A fuzzy-random extension of the Lee-Carter mortality prediction model

Jorge de Andrés-Sánchez[a,*] and Laura González-Vila Puchades[b]

[a] *Social and Business Research Laboratory. Business Management Department. Rovira i Virgili University. Av. Universitat 1, 43204 Reus (Spain).*

[b] *Department of Mathematics for Economics, Finance and Actuarial Science. University of Barcelona, Av. Diagonal 690, 08034 Barcelona (Spain).*

* Corresponding author. Tel. 0034 977759832. Fax. 0034 977759810 Email: jorge.deandres@urv.cat.

Abstract. The Lee-Carter model is a useful dynamic stochastic model to represent the evolution of central mortality rates throughout time. This model only considers the uncertainty about the coefficient related to the mortality trend over time but not to the age-dependent coefficients. This paper proposes a fuzzy-random extension of the Lee-Carter model that allows quantifying the uncertainty of both kinds of parameters. As it is commonplace in actuarial literature, the variability of the time-dependent index is modelled as an ARIMA time series. Likewise, the uncertainty of the age-dependent coefficients is also quantified, but by using triangular fuzzy numbers. The consideration of this last hypothesis requires developing and solving a fuzzy regression model. Once the fuzzy-random extension has been introduced, it is also shown how to obtain some variables linked with central mortality rates such as death probabilities or life expectancies by using fuzzy numbers arithmetic. It is simultaneously shown the applicability of our developments with data of Spanish male population in the period 1970-2012. Finally, we make a comparative assessment of our method with alternative Lee-Carter model estimates on 16 Western Europe populations.

Keywords: Lee-Carter model, Fuzzy Numbers, Fuzzy Regression, Fuzzy-random modelling.

1. Introduction

Classical actuarial methods graduate mortality by only taking into account the age of persons without calendar year considerations. Due to the progressive increase of life expectancy in all developed countries, this kind of methods systematically overestimate the mortality rates and, as a consequence, may increase the longevity risk when pricing life annuities.

In the last decades of the 20th century, several papers developed dynamic stochastic approaches for the evolution of mortality rates throughout calendar time and, so, projecting mortality to the future with these models became more accurate. In this way, the method in [26], that we will name LC, is one of the most extended methodologies. The LC model proposed adjusting a linear function to the logarithm of central mortality rates of each year and age, $m_{x,t}$. The coefficients of the linear function depend on the age x whereas the independent variable is a non-observed intensity index k_t associated to the time calendar t. Once the parameters of the model have been adjusted, to make predictions on mortality dynamics it is necessary projecting k_t to the future. It is commonly made with an ARIMA model.

There are two main reasons why the LC model boasts great acceptance. On the one hand, it has been applied in many countries with good results [7,8,10,13,21,23,25,30,37]. Likewise, the LC method is relatively easy to compute in its seminal version, either by using the singular value decomposition (SVD) method or with the approximation to the SVD solution suggested in [26].

Several papers proposed technical extensions to the original LC model as [8,12,13,21,23,31,36]. All these extensions have two common features. Firstly, more

completeness and sophistication of the model suppose more computational effort. Secondly, all of them consider that the age-specific historical influences not captured by the model are due to a stochastic error-term, as the LC model does. However, stochastic variability may not be the unique source of uncertainty since it can also come from fuzziness (e.g. due to incomplete or imprecise information) and, consequently, it can be modelled with Fuzzy Sets Theory tools. In this way, [22] developed two alternative fuzzy formulations of the LC model. The first model considers that all the parameters are fuzzy numbers (FNs) and arithmetical operations are carried out by means of the weakest t-norm. This first approach was object of several refinements in [14,24]. In the second approach, the centres and spreads of the FNs that estimate the parameters of the LC model are supposed to be random variables and are estimated with Bayesian methods. The comparison between the fuzzy and the fuzzy-stochastic models seems to show very similar results, but the second model requires much more computational effort.

Mixing fuzziness and randomness in actuarial modelling is not new. [33] described fuzzy random variables with actuarial applications in view and [19] developed a non-life individual risk model where the number of claims follows a Poisson process and their amount is estimated with a triangular FN (TFN). In a life insurance context, [5,6] used fuzzy random variables for the valuation of life contingencies.

This paper also blends fuzziness and randomness and proposes a fuzzy-random approach of the LC model which is conceptually different to those developed in [22]. We consider that the behaviour of the independent variable k_t follows an ARIMA stochastic process. Likewise, we assume that the variability of the age-specific coefficients is due to fuzziness and it is captured by means of FNs. Under these hypotheses, for a given outcome of the random variable k_t, we will have a concrete result of the central mortality rate, $m_{x,t}$, which will be given by a FN.

In order to adjust the fuzzy coefficients of the logarithm of $m_{x,t}$, we use the model of fuzzy regression (FR) developed by [20], that mixes Ordinary Least Squares (OLS) regression and the FR method by [35], but also allows a non-symmetrical shape for the coefficients.

The rest of the paper is organized as follows. We firstly make a brief review of the LC model. Then, we describe some concepts of FNs and FR that will be necessary to develop our work. Section 4 features our fuzzy-random extension to the LC model, exposes how to project future mortality from this formulation and shows an empirical application for Spanish male population, evaluating both the capability of the model to adjust central mortality rates to sample data and to predict out-of-sample data. In the fifth section, we show how some mortality tables variables can be obtained from the results of previous sections. Subsequently, focused on life expectancies of Spanish male population, we test the capability of our model to predict future values. Section 6 includes a complete comparison of the predictive performance of the proposed method with both the basic LC method in [26] and the fuzzy extension by Koissi and Shapiro in [22]. This comparison shows the advantages of our fuzzy-random extension of the LC model. We finish the work by pointing out the main conclusions and suggesting possible extensions.

2. Overview of the Lee-Carter model

Lee and Carter in [26] proposed modelling the logarithm of the central death rate for each specific age and each year with a linear function. In such a way, if $m_{x,t}$ is the central death rate of a person aged x in the calendar year t, the Lee-Carter (LC) model considers:

$$\ln(m_{x,t}) = a_x + b_x k_t + \varepsilon_{x,t} \quad (1)$$

or, equivalently:

$$m_{x,t} = \exp(a_x + b_x k_t + \varepsilon_{x,t}) \quad (2)$$

where:

- $\exp(a_x)$ is the specific value of the central mortality rate at age *x* regardless of the time calendar *t*,
- b_x quantifies the sensitivity of the central death logarithm rate for age x in year t respect to changes in k_t $\left(\frac{d\ln(m_{x,t})}{dt} = b_x \frac{dk_t}{dt}\right)$,
- k_t is a specific mortality index for each year t that represents the trend of the mortality across time,
- $\varepsilon_{x,t}$ is a random error term, with mean 0 and standard deviation σ_ε , which reflects particular age-specific historical influences not captured by the model.

Notice that whereas the parameters a_x and b_x are age-dependent, the parameter k_t is time-dependent. To estimate the model for a given matrix of rates $m_{x,t}$, the authors seek the least squares solution to the equation (1). This model is undetermined since, given a solution (a_x, b_x, k_t), any transformation of the type $(a_x, b_x/c\,, ck_t)$ or $(a_x + cb_x, b_x, k_t - c)$, $\forall c \in \Re$, is also a solution. In order to avoid this issue [26] introduced the constraints $\sum_x b_x = 1$ and $\sum_t k_t = 0$. So,

the estimations of a_x are simply the averages over time of $ln(m_{x,t})$, i.e.:

$$a_x^* = \frac{\sum_{t=0}^{T} \ln(m_{x,t})}{T+1} \quad (3a)$$

being the considered calendar years $t = 0,1,\dots,T$ and $x = 0,1,\dots,\omega$ the different ages with ω the maximum attainable age.

The model (1) cannot be fitted by ordinary regression techniques because on its right side there are two parameters to be estimated but k_t is unobservable. However, [26] showed that a least squares solution can be obtained by applying singular value decomposition (SVD) to the matrix $Z_{x,t} = \ln(m_{x,t}) - a_x^*$. Alternatively, in their paper (appendix A) it is also proposed an approximation to the SVD solution. As the parameters b_x are assumed to follow the constraint $\sum_x b_x = 1$, the parameter k_t, for each $t = 0,1,\dots,T$, can be fitted as:

$$k_t^* = \sum_x \ln(m_{x,t}) - \sum_x a_x^* \quad (3b)$$

And finally, each b_x can be found by regressing through Ordinary Least Squares (OLS) the linear model $Z_{x,t} = b_x k_t + \varepsilon_{x,t}$. So:

$$b_x^* = \frac{\sum_t Z_{x,t} k_t^*}{\sum_t (k_t^*)^2} = \frac{\sum_t (\ln(m_{x,t}) - a_x^*) k_t^*}{\sum_t (k_t^*)^2} \quad (3c)$$

Once the parameters of the model have been fitted, for $t = 0,1,\dots,T$ and $x = 0,1,\dots,\omega$, the trend of the mortality across time, $\boldsymbol{k_t}$, for $t = T+1,\dots$ can be forecasted with an ARIMA model and related variables, as confidence intervals for $\boldsymbol{k_t}$, life and mortality probabilities or life expectancies, can be obtained.

