(\frac{X_{1,1}}{g(X_{i,1})}), log(\frac{X_{2,1}}{g(X_{i,2})}), log(\frac{X_{3,1}}{g(X_{i,3})})] \quad (2)$$

where $X_{i,j}$ are the different value for component i in each sample-row j of the matrix B and $g(X_{i,j})$ is the row-geometric mean (over ages). We obtain a new transformed matrix, C.

4. At this step, any forecast method can be applied, in theory. In this study, several methods will be tested and will be defined subsequently. Once the forecast method is applied, the centered log-ratio matrix, with the new forecasted sample-row, is constructed. We obtain then a new matrix (C^*) similar to the matrix C, but including the forecasted rows.

5. To transform back the matrix into compositional data, B^*, the inverse centered log-ratio is used.

6. The last step is to add back the geometric means, to obtain the matrix A^*.

3.2 Li and Lee method

The Lee-Carter model uses a stochastic process to forecast mortality. This model summarizes the log of death rates by age (m_x) in terms of vectors α and β on an age dimension and vector κ on a time dimension. The coefficient α is the empirical average of the age profile (Girosi and King, 2007; Lee and Carter, 1992). This model expresses the log of death rate at age x at time t as:

$$log(m(x,t)) = \alpha_x + \beta_x \kappa_t + \epsilon_{xt} \quad (3)$$

where ϵ_{xt} is the error terms. To forecast κ_t, Lee and Carter suggested using the model ARIMA(0,1,0). However, other ARIMA models might be used. For example, Oeppen (2008) used ARIMA(0,2,2) which offers a better fit to the cause of death trends. This might be also the case for forecasts by country.

Li and Lee (2005) modified the Lee-Carter model to forecast groups of a same population, in a coherent way. This model uses a common factor, representing general mortality trends for the whole group of countries.

$$log(m(x,t,i)) = \alpha(x,i) + \beta(x)\kappa(t) + b(x,i)k(t,i) + \varepsilon(x,t,i) \quad (4)$$

where $\beta(x)\kappa(t)$ is the common factor for each population (it is obtained by applying the ordinary Lee-Carter method to the whole group) and $\alpha(x,i)$ is the average log mortality at a given age x and population i. The term $b(x,i)k(t,i)$ represents the difference between the death rates of population i and the rates implied by the common factor (Li and Lee, 2005). This method is "*taking advantage of commonalities in their historical experience and age patterns, while acknowledging their individual differences in levels, age patterns, and trends*" (Li and Lee, 2005, p.590). This method also allows more coherent forecasts when forecasting for disaggregated data.

To assess which model best describes the past and future pattern of mortality in industrialized countries the Li and Lee (2005) and Oeppen (2008) methods will first be compared.

Then different ARIMA models will be introduced within Li and Lee (2005) and Oeppen (2008) methods. Both methods will finally be combined.

4 Preliminary results, the case of France

The following figures show an application of CoDa for a traditional Lee-Carter model using ARIMA (0,1,0) to forecasts $\kappa(t)$, in comparison with a Lee-Carter model. The life expectancy is forecasted until 2040, based on the reference period 1960-2010. Figure 1 shows that the CoDa forecast is more optimistic about future mortality in France than the Lee-Carter model. The forecast for life expectancy at birth in 2040 is 89.69 years with the CoDa method and 86.70 years with the Lee-Carter model. One way to evaluate the performance of a forecast method, is to use an older reference period to forecast actual mortality level. We can observe that the application of CoDa appears to be quite efficient predicting 2010 mortality levels based on trends from 1950 to 1990, as showed in Figure 2.

Figure 1
French Life expectancy at birth,
observed from 1960 to 2010 and forecasted from 2011 to 2040
using the data 1960−2010 under two models

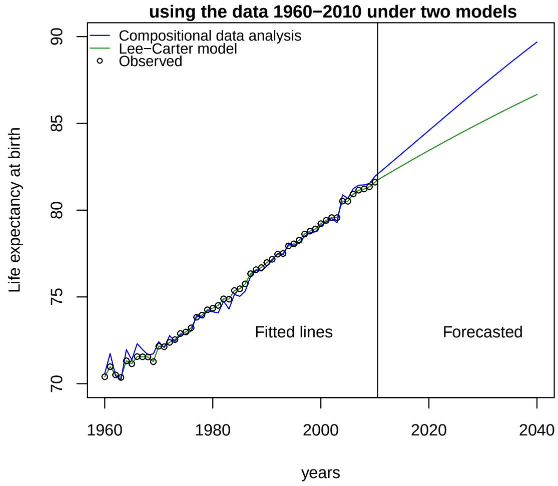
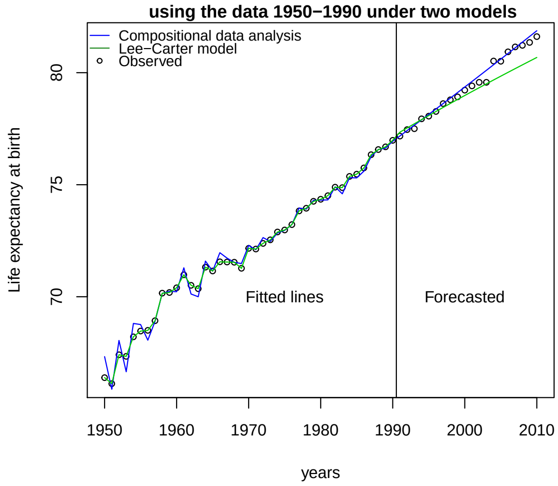


Figure 2
French Life expectancy at birth,
observed from 1950 to 2010 and forecasted from 1991 to 2010
using the data 1950−1990 under two models



References

Aitchison, J. (1986). *The statistical analysis of compositional data.* Chapman & Hall, Ltd., 416p.

Girosi, F. and G. King (2007). *Understanding the Lee-Carter mortality forecasting method.* Harvard University, 33p.

HMD (2013, January). Human mortality database, university of california, berkeley (usa), and max planck institute for demographic research (germany), url: http://www.mortality.org.

Lee, R. D. and L. R. Carter (1992). Modeling and forecasting us mortality. *Journal of the American statistical association 87*(419), 659–671.

Li, N. and R. Lee (2005). Coherent mortality forecasts for a group of populations: An extension of the lee-carter method. *Demography 42*(3), 575–594.

Oeppen, J. (2008, July). Coherent forecasting of multiple-decrement life tables: a test using japanese cause of death data. In *European Population Conference 2008*. European Association for Population Studies.

Pawlowsky-Glahn, V. and J. Egozcue (2006). Compositional data and their analysis: an introduction. *Geological Society, London, Special Publications 264*(1), 1–10.

White, K. M. (2002). Longevity advances in high-income countries, 1955–96. *Population and Development Review 28*(1), 59–76.

Wilmoth, J. R. (1995). Are mortality projections always more pessimistic when disaggregated by cause of death? *Mathematical Population Studies 5*(4), 293–319.

Wilson, C. (2001). On the scale of global demographic convergence 1950–2000. *Population and Development Review 27*(1), 155–171.

Wilson, C. (2011). Understanding global demographic convergence since 1950. *Population and Development Review 37*(2), 375–388.