3. Fuzzy numbers and fuzzy regression

3.1. Fuzzy numbers and their arithmetic

This paper quantifies uncertain quantities as a common type of Fuzzy Number (FN), Triangular Fuzzy Numbers (TFNs), that will be symbolized as $\widetilde{A} = (A, l_A, r_A)$ being A the core of the TFN ($\mu_{\widetilde{A}}(A) = 1$) and l_A and r_A its left and right spreads, respectively. The α-cuts of this kind of FNs are closed and bounded intervals $\forall \alpha \in [0,1]$:

$$A_\alpha = [\underline{A}(\alpha), \overline{A}(\alpha)] = [A - l_A(1-\alpha), A + r_A(1-\alpha)] \quad (4a)$$

The expected interval of a FN $\widetilde{A}$, $e_I(\widetilde{A})$, is a crisp interval that in the case of TFNs is:

$$e_I(\widetilde{A}) = \left[\int_0^1 \underline{A}(\alpha)\, d\alpha, \int_0^1 \overline{A}(\alpha) d\alpha\right] = \left[A - \frac{l_A}{2}, A\frac{r_A}{2}\right] \quad (4b)$$

Let f be a continuous real-valued function of n-real variables x_j, $j = 1,2,\dots,n,$. If x_j are not crisp numbers, but the FNs $\tilde{A}_j$, j=1,2,..,n, f induces the FN $\tilde{B}$ in such a way that $\tilde{B} = f(\tilde{A}_1, \tilde{A}_2, \dots, \tilde{A}_n)$. In order to obtain the α-cuts of $\tilde{B}$, B_α, the results of [9] can be used. If the function f is increasing respect to the first m variables, $m \le n$, and decreasing respect to the last $n - m$ variables:

$$B_\alpha = [\underline{B}(\alpha), \overline{B}(\alpha)] = \\ = \left[f\left(\underline{A}_1(\alpha), \underline{A}_2(\alpha), \dots, \underline{A}_m(\alpha), \overline{A}_{m+1}(\alpha), \overline{A}_{m+2}(\alpha), \dots \overline{A}_n(\alpha)\right), \right. \\ \left. f\left(\overline{A}_1(\alpha), \overline{A}_2(\alpha), \dots \overline{A}_m(\alpha), \underline{A}_{m+1}(\alpha), \underline{A}_{m+2}(\alpha), \dots, \underline{A}_n(\alpha)\right)\right] \quad (5)$$

The result of evaluating non-linear functions with TFNs is not a TFN. In this way, [16] proposed a TFN approximation for any real-valued function, derivable and increasing (decreasing) respect to the first (last) m ($n - m$) variables, built up from the first-order Taylor polynomial expansion from the 1-cut to any α-cut. It can be demonstrated that in (5) $\tilde{B} \approx (B, l_B, r_B)$ where, naming the vector that comprises the centres of $\tilde{A}_j$, $j = 1,2,\dots,n$, $A_C = (A_1, A_2, \dots, A_n)$:

$$B = f(A_C)$$
$$l_B = \sum_{j=1}^{m} \frac{\partial f(A_C)}{\partial x_j} l_{A_j} - \sum_{j=m+1}^{n} \frac{\partial f(A_C)}{\partial x_j} r_{A_j}$$
$$r_B = \sum_{j=1}^{m} \frac{\partial f(A_C)}{\partial x_j} r_{A_j} - \sum_{j=m+1}^{n} \frac{\partial f(A_C)}{\partial x_j} l_{A_j} \quad (6)$$

Arithmetic operations between real numbers can be extended to FNs by using the appropriate real-valued function. Since this work uses TFNs, when this function is linear the result of the arithmetic operation will also be a TFN. Otherwise, the result will be approximated by using (6). So, it is obtained:

- Addition:

$$\tilde{B} = \sum_{j=1}^{n} \widetilde{A}_j = \left(\sum_{j=1}^{n} A_j, \sum_{j=1}^{n} l_{A_j}, \sum_{j=1}^{n} r_{A_j}\right)$$

- Scalar multiplication:

$$\tilde{B} = k\widetilde{A} = \begin{cases} (kA, kl_A, kr_A) & k \ge 0 \\ (kA, |k|r_A, |k|l_A) & k < 0 \end{cases}$$

- Product of two positive TFNs (i.e. their supports are contained within $\Re^+$):

$$\tilde{B} = \widetilde{A}_1 \cdot \widetilde{A}_2 = (A_1A_2, A_2l_{A_1} + A_1l_{A_2}, A_2r_{A_1} + A_1r_{A_2})$$

- Division of two positive TFNs:

$$\tilde{B} = \frac{\widetilde{A}_1}{\widetilde{A}_2} = \left(\frac{A_1}{A_2}, \frac{l_{A_1}}{A_2} + \frac{A_1r_{A_2}}{(A_2)^2}, \frac{r_{A_1}}{A_2} + \frac{A_1l_{A_2}}{(A_2)^2}\right)$$

- Exponential function:

$$\tilde{B} = \exp(\tilde{A}) \approx (\exp(A), \exp(A)l_A, \exp(A)r_A)$$

- Logarithmic function:

$$\tilde{B} = \ln(\tilde{A}) \approx \left(\ln(A), \frac{l_A}{A}, \frac{r_A}{A}\right)$$

3.2. Fuzzy regression model with asymmetric coefficients

This paper uses the fuzzy regression (FR) model of [20] that combines the least squares method with the minimum fuzziness principle in [35]. This type of FR method has been used in financial and actuarial applications like fitting options volatility smile [3,28] or calculating claim reserves [1]. In the actuarial field, a wide survey of FR models can be found in [2].

Let us suppose that for the j-th observation of the sample, $j = 0,1,\dots,n$, the pair of the dependent variable (that may be a FN) and the independent variables (that we suppose crisp) is $(\tilde{y}_j, x_j)_{j=1,2,\dots,n}$ where $x_j = (x_{1,j}, x_{2,j}, \dots, x_{m,j})$, $\tilde{y}_j = \left(y_j, l_{y_j}, r_{y_j}\right)$, $x_{i,j} \in \Re$. Likewise, we suppose a linear relationship and also that the coefficients of the linear function are TFNs $\tilde{a}_i = \left(a_i, l_{a_i}, r_{a_i}\right), i = 0,1,\dots,m$. So:

$$\tilde{y}_j = \tilde{a}_0 + \sum_{i=1}^{m} \tilde{a}_i x_{i,j}$$

and then:

$$\left(y_j, l_{y_j}, r_{y_j}\right) = \left(a_0, l_{a_0}, r_{a_0}\right) + \sum_{i=1}^{m} \left(a_i, l_{a_i}, r_{a_i}\right) x_{i,j}$$

where:

$$y_j = a_0 + \sum_{i=1}^{m} a_i\, x_{i,j}$$

$$l_{y_j} = l_{a_0} + \sum_{\substack{i=1 \\ x_{i,j} \geq 0}}^{m} |x_{i,j}| l_{a_i} + \sum_{\substack{i=1 \\ x_{i,j} < 0}}^{n} |x_{i,j}| r_{a_i}$$

$$r_{y_j} = r_{a_0} + \sum_{\substack{i=1 \\ x_{i,j} \geq 0}}^{m} |x_{i,j}| r_{a_i} + \sum_{\substack{i=1 \\ x_{i,j} < 0}}^{n} |x_{i,j}| l_{a_i}$$

The final objective is obtaining the estimates of $\tilde{a}_i = \left(a_i, l_{a_i}, r_{a_i}\right), i = 0,1,\dots m$, that will be denoted by $\tilde{a}_i^* = \left(a_i^*, l_{a_i^*}, r_{a_i^*}\right)$. Following [20], we implement the following steps:

Step 1. By taking the centres of the dependent variable, $y_j, j = 0,1,\dots,n$, we fit the centres of the fuzzy coefficients $\tilde{a}_i^*$, $a_i, i = 0,1,\dots m$, by using OLS on the expression $y_j = a_0 + \sum_{i=1}^{m} a_i x_{i,j}$. In such a way, we obtain the estimates $\left(a_0^*, a_{1,\dots,}^* a_m^*\right)$. To solve this step we can use (3a)-(3c).

Step 2. We fit the spreads of parameters applying the minimum fuzziness criterion in [35]. So, spread estimates must minimize the uncertainty of the estimated outputs and simultaneously these estimated outputs have to contain the real observations, with a membership level of at least α. If we symbolize the estimates of the spreads as $l_{a_i}^*$ and $r_{a_i}^*$ $i = 0,1,\dots m$, the estimated output for $\tilde{y}_j$ will be $\tilde{y}_j^* = \left(y_{j,}^*, l_{y_j}^*, r_{y_j}^*\right)$, where $y_j^* = a_0^* + \sum_{i=1}^{m} a_i^* x_{i,j}$.

Considering, as in [35], that $\tilde{y}_j \subseteq_\alpha \tilde{y}_j^* \Leftrightarrow y_{j_\alpha} \subseteq y_{j_\alpha}^*$, the spreads $l_{a_i^*}$ and $r_{a_i^*}$, $i = 0,1,\dots m$, minimise for a prefixed level α:

$$\min_{l_{a_i}, r_{a_i}} z = \sum_{j=1}^{n} l_{y_j} + \sum_{j=1}^{n} r_{y_j}$$

And accomplish the constraints:

$$y_{j_\alpha} \subseteq y_{j_\alpha}^*, j = 1,2,\dots,n$$

$$l_{a_i}^*, r_{a_i}^* \geq 0, \quad i = 0,1,\dots,m \tag{7}$$

[11] proposed a rule to choose α when the observations on inputs are crisp. α must reach a compromise between containing observed outputs in $\tilde{y}_j^*$ reasonably well but, likewise, $\tilde{y}_j^*$ must be narrow enough in order to be a useful prediction.

If we name as $\tilde{y}_j^{*\alpha} = \left(y_j^{*\alpha}, l_{yj}^{*\alpha}, r_{y_j}^{*\alpha}\right)$ the estimate of the jth observation of the dependent variable at a given α, we can define the credibility level c_j^α as:

$$c_j^\alpha = \frac{\mu_{\tilde{y}_j^{*\alpha}}(y_j)}{l_{y_j}^{*\alpha} + r_{y_j}^{*\alpha}}$$

Thus, the credibility for the entire sample c^α is $c^\alpha = \sum_{j=0}^{n} \frac{\mu_{\tilde{y}_j^{*\alpha}}(y_j)}{l_{y_j}^{*\alpha} + r_{y_j}^{*\alpha}}$. [11] showed that maximizing c^α is equivalent to solve the following quadratic linear programming problem:

$$\max c^\alpha = -p^0 \alpha^2 + (p^0 - c^0)\alpha, \alpha \in [0,1]$$

where:

$$p^0 = \sum_{j=0}^{n} \frac{1 - \mu_{\tilde{y}_j^{*0}}(y_j)}{l_{y_j}^{*0} + r_{y_j}^{*0}}$$

$$c^0 = \sum_{j=0}^{n} \frac{\mu_{\tilde{y}_j^{*0}}(y_j)}{l_{y_j}^{*0} + r_{y_j}^{*0}}$$

being the solution of this problem:

$$\alpha = \begin{cases} \frac{1}{2}\left(1 - \frac{c^0}{p^0}\right) & c^0 < p^0 \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

Therefore, the process that we follow to fit the fuzzy coefficients consists in implementing Step 2 for $\alpha = 0$. Once the spreads $l_{a_i}^{*0}$ and $r_{a_i}^{*0}$, $i = 0,1,\dots,m$, have been obtained, the optimal value of α, α', will be calculated by using the expression (8). Following [27], the final value of $l_{a_i}^*$ and $r_{a_i}^*$, $i = 0,1,\dots m$ is, simply:

$$l_{a_i}^* = l_{a_i}^{*\alpha'} = \frac{l_{a_i}^{*0}}{1-\alpha'} \text{ and } r_{a_i}^* = r_{a_i}^{*\alpha'} = \frac{r_{a_i}^{*0}}{1-\alpha'}$$

4. Fuzzy-random approach of the Lee-Carter model

4.1. Fuzzy-random fitting of the Lee-Carter model

Our fuzzy-random approach of the LC model considers two different sources of uncertainty:

1. It is supposed that historical influences of each specific age are due to fuzziness in the model structure. As a consequence, both, the coefficient that describes the average age-specific pattern of mortality and the coefficient which reflects the variation in the central death across time, turn into FNs and so $\tilde{a}_x = (a_x, l_{a_x}, r_{a_x})$ and $\tilde{b}_x = (b_x, l_{b_x}, r_{b_x})$.

2. The mortality index k_t follows an ARIMA stochastic process, i.e., k_t is an outcome of a random variable (RV) $\boldsymbol{k_t}$.

Under these assumptions, once the average pattern of mortality, $\tilde{a}_x$, and the decline in mortality, $\tilde{b}_x$, have been estimated we can obtain for an outcome of the RV $\boldsymbol{k_t}$, k_t, the central rate of mortality (and its logarithm) as:

$$\ln(\tilde{m}_{x,t}) = \tilde{a}_x + \tilde{b}_x k_t \quad (9a)$$

where:

$$\ln(\tilde{m}_{x,t}) = \left(\ln(m_{x,t}), l_{\ln(m_{x,t})}, r_{\ln(m_{x,t})}\right) \quad (9b)$$

being:

$$\ln(m_{x,t}) = a_x + b_x k_t \quad (9c)$$

$$l_{\ln(m_{x,t})} = \begin{cases} l_{a_x} + k_t l_{b_x} \text{ for } k_t > 0 \\ l_{a_x} - k_t r_{b_x} \text{ for } k_t \leq 0 \end{cases} \quad (9d)$$

$$r_{\ln(m_{x,t})} = \begin{cases} r_{a_x} + k_t r_{b_x} \text{ for } k_t > 0 \\ r_{a_x} - k_t l_{b_x} \text{ for } k_t \leq 0 \end{cases} \quad (9e)$$

In order to fit the estimates $\tilde{a}_x^* = (a_x^*, l_{a_x}^*, r_{a_x}^*)$, $\tilde{b}_x^* = (b_x^*, l_{b_x}^*, r_{b_x}^*)$ and $k_t^*, t \leq T$, we follow the process described in section 3.2 as follows:

Step 1. By taking the centres of $\ln(\tilde{m}_{x,t})$, $t = 0,1,\dots,T$ and $x = 0,1,\dots,\omega$, we fit the centres of $\tilde{a}_x^*$ and $\tilde{b}_x^*$ and the outcomes of the RVs $\boldsymbol{k_t}$, k_t^*, as we described in section 2. In this step, it is necessary to point out that the observed values of the central rate of mortality (and its logarithms) in which we will base our work are crisp. So, $\ln(\tilde{m}_{x,t}) = \ln(m_{x,t})$.

Step 2. We have to calculate the values $l_{a_x}^{*0}$, $l_{b_x}^{*0}$, $r_{a_x}^{*0}$ and $r_{b_x}^{*0}$ by solving the linear problem (7) for $\alpha = 0$, i.e.:

$$\min_{l_{a_x}, l_{b_x}, r_{a_x}, r_{b_x}} (T+1)\sum_x (l_{a_x} + r_{a_x}) + \sum_t |k_t| \sum_x (l_{b_x} + r_{b_x})$$

subject to:

$a_x^* + b_x^* k_t^* + r_{a_x} + k_t r_{b_x} \geq \ln(m_{x,t})$ for $k_t > 0$

$a_x^* + b_x^* k_t^* + r_{a_x} - k_t l_{b_x} \geq \ln(m_{x,t})$ for $k_t \leq 0$

$a_x^* + b_x^* k_t^* - l_{a_x} - k_t l_{b_x} \leq \ln(m_{x,t})$ for $k_t > 0$

$a_x^* + b_x^* k_t^* - l_{a_x} + k_t r_{b_x} \leq \ln(m_{x,t})$ for $k_t \leq 0$

$b_x^* - l_{b_x} \geq 0$ if $b_x^* \geq 0$

$b_x^* + r_{b_x} < 0$ if $b_x^* < 0$

$l_{a_x}, l_{b_x}, r_{a_x}, r_{b_x} \geq 0,\ x = 0,1,2,\dots,\omega$

Let us remark that the constraints $b_x^* - l_{b_x} \geq 0$ and $b_x^* + r_{b_x} < 0$, ensure that the estimate of the centre of $\tilde{b}_x$, $\tilde{b}_x^*$, will have clearly defined its sign. It will make easier to fit fuzzy-probabilistic confidence intervals for out-of-sample predictions.

Step 3. We obtain the optimal value α' from (8). Finally, the spreads $l_{a_x}^*, l_{b_x}^*$, $r_{a_x}^*$ and $r_{b_x}^*$ are obtained, simply, dividing $l_{a_x}^{*0}, l_{b_x}^{*0}$, $r_{a_x}^{*0}$ and $r_{b_x}^{*0}$ by $1 - \alpha'$.

By using $\tilde{a}_x^*$, $\tilde{b}_x^*$ and k_t^*, it is possible to have a fuzzy estimate for the observed central mortality rates. In fact, from (9a)-(9e) we obtain:

$$\ln(\tilde{m}_{x,t}^*) = \left(\ln(m_{x,t}^*), l_{\ln(m_{x,t}^*)}, r_{\ln(m_{x,t}^*)}\right) = (a_x^*, l_{a_x}^*, r_{a_x}^*) + k_t^*(b_x^*, l_{b_x}^*, r_{b_x}^*) \quad (10a)$$

where:

$$\ln(m_{x,t}^*) = a_x^* + b_x^* k_t^* \quad (10b)$$

$$l_{\ln(m_{x,t}^*)} = \begin{cases} l_{a_x}^* + k_t^* l_{b_x}^* \text{ for } k_t^* > 0 \\ l_{a_x}^* - k_t^* r_{b_x}^* \text{ for } k_t \leq 0 \end{cases} \quad (10c)$$

$$r_{\ln(m_{x,t}^*)} = \begin{cases} r_{a_x}^* + k_t^* r_{b_x}^* \text{ for } k_t^* > 0 \\ r_{a_x}^* - k_t^* l_{b_x}^* \text{ for } k_t^* \leq 0 \end{cases} \quad (10d)$$

and, consequently, $\tilde{m}_{x,t}^* = \exp(\tilde{a}_x^* + \tilde{b}_x^* k_t^*)$. Using the results in section 3.1., $\tilde{m}_{x,t}^*$ can be approximated by a TFN:

$$\tilde{m}_{x,t}^* \approx \left(m_{x,t}^*, l_{m_{x,t}^*}, r_{m_{x,t}^*}\right) = \left(\exp(a_x^* + b_x^* k_t^*), \exp(a_x^* + b_x^* k_t^*) l_{\ln(m_{x,t}^*)}, \exp(a_x^* + b_x^* k_t^*) r_{\ln(m_{x,t}^*)}\right) \quad (11)$$

4.2. Forecasting with the fuzzy-random Lee-Carter model

To forecast future central mortality rates and related variables, it is necessary projecting the values of the index k_t. In our approach, these values are the outcomes of the RVs $\boldsymbol{k_t}$ for each year $t > T$ that actuarial literature commonly fits by an ARIMA(p,1,q) on the data set $\{k_t^*\}, t \leq T$. Subsequently, these projections must be combined with (10a)-(10d) and (11).

[39] developed a framework for predictions that mixes conventional regression and fuzzy parameters. Following those developments, the predicted values of $\boldsymbol{k_t}$, $t > T$, that may be point values or statistical confidence intervals with a linked significance level, will allow obtaining predictions for the central rate of mortality. The values for the central rates of mortality that

we obtain from a point prediction of $\boldsymbol{k_t}$ are FNs given the fuzziness of $\tilde{a}_x$ and $\tilde{b}_x$. If we use a probabilistic confidence interval of $\boldsymbol{k_t}$, the prediction of the central rate of mortality is a fuzzy-probabilistic confidence interval, i.e. a probabilistic interval whose lower and upper bounds are FNs.

We can use three different estimates for $\boldsymbol{k_t}$, $t > T$, and so, forecasted central rates of mortality change:

- If we use the mathematical expectation, $E^*(\boldsymbol{k_t})$, the mathematical expectation of the logarithm of the central rate of mortality, $\ln(\tilde{m}_{x,t})$, is denoted by $\tilde{E}^*\left(\ln(\tilde{m}_{x,t})\right)$, and:

$$\tilde{E}^*\left(\ln(\tilde{m}_{x,t})\right) = \left(E^*\left(\ln(m_{x,t})\right), l_{E^*\left(\ln(\tilde{m}_{x,t})\right)}, r_{E^*\left(\ln(\tilde{m}_{x,t})\right)}\right) = \\ = (a_x^*, l_{a_x}^*, r_{a_x}^*) + E^*(\boldsymbol{k_t})(b_x^*, l_{b_x}^*, r_{b_x}^*) \quad (12a)$$

that can be calculated with (10a)-(10d). So, the central rate of mortality obtained from $E^*(\boldsymbol{k_t})$, $\tilde{E}^*(\tilde{m}_{x,t})$, is:

$$\tilde{E}^*(\tilde{m}_{x,t}) = \exp\left(\tilde{E}^*\left(\ln(\tilde{m}_{x,t})\right)\right) \quad (12b)$$

which can be approximated by a TFN using (11).

- If we estimate $\boldsymbol{k_t}$ by its ε-percentile, $k_t^{*,\varepsilon}$, the fuzzy forecast of the logarithm of the central rate of mortality, $\ln(\tilde{m}_{x,t})$, is denoted by $\ln(\tilde{m}_{x,t}^*)^{\varepsilon}$, being:

  - If $b_x^* - l_{b_x}^* \geq 0$:

$$\ln(\tilde{m}_{x,t}^*)^{\varepsilon} = \left(\ln(m_{x,t}^*)^{\varepsilon}, l_{\ln(m_{x,t}^*)^{\varepsilon}}, r_{\ln(m_{x,t}^*)^{\varepsilon}}\right) = \\ (a_x^*, l_{a_x}^*, r_{a_x}^*) + k_t^{*,\varepsilon}(b_x^*, l_{b_x}^*, r_{b_x}^*) \quad (13a)$$

  - If $b_x^* + r_{b_x}^* < 0$:

$$\ln(\tilde{m}_{x,t}^*)^{\varepsilon} = \left(\ln(m_{x,t}^*)^{\varepsilon}, l_{\ln(m_{x,t}^*)^{\varepsilon}}, r_{\ln(m_{x,t}^*)^{\varepsilon}}\right) = \\ (a_x^*, l_{a_x}^*, r_{a_x}^*) + k_t^{*,1-\varepsilon}(b_x^*, l_{b_x}^*, r_{b_x}^*) \quad (13b)$$

And it can be implemented with (10a)-(10d). The central rate of mortality obtained from $k_t^{*,\varepsilon}$, $\tilde{m}_{x,t}^{*,\varepsilon}$, is:

$$\tilde{m}_{x,t}^{*,\varepsilon} = \exp\left(\ln(\tilde{m}_{x,t}^*)^{\varepsilon}\right) \approx \left(m_{x,t}^{*,\varepsilon}, l_{m_{x,t}^{*,\varepsilon}}, r_{m_{x,t}^{*,\varepsilon}}\right) \quad (13c)$$

and this FN can also be approximated with (11).

Of course, with this procedure, we maintain the fuzzy uncertainty of $\tilde{a}_x^*$ and $\tilde{b}_x^*$ but the probabilistic uncertainty of $\boldsymbol{k_t}$ is reduced to a point estimation.

- If we take for $\boldsymbol{k_t}$ its probabilistic $1-\varepsilon$ confidence interval, $\widetilde{k_t^{*,\varepsilon}} = \left[k_t^{*,\frac{\varepsilon}{2}}, k_t^{*,1-\frac{\varepsilon}{2}}\right]$, following [39], the 1-$\varepsilon$ confidence interval prediction for $\ln(\tilde{m}_{x,t})$ –or $\tilde{m}_{x,t}$–, that we denote $\widetilde{\ln(m_{x,t}^*)^{\varepsilon}}$ - or $\widetilde{\tilde{m}_{x,t}^{*,\varepsilon}}$- is a fuzzy-probabilistic confidence interval.

  - If $b_x^* - l_{b_x}^* \geq 0$, the lower bound of $\widetilde{\ln(m_{x,t}^*)^{\varepsilon}}$ is the FN $\ln(\tilde{m}_{x,t}^*)^{\frac{\varepsilon}{2}} = (a_x^*, l_{a_x}^*, r_{a_x}^*) + k_t^{*,\frac{\varepsilon}{2}}(b_x^*, l_{b_x}^*, r_{b_x}^*)$ whereas the upper bound of $\widetilde{\ln(m_{x,t}^*)^{\varepsilon}}$ is $\ln(\tilde{m}_{x,t}^*)^{1-\frac{\varepsilon}{2}} = (a_x^*, l_{a_x}^*, r_{a_x}^*) + k_t^{*,1-\frac{\varepsilon}{2}}(b_x^*, l_{b_x}^*, r_{b_x}^*)$.

  - If $b_x^* + r_{b_x}^* < 0$, the lower bound of $\widetilde{\ln(m_{x,t}^*)^{\varepsilon}}$ is $\ln(\tilde{m}_{x,t}^*)^{\frac{\varepsilon}{2}} = (a_x^*, l_{a_x}^*, r_{a_x}^*) + k_t^{*,1-\frac{\varepsilon}{2}}(b_x^*, l_{b_x}^*, r_{b_x}^*)$ and the upper bound of $\widetilde{\ln(m_{x,t}^*)^{\varepsilon}}$ is $\ln(\tilde{m}_{x,t}^*)^{1-\frac{\varepsilon}{2}} = (a_x^*, l_{a_x}^*, r_{a_x}^*) + k_t^{*,\frac{\varepsilon}{2}}(b_x^*, l_{b_x}^*, r_{b_x}^*)$

  In both cases, we have to apply (13a)-(13b).

  To calculate the bounds of the fuzzy-probabilistic confidence interval $\widetilde{\tilde{m}_{x,t}^{*,\varepsilon}}$ we implement:

  $\tilde{m}_{x,t}^{*,\frac{\varepsilon}{2}} = \exp\left(\ln(\tilde{m}_{x,t}^*)^{\frac{\varepsilon}{2}}\right)$ and $\tilde{m}_{x,t}^{*,1-\frac{\varepsilon}{2}} = \exp\left(\ln(\tilde{m}_{x,t}^*)^{1-\frac{\varepsilon}{2}}\right)$

  as we do in (13c).

  Let us remark that [26] did not take into account the uncertainty of a_x and b_x but only that from $\boldsymbol{k_t}$. In our model, it is the particular case where $\tilde{a}_x$ and $\tilde{b}_x$ have null spreads and the expressions of the fuzzy-probabilistic $1-\varepsilon$ confidence intervals of the logarithm of the central rate of mortality turn into conventional confidence intervals:

$$\widetilde{\ln(\tilde{m}_{x,t}^*)}^{\varepsilon} = \left[a_x^* + k_t^{*,\frac{\varepsilon}{2}} b_x^*, a_x^* + k_t^{*,1-\frac{\varepsilon}{2}} b_x^*\right] \text{ if } b_x^* \geq 0 \quad (14a)$$

$$\widetilde{\ln(\tilde{m}_{x,t}^*)}^{\varepsilon} = \left[a_x^*, +k_t^{*,1-\frac{\varepsilon}{2}} b_x^*, a_x^* + k_t^{*,\frac{\varepsilon}{2}} b_x^*\right] \text{ if } b_x^* < 0 \quad (14b)$$

4.3. An empirical application of the fuzzy-random Lee-Carter model: the case of Spanish male population

We apply our extension of the LC model to Spanish male population within the period 1970-2000 and we test its out-of-sample performance during 2001-2012. Central mortality rates have been collected from the "Human Mortality Database", [38] (http://www.mortality.org). Ages are grouped in 5 year intervals, except for ages less than one year, for ages from 1 to 5 years and for ages greater or equal to 110 years. The values of the estimates $\tilde{a}_x^*$ and $\tilde{b}_x^*$ are in Table 1, whereas the estimates of the behaviour of $\boldsymbol{k_t}$ are in Figure 1.

The unit root test [15] on $\{k_t^*\}_{t=1970,1971,\ldots,2000}$, suggests that it is $I(1)$. We cannot reject the null hypothesis of one unit root on the level (the Students' t

is -0.202) but we reject that null hypothesis on the first difference because the Students' t is -7.563.

Table 1. Parameters $\tilde{a}_x^*$ and $\tilde{b}_x^*$ for Spanish male population for the period 1970-2000.

	a_x^*			b_x^*		
Age	Centre	Left spread	Right spread	Centre	Left spread	Right spread
[0, 1)	-4.49273	0.30688	0.25300	0.17351	0.00000	0.00000
[1, 5)	-7.48194	0.20455	0.18860	0.12731	0.00000	0.00000
[5,10)	-8.10376	0.21022	0.22686	0.11147	0.00000	0.00000
[10,15)	-8.11329	0.12371	0.11770	0.08472	0.00000	0.01665
[15,19)	-7.17041	0.12666	0.21927	0.03932	0.00000	0.02229
[20,24)	-6.77416	0.12234	0.33523	0.02428	0.00000	0.02060
[25,29)	-6.64539	0.19804	0.43116	0.00113	0.00000	0.00113
[30,34)	-6.47171	0.30079	0.29775	-0.01338	0.01338	0.00000
[35,39)	-6.25015	0.17167	0.12575	0.00356	0.00356	0.00000
[40,44)	-5.89617	0.07598	0.04676	0.02483	0.00243	0.00000
[45,49)	-5.45921	0.06226	0.02487	0.03075	0.00000	0.00520
[50,54)	-5.00591	0.05166	0.05161	0.03864	0.00000	0.00000
[55,59)	-4.55867	0.06395	0.05231	0.04121	0.00000	0.00087
[60,64)	-4.10372	0.08045	0.05451	0.04445	0.00000	0.00000
[65,69)	-3.64283	0.07091	0.04687	0.04724	0.00000	0.00523
[70,74)	-3.15519	0.07888	0.05236	0.05065	0.00030	0.00000
[75,79)	-2.66456	0.09569	0.05943	0.04685	0.00210	0.00000
[80,84)	-2.18259	0.07938	0.05492	0.04257	0.00203	0.00001
[85,89)	-1.72857	0.06846	0.04214	0.03342	0.00580	0.00000
[90,94)	-1.32104	0.06281	0.04782	0.02257	0.00465	0.00028
[95,99)	-0.97328	0.04833	0.02706	0.01436	0.00810	0.00692
[100,104)	-0.68435	0.05068	0.02508	0.00759	0.00636	0.00759
[105,109)	-0.46013	0.04953	0.03142	0.00277	0.00000	0.00277
[110,∞)	-0.31708	0.04852	0.03221	0.00017	0.00000	0.00017

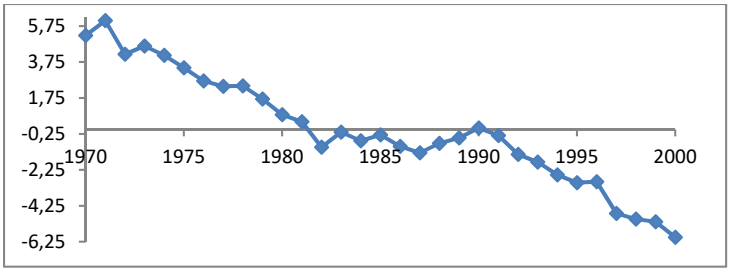


Fig. 1. Evolution of $\boldsymbol{k_t}$ for Spanish male population in the period 1970-2000.

Table 2. Autocorrelation statistics for the time series $\{k_t^*\}_{t=1970,1971,\ldots,2000}$.

Lag	Autocorrelation	Partial Autocorrelation	Q-Statistic	p-value
1	-0.305	-0.305	30.004	0.120
2	0.145	0.005	38.019	0.149
3	-0.01	0.054	38.053	0.283
4	-0.101	-0.109	41.836	0.382
5	0.174	0.123	53.422	0.376
6	-0.023	0.092	53.638	0.498
7	-0.002	-0.004	53.927	0.612
8	-0.112	-0.174	59.353	0.654
9	-0.333	-0.437	11.001	0.276
10	0.014	-0.055	12.450	0.256
11	0.22	-0.152	14.886	0.188
12	0.11	-0.023	15.530	0.215

In Table 2, Ljung and Box Q-statistic suggests that a pure random walk for the first difference is acceptable. So, we model $\boldsymbol{k_t}$ as $\Delta \boldsymbol{k_t}^* = -0.375 + \varepsilon_t^*$, where the estimate for the standard deviation of ε_t^* is 0.68. Figure 2 represents the evolution that we predict for $\boldsymbol{k_t}$ for years 2001-2012 which has been elaborated by using the bootstrapping procedure for ARIMA time series described in [29].

We now check the capability of our extension of the LC model to fit the central rate of mortality, $m_{x,t}$, into the sample used to adjust the coefficients, $t = 1970,1971,\ldots,2000$ but also its performance in out-of-sample predictions at t= 2001,2002,…,2012. We measure this capability with the membership level that the actual central mortality rate $m_{x,t}$ has in its fuzzy estimate $\tilde{m}_{x,t}^*$, $\mu_{\tilde{m}_{x,t}^*}(m_{x,t})$. Figure 3 shows the average of grades of membership, for all age groups, for the period 1970-2000, i.e., $\overline{\mu_t} = \frac{\sum_x \mu_{\tilde{m}_{x,t}^*}(m_{x,t})}{N}$, with N the number of age groups that have been considered ($N = 24$). Likewise, Figure 4 represents, for $t = 2001,\ldots,2012$, the values of $\overline{\mu_t} = \frac{\sum_x \mu_{\tilde{E}^*(\tilde{m}_{x,t})}(m_{x,t})}{N}$, where the central rate of mortality has been forecasted by using $E^*(\boldsymbol{k_t})$, and so, with (12a)-(12b).

Figure 3 shows that the mean grade of membership until the middle of the 80s oscillates, depending on the year, between 0.4 and 0.8. Subsequently, $\overline{\mu_t}$ remains always around 0.6. In Figure 4, where we also predict central mortality rates with (12a)-(12b), we can check that with the exception of 2003 and 2005, the average grade of membership of the real observed central rates of mortality in $\tilde{E}^*(\tilde{m}_{x,t})$ is at least 0.4. Therefore, it can be said that the capability of the model to fit the central mortality rates in the sample as well as to extrapolate them for a period of more than 10 years is reasonably good.

Table 3 shows the TFN predictions for the central mortality rates of year 2010 that come from $E^*(\boldsymbol{k_t})$ and from the bounds of $\widehat{k_t^{*,10\%}} = [k_t^{*,5\%}, k_t^{*,95\%}]$, i.e. we forecast the mathematical expectation and the 10% fuzzy-probabilistic confidence interval of $m_{x,2010}$. For example, if we consider the age group [30,34):

$\tilde{E}^*(\tilde{m}_{[30,34),2010}) = (0.00176, 0.00053, 0.00087)$

which means that the forecasted mean of the central rate of mortality in the year 2010 is approximately 0.00176, although it can vary between 0.00123 and 0.00263. Likewise, $\widehat{\tilde{m}_{[30,34),2010}^{*,10\%}}$ has as lower and upper bounds, $\tilde{m}_{[30,34),2010}^{*,5\%}$ and $\tilde{m}_{[30,34),2010}^{*,95\%}$, the TFNs (0.00168, 0.00051, 0.00075) and (0.00184, 0.00055, 0.00100), respectively, i.e. the real central mortality rate in the year 2010 is contained, with a probability of 90%, between approximately 0.00168, in the most optimistic scenario, and approximately 0.00184, in the most pessimistic scenario. If we were using the basic LC model, which only takes into account the uncertainty related to the index $\boldsymbol{k_t}$, the results would not be FNs but the real numbers: 0.00176 for the mathematical expectation and [0.00168, 0.00184], for its 90% confidence interval.

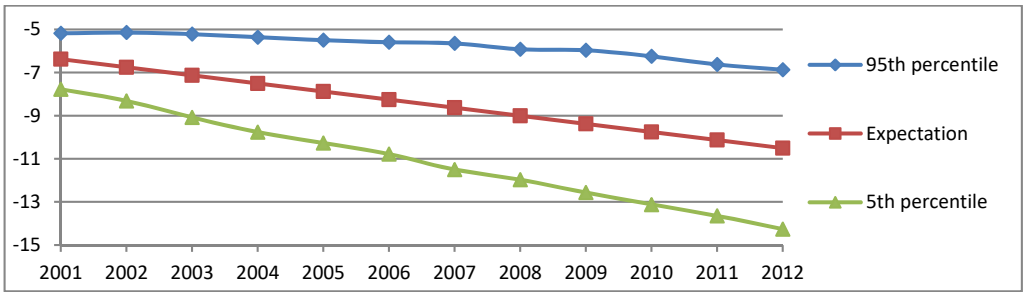


Fig. 2. Estimation of the evolution of $\boldsymbol{k_t}$ for Spanish male population in the period 2001-2012.

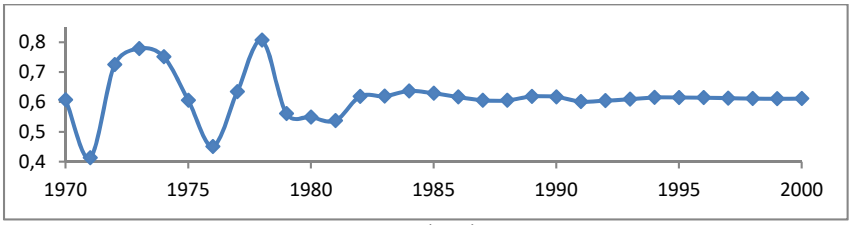


Fig. 3. Values of $\bar{\mu}_t = \frac{\sum_x \mu_{\widetilde{m}^*_{x,t}}(m_{x,t})}{N}$ in the period 1970-2000.

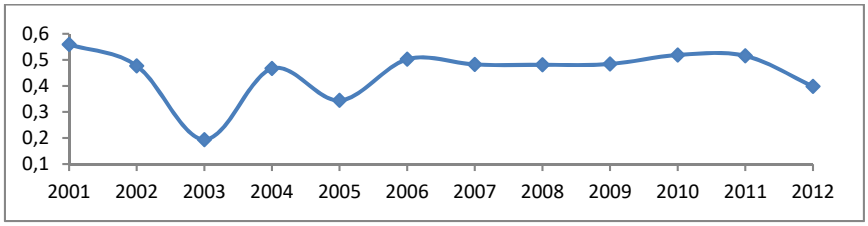
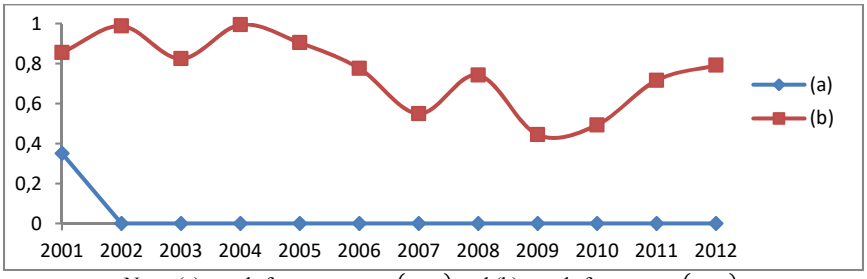


Fig. 4. Values of $\bar{\mu}_t = \frac{\sum_x \mu_{\tilde{E}^*(\widetilde{m}_{x,t})}(m_{x,t})}{N}$ in the period 2001-2012.

Table 3. TFN approximations of the estimates of $\tilde{E}^*(\widetilde{m}_{x,t})$ and the bounds of $\widetilde{m}^{*,10\%}_{x,t}$ ($\widetilde{m}^{*,5\%}_{x,t}$ and $\widetilde{m}^{*,95\%}_{x,t}$) for year 2010.

	$\tilde{E}^*(\widetilde{m}_{x,t})$			$\widetilde{m}^{*,5\%}_{x,t}$			$\widetilde{m}^{*,95\%}_{x,t}$		
Age	Centre	Left spread	Right spread	Centre	Left spread	Right spread	Centre	Left spread	Right spread
[0, 1)	0.00206	0.00063	0.00052	0.00115	0.00035	0.00029	0.00379	0.00116	0.00096
[1, 5)	0.00016	0.00003	0.00003	0.00011	0.00002	0.00002	0.00025	0.00005	0.00005
[5,10)	0.00010	0.00002	0.00002	0.00007	0.00001	0.00002	0.00015	0.00003	0.00003
[10,15)	0.00013	0.00004	0.00002	0.00010	0.00003	0.00001	0.00018	0.00004	0.00002
[15,19)	0.00052	0.00018	0.00011	0.00046	0.00019	0.00010	0.00060	0.00016	0.00013
[20,24)	0.00090	0.00029	0.00030	0.00083	0.00033	0.00028	0.00098	0.00025	0.00033
[25,29)	0.00129	0.00037	0.00055	0.00128	0.00038	0.00055	0.00129	0.00037	0.00056
[30,34)	0.00176	0.00053	0.00087	0.00168	0.00051	0.00075	0.00184	0.00055	0.00100
[35,39)	0.00186	0.00032	0.00047	0.00184	0.00032	0.00049	0.00189	0.00032	0.00045
[40,44)	0.00216	0.00016	0.00015	0.00199	0.00015	0.00016	0.00235	0.00018	0.00015
[45,49)	0.00315	0.00036	0.00008	0.00284	0.00037	0.00007	0.00351	0.00033	0.00009
[50,54)	0.00460	0.00024	0.00024	0.00403	0.00021	0.00021	0.00526	0.00027	0.00027
[55,59)	0.00701	0.00051	0.00037	0.00610	0.00046	0.00032	0.00810	0.00056	0.00042
[60,64)	0.01070	0.00086	0.00058	0.00922	0.00074	0.00050	0.01251	0.00101	0.00068
[65,69)	0.01651	0.00201	0.00077	0.01409	0.00197	0.00066	0.01949	0.00202	0.00091
[70,74)	0.02602	0.00205	0.00144	0.02194	0.00173	0.00124	0.03107	0.00245	0.00169
[75,79)	0.04410	0.00422	0.00352	0.03766	0.00360	0.00327	0.05196	0.00497	0.00377
[80,84)	0.07445	0.00592	0.00556	0.06450	0.00513	0.00526	0.08642	0.00687	0.00584
[85,89)	0.12816	0.00877	0.01264	0.11452	0.00784	0.01353	0.14409	0.00986	0.01129
[90,94)	0.21414	0.01404	0.01994	0.19846	0.01321	0.02159	0.23176	0.01497	0.01781
[95,99)	0.32848	0.03803	0.03485	0.31297	0.04353	0.04175	0.34543	0.03162	0.02684
[100,104)	0.46845	0.05871	0.04225	0.45663	0.06889	0.05097	0.48107	0.04750	0.03266
[105,109)	0.61439	0.05071	0.03279	0.60868	0.05592	0.03249	0.62038	0.04519	0.03311
[110,∞)	0.72704	0.03900	0.04508	0.72661	0.03941	0.04506	0.72748	0.03858	0.04511



Note: (a) stands for $\mu_{\tilde{E}^*(\tilde{m}_{[0,1),t})}(m_{x,t})$ and (b) stands for $\mu_{\tilde{m}^{*,95\%}_{[0,1),t}}(m_{x,t})$.

Fig. 5. Membership levels $\mu_{\tilde{E}^*(\tilde{m}_{[0,1),t})}(m_{x,t})$ and $\mu_{\tilde{m}^{*,95\%}_{[0,1),t}}(m_{x,t})$ in 2001-2012.

Figure 5 depicts the membership level of true observed values of $m_{[0,1),t}$ for the period 2001-2012 in the TFNs $\tilde{E}^*(\tilde{m}_{[0,1),t})$ and $\tilde{m}^{*,95\%}_{[0,1),t}$. Such important age group is not well predicted when the mathematical expectation of the parameter $\boldsymbol{k_t}$ is used because from the year 2002 on the forecasted rates never contain the real values, i.e. their grade of membership is always 0. Nevertheless, it does not happen when considering the upper bound of the fuzzy-probabilistic 90% confidence interval of $\tilde{m}^{*}_{[0,1),t}{}^{95\%}$. In this case, membership levels are never lower than 0.4. We statistically test the capability of $\tilde{E}^*(\tilde{m}_{x,t})$ to predict actual central mortality rates for each year t = 2001, 2002,…,2012. Table 4 shows the results. Following [4], it is desirable that the observed rates $m_{x,t}$ attain membership levels of at least 0.5, in the fuzzy prediction $\tilde{E}^*(\tilde{m}_{x,t})$ i.e, $\mu_{\tilde{E}^*(\tilde{m}_{x,t})}(m_{x,t})\geq 0.5$. So, for each year t = 2001, 2002,…,2012 we implement a Wilcoxon rank test for the null hypothesis that the median value of $\mu_{\tilde{E}^*(\tilde{m}_{x,t})}(m_{x,t})$ is 0.5. It can be seen in Table 4 that only in years 2003 and 2012 the median of $\mu_{\tilde{E}^*(\tilde{m}_{x,t})}(m_{x,t})$ is under 0.5 and the null hypothesis is rejected at standard significant levels.

Table 4. Assessment of the capability prediction of $\tilde{E}^*(\tilde{m}_{x,t})$ with a Wilcoxon rank test in 2001-2012.

Year	W	Median	Mean	Year	W	Median	Mean
2001	92^{*}	0.602	0.656	2007	93	0.414	0.381
2002	141	0.488	0.643	2008	141	0.438	0.256
2003	77^{**}	0.181	0.439	2009	59^{**}	0.534	0.246
2004	140	0.454	0.431	2010	122	0.462	0.233
2005	117	0.165	0.337	2011	67^{***}	0.521	0.255
2006	61^{***}	0.528	0.320	2012	62^{***}	0.202	0.182

Notes: (1) W stands for the value of Wilcoxon rank test statistic. (2) "*", "**" and "***" stand for the rejection of the null hypothesis that the median value of $\mu_{\tilde{E}^*(\tilde{m}_{x,t})}(m_{x,t})$ is 0.5 with a significance level of 10%, 5% and 1% respectively.

5. Forecasting life expectancy with the fuzzy-random Lee-Carter model

5.1. *Calculating life expectancy from fuzzy estimates of central mortality rates*

We now compute probabilities of death or survival and life expectancies after calculating estimates of central mortality rates. Let us denote the width of the age group as n_x years.

To obtain the probability that a person in the age group x, at calendar year t, does not reach the following age group, ${}_{n_x}q_{x,t}$,we have to take into account that it is a function of $m_{x,t}$.

$$ {}_{n_x}q_{x,t} = \frac{n_x m_{x,t}}{1+n_x\left(1-\gamma_{x,t}\right)m_{x,t}} \tag{15}$$

where $\gamma_{x,t} \in [0,1]$ is the average fraction of the n_x-year period lived by those who died in that period and we will suppose that this coefficient is fixed beforehand. Given that:

$$\frac{\partial {}_{n_x}q_{x,t}}{\partial m_{x,t}} = \frac{n_x}{\left[1+n_x\left(1-\gamma_{x,t}\right)m_{x,t}\right]^2} > 0$$

by considering (15) and bearing in mind that ${}_nx\tilde{q}^*_{x,t}{}_\alpha$ is a prediction of a probability, i.e. ${}_nx\tilde{q}^*_{x,t}{}_\alpha \subseteq [0,1]$, we can obtain the α-cuts of ${}_nx\tilde{q}^*_{x,t}$, ${}_nx\tilde{q}^*_{x,t}{}_\alpha$:

$$
\begin{aligned}
{}_nx\tilde{q}^*_{x,t}{}_\alpha &= \left[\underline{{}_nx q^*_{x,t}}(\alpha), \overline{{}_nx q^*_{x,t}}(\alpha)\right] \\
&= \left[\max\left\{0, \frac{n_x\left[m^*_{x,t} - l_{m^*_{x,t}}(1-\alpha)\right]}{1+n_x\left(1-\gamma_{x,t}\right)\left[m^*_{x,t} - l_{m^*_{x,t}}(1-\alpha)\right]}\right\},\right. \\
&\quad \left.\min\left\{1, \frac{n_x\left[m^*_{x,t} + r_{m^*_{x,t}}(1-\alpha)\right]}{1+n_x\left(1-\gamma_{x,t}\right)\left[m^*_{x,t} + r_{m^*_{x,t}}(1-\alpha)\right]}\right\}\right]
\end{aligned}
$$

It may be useful to obtain a triangular approximation for ${}_nx\tilde{q}^*_{x,t}$, ${}_nx\tilde{q}^*_{x,t} \approx \left({}_nx q^*_{x,t}, l_{{}_nx q^*_{x,t}}, r_{{}_nx q^*_{x,t}}\right)$, with:

$${}_nx q^*_{x,t} = \frac{n_x m^*_{x,t}}{1+n_x\left(1-\gamma_{x,t}\right)m^*_{x,t}} \tag{16a}$$

In order to obtain the support of ${}_nx\tilde{q}^*_{x,t}$, we have to take into account that it is a probability and so, its support must be within the interval [0, 1]. Then:

$$l_{{}_nx q^*_{x,t}} = \min\left\{{}_nx q^*_{x,t}, \frac{n_x l_{m^*_{x,t}}}{\left[1+n_x\left(1-\gamma_{x,t}\right)m_{x,t}\right]^2}\right\} \tag{16b}$$

and:

$$r_{{}_nx q^*_{x,t}} = \min\left\{1 - {}_nx q^*_{x,t}, \frac{n_x r_{m^*_{x,t}}}{\left[1+n_x\left(1-\gamma_{x,t}\right)m_{x,t}\right]^2}\right\} \tag{16c}$$

To determine the probability that a person in the age group x, at calendar year t, reaches the following age group, ${}_nx p_{x,t}$, from the crisp relationship ${}_nx p_{x,t} = 1 - {}_nx q_{x,t}$, under fuzziness we state ${}_nx\tilde{p}_{x,t} = 1 - {}_nx\tilde{q}_{x,t}$ where:

$${}_nx p^*_{x,t}{}_\alpha = \left[\underline{{}_nx p^*_{x,t}}(\alpha), \overline{{}_nx p^*_{x,t}}(\alpha)\right] = \left[1 - \underline{{}_nx q^*_{x,t}}(\alpha), 1 - \overline{{}_nx q^*_{x,t}}(\alpha)\right]$$

From (15a)-(15c), ${}_nx\tilde{p}^*_{x,t} \approx \left({}_nx p^*_{x,t}, l_{{}_nx p^*_{x,t}}, r_{{}_nx p^*_{x,t}}\right)$:

$${}_nx p^*_{x,t} = 1 - {}_nx q^*_{x,t};\ l_{{}_nx p^*_{x,t}} = l_{{}_nx q^*_{x,t}};\ r_{{}_nx p^*_{x,t}} = r_{{}_nx q^*_{x,t}}$$

The life expectancy of a person in the age group x, at calendar year t, $e_{x,t}$ can be calculated with the expression:

$$e_{x,t} = \sum_{i\geq x}\prod_{x\leq j<i}\left(1 - {}_nj q_{j,t}\right)\left[n_i - \left(n_i - \gamma_{i,t}\right){}_ni q_{i,t}\right] \tag{17}$$

As:

$$
\begin{aligned}
\frac{\partial e_{x,t}}{\partial {}_ni q_{i,t}} = -\left\{\prod_{x\leq j<i}\left(1 - {}_nj q_{j,t}\right)\left[n_i - \left(n_i - \gamma_{i,t}\right)\right]\right. \\
\left. + \sum_{s>i}\prod_{\substack{x\leq j<s \\ j\neq i}}\left(1 - {}_nj q_{j,t}\right)\left[n_s - \left(n_s - \gamma_{s,t}\right){}_ns q_{s,t}\right]\right\}
\end{aligned}
$$

it turns out that $e_{x,t}$ is a decreasing function of ${}_ni q_{i,t}$ (and so, of its linked central mortality rate).

By evaluating (17) with ${}_ni\tilde{q}^*_{i,t}$, we will obtain a fuzzy estimate for $e_{x,t}$, $\tilde{e}^*_{x,t}$. Moreover, it is straightforward to see that its α –cuts, $e^*_{x,t}{}_\alpha$, are:

$$e^*_{x,t}{}_\alpha = \left[\underline{e^*_{x,t}}(\alpha), \overline{e^*_{x,t}}(\alpha)\right]$$

$$
\begin{aligned}
&= \left[\sum_{i\geq x}\prod_{x\leq j<i}\left(1 - \overline{{}_nj q_{j,t}}(\alpha)\right)\left[n_i - \left(n_i - \gamma_{i,t}\right)\overline{{}_nj q_{j,t}}(\alpha)\right],\right. \\
&\quad \left.\sum_{i\geq x}\prod_{x\leq j<i}\left(1 - \underline{{}_nj q_{j,t}}(\alpha)\right)\left[n_i - \left(n_i - \gamma_{i,t}\right)\underline{{}_ni q_{i,t}}(\alpha)\right]\right]
\end{aligned}
$$

A TFN approximation of $\tilde{e}^*_{x,t}$, $\tilde{e}^*_{x,t} \approx \left(e^*_{x,t}, l_{e^*_{x,t}}, r_{e^*_{x,t}}\right)$, can be obtained by using (6):

$$e^*_{x,t} = \sum_{i\geq x}\prod_{x\leq j<i}\left(1 - {}_nj q^*_{j,t}\right)\left[n_i - \left(n_i - \gamma_{i,t}\right){}_ni q^*_{i,t}\right] \tag{18a}$$

$$
\begin{aligned}
l_{e^*_{x,t}} = \sum_{i\geq x}\left\{\prod_{x\leq j<i}\left(1 - {}_nj q^*_{j,t}\right)\left(n_i - \gamma_{i,t}\right)\right. \\
\left. + \sum_{s>i}\prod_{\substack{x\leq j<s \\ j\neq i}}\left(1 - {}_nj q^*_{j,t}\right)\left[n_s - \left(n_s - \gamma_{s,t}\right){}_ns q^*_{s,t}\right]\right\} r_{{}_ni q^*_{i,t}}
\end{aligned} \tag{18b}
$$

$$
\begin{aligned}
r_{e^*_{x,t}} = \sum_{i\geq x}\left\{\prod_{x\leq j<i}\left(1 - {}_nj q^*_{j,t}\right)\left(n_i - \gamma_{i,t}\right)\right. \\
\left. + \sum_{s>i}\prod_{\substack{x\leq j<s \\ j\neq i}}\left(1 - {}_nj q^*_{j,t}\right)\left[n_s - \left(n_s - \gamma_{s,t}\right){}_ns q^*_{s,t}\right]\right\} l_{{}_ni q^*_{i,t}}
\end{aligned} \tag{18c}
$$

Of course, if in (18a)-(18c) we take as a prediction of the index $\boldsymbol{k_t}$ its mathematical expectation, $k^*_t = E^*(\boldsymbol{k_t})$, we will obtain a fuzzy estimate of life expectation that we symbolize as $\tilde{E}^*\left(\tilde{e}^*_{x,t}\right)$.

If the prediction of the mortality trend comes from its probabilistic confidence interval, $\widetilde{k^{*,\varepsilon}_t} = \left[k^{*,\frac{\varepsilon}{2}}_t, k^{*,1-\frac{\varepsilon}{2}}_t\right]$, we can built up a fuzzy-probabilistic confidence interval of the life expectancy $\widetilde{e^{*,\varepsilon}_{x,t}} = \left[e^{*,\frac{\varepsilon}{2}}_{x,t}, e^{*,1-\frac{\varepsilon}{2}}_{x,t}\right]$. In the common case where the sensitivity of the central rate of mortality respect to changes in the index $\boldsymbol{k_t}$ is strictly positive, i.e. $\tilde{b}^*_x > 0$ ($b^*_x - l_{b^*_x} \geq 0$), $k^{*,\frac{\varepsilon}{2}}_t$ will determinate $e^{*,1-\frac{\varepsilon}{2}}_{x,t}$, whereas $k^{*,1-\frac{\varepsilon}{2}}_t$ will define the lower life expectancy $e^{*,\frac{\varepsilon}{2}}_{x,t}$.

5.2. *Predicting life expectancies of Spanish male population in 2001-2012*

Tables 5 and 6 show the estimates for the mean value and the 90% confidence fuzzy-probabilistic interval of the life expectancy for the age groups [0, 1) and [65, 69) during the period 2001-2012. These ages are significantly important because they are considered in order to quantify life expectancy at birth and at retirement, respectively. If only the centres of the fuzzy life expectancies are considered, predictions that come from the basic LC method are found. So, for example, the point estimate for $e_{[0,1),2012}$ is 78.19 years and the 90% confidence probabilistic interval is [76.56, 79.75] years. The fuzzy-random extension of the LC

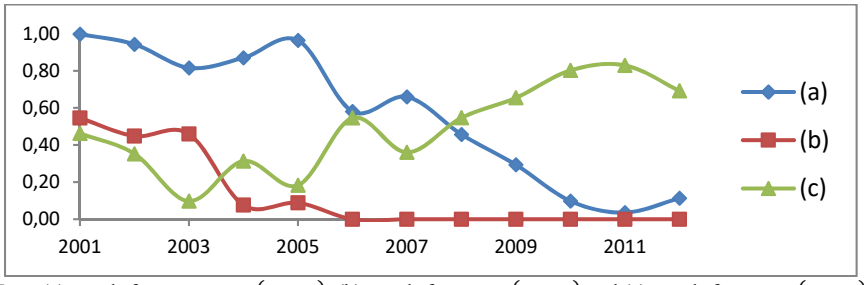
model allows introducing the fuzziness into the coefficients a_x and b_x and, as a consequence, point predictions and probabilistic interval predictions, as well, are fuzzified. So, the projection of $e_{[0,1),2012}$ is the TFN (78.19, 1.25, 1.30) years, whereas the lower and upper bounds of the 90% confidence fuzzy probabilistic interval are, respectively (76.56, 1.19, 1.28) and (79.75, 1.34, 1.33).

Table 5. TFN approximation of the estimates of $\tilde{E}^*(\tilde{e}^*_{[0,1),t})$ and $\widetilde{e^{*,10\%}_{[0,1),t}}$, with lower and upper bounds $\tilde{e}^{*,5\%}_{[0,1),t}$ and $\tilde{e}^{*,95\%}_{[0,1),t}$, in 2001-2012.

	$\tilde{E}^*(\tilde{e}^*_{[0,1),t})$			$\tilde{e}^{*\ 5\%}_{[0,1),t}$			$\tilde{e}^{*\ 95\%}_{[0,1),t}$		
Year	Centre	Left spread	Right spread	Centre	Left spread	Right spread	Centre	Left spread	Right spread
2001	76.31	1.18	1.28	75.73	1.17	1.28	76.96	1.20	1.28
2002	76.49	1.19	1.28	75.72	1.17	1.28	77.20	1.21	1.29
2003	76.66	1.19	1.28	75.75	1.17	1.28	77.54	1.22	1.29
2004	76.84	1.20	1.28	75.82	1.18	1.28	77.85	1.23	1.30
2005	77.01	1.20	1.29	75.89	1.18	1.28	78.07	1.24	1.30
2006	77.18	1.21	1.29	75.94	1.18	1.28	78.29	1.25	1.30
2007	77.35	1.21	1.29	75.94	1.18	1.28	78.60	1.27	1.31
2008	77.52	1.22	1.29	75.97	1.18	1.28	78.80	1.28	1.31
2009	77.69	1.23	1.29	76.10	1.18	1.28	79.05	1.29	1.32
2010	77.85	1.23	1.30	76.13	1.18	1.28	79.28	1.31	1.32
2011	78.02	1.24	1.30	76.26	1.18	1.28	79.50	1.32	1.33
2012	78.19	1.25	1.30	76.56	1.19	1.28	79.75	1.34	1.33

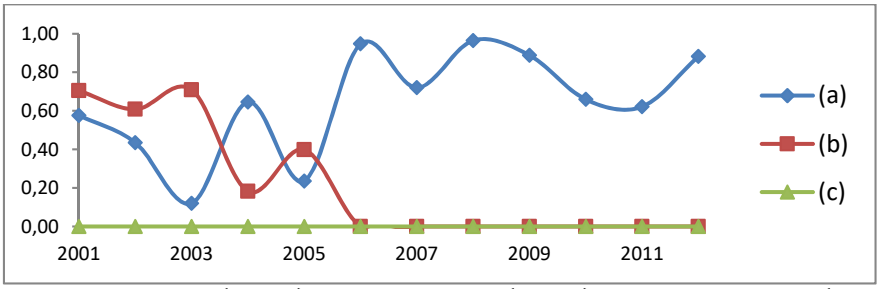
Table 6. TFN approximation of the estimates of $\tilde{E}^*(\tilde{e}^*_{[65,69),t})$ and $\widetilde{e^{*,10\%}_{[65,69),t}}$, with lower and upper bounds $\tilde{e}^{*,5\%}_{[65,69),t}$ and $\tilde{e}^{*,95\%}_{[65,69),t}$, in 2001-2012.

	$\tilde{E}^*(\tilde{e}^*_{[65,69),t})$			$\tilde{e}^{*\ 5\%}_{[65,69),t}$			$\tilde{e}^{*\ 95\%}_{[65,69),t}$		
Year	Centre	Left spread	Right spread	Centre	Left spread	Right spread	Centre	Left spread	Right spread
2001	17.06	0.48	0.64	16.67	0.46	0.63	17.51	0.51	0.65
2002	17.18	0.49	0.64	16.66	0.46	0.63	17.68	0.52	0.65
2003	17.30	0.49	0.64	16.68	0.46	0.63	17.92	0.53	0.66
2004	17.42	0.50	0.65	16.73	0.46	0.63	18.15	0.55	0.66
2005	17.54	0.51	0.65	16.78	0.46	0.63	18.31	0.56	0.67
2006	17.67	0.52	0.65	16.82	0.47	0.63	18.47	0.57	0.67
2007	17.79	0.52	0.66	16.82	0.47	0.63	18.70	0.58	0.67
2008	17.91	0.53	0.66	16.83	0.47	0.63	18.85	0.59	0.68
2009	18.03	0.54	0.66	16.92	0.47	0.63	19.04	0.61	0.68
2010	18.15	0.55	0.66	16.94	0.47	0.64	19.21	0.62	0.68
2011	18.27	0.55	0.67	17.03	0.48	0.64	19.38	0.63	0.69
2012	18.40	0.56	0.67	17.24	0.49	0.64	19.57	0.64	0.69



Note: (a) stands for $\mu_{\tilde{E}^*(\tilde{e}^*_{[0,1),t})}(e_{[0,1),t})$, (b) stands for $\mu_{\tilde{e}^{*,5\%}_{[0,1),t}}(e_{[0,1),t})$ and (c) stands for $\mu_{\tilde{e}^{*,95\%}_{[0,1),t}}(e_{[0,1),t})$.

Fig. 6. Membership levels $\mu_{\tilde{E}^*(\tilde{e}^*_{[0,1),t})}(e_{[0,1),t})$, $\mu_{\tilde{e}^{*,5\%}_{[0,1),t}}(e_{[0,1),t})$ and $\mu_{\tilde{e}^{*,95\%}_{[0,1),t}}(e_{[0,1),t})$ in 2001-2012.



Note: (a) stands for $\mu_{\tilde{E}^*\left(\tilde{e}^*_{[65,69),t}\right)}\left(e_{[65,69),t}\right)$, (b) stands for $\mu_{\tilde{e}^{*,5\%}_{[65,69),t}}\left(e_{[65,69),t}\right)$ and (c) stands for $\mu_{\tilde{e}^{*,95\%}_{[65,69),t}}\left(e_{[65,69),t}\right)$.

Fig. 7. Membership levels $\mu_{\tilde{E}^*\left(\tilde{e}^*_{[65,69),t}\right)}\left(e_{[65,69),t}\right)$, $\mu_{\tilde{e}^{*,5\%}_{[65,69),t}}\left(e_{[65,69),t}\right)$ and $\mu_{\tilde{e}^{*,95\%}_{[65,69),t}}\left(e_{[65,69),t}\right)$ in 2001-2012.

Table 7. Assessment of the capability prediction of $\tilde{E}^*\left(\tilde{e}^*_{x,t}\right)$ with a Wilcoxon rank test for the period 2001-2012.

Capability prediction of $\tilde{E}^*\left(\tilde{e}^*_{x,t}\right)$ per years							
Year	W	Median	Mean	Year	W	Median	Mean
2001	92*	0.602	0.718	2007	93	0.414	0.589
2002	141	0.488	0.656	2008	141	0.438	0.586
2003	77**	0.181	0.338	2009	59***	0.534	0.555
2004	140	0.454	0.653	2010	122	0.462	0.445
2005	117	0.165	0.440	2011	67***	0.521	0.443
2006	61***	0.528	0.699	2012	62***	0.202	0.409

Capability prediction of $\tilde{E}^*\left(\tilde{e}^*_{[0,1),t}\right)$ and $\tilde{E}^*\left(\tilde{e}^*_{[65,69),t}\right)$ on life expectancy at birth and retirement			
	Median	Mean	W
$\tilde{E}^*\left(\tilde{e}^*_{[0,1),t}\right)$	0.628	0.567	37
$\tilde{E}^*\left(\tilde{e}^*_{[65,69),t}\right)$	0.655	0.609	19

Notes: (1) Each year has 24 predictions on life expectations, one per age group. (2) Each age group has 12 predictions available, one for each assessed year. (3) W stands for the value of Wilcoxon rank test statistic. (4) "*", "**" and "***" stand for the rejection of the null hypothesis that median value of $\mu_{\tilde{E}^*\left(\tilde{e}^*_{x,t}\right)}\left(e_{x,t}\right)$ is 0.5 with a significance level of 10%, 5% and 1% respectively.

Figures 6 and 7 represent the membership level of true observed values for life expectancies, $e_{x,t}$, into their estimates $\tilde{E}^*\left(\tilde{e}^*_{x,t}\right)$, $\tilde{e}^{*,5\%}_{x,t}$ and $\tilde{e}^{*,95\%}_{x,t}$. Concretely, we take the estimates calculated in Tables 5 and 6. We can check that $e_{[0,1),t}$ is fitted quite accurately by $\tilde{E}^*\left(\tilde{e}^*_{[0,1),t}\right)$ from 2001 to 2005. On the other hand, from 2006 to 2012, $\mu_{\tilde{E}^*\left(\tilde{e}^*_{[0,1),t}\right)}\left(e_{[0,1),t}\right)$ decreases to values near 0. However, we can also remark that in those years the membership level of $e_{[0,1),t}$ into the TFN $\tilde{e}^{*,95\%}_{[0,1),t}$ stands clearly up to 0.5. Likewise, for [65, 69), Figure 7 shows that $\tilde{E}^*\left(\tilde{e}^*_{[65,69),t}\right)$ fits $e_{[65,69),t}$, from 2006 to 2012, with clearly satisfactory membership levels that are usually up 0.8. It is true that $\mu_{\tilde{E}^*\left(\tilde{e}^*_{[65,69),t}\right)}\left(e_{[65,69),t}\right)$ has low values in the years 2003 and 2005 but they are compensated by the greater membership levels of $\mu_{\tilde{e}^{*,5\%}_{[65,69),t}}\left(e_{[65,69),t}\right)$. In a similar way as in subsection 4.3., we statistically test the capability of fuzzy mean life expectancies to predict actual life expectations at t = 2001,2002,…, 2012. Again, for each year t =2001,2002,…,2012 we implement a Wilcoxon rank test with the null hypothesis that the median value of $\mu_{\tilde{E}^*\left(\tilde{e}^*_{x,t}\right)}\left(m_{x,t}\right) = 0.5$. Table 7 shows that, except for the year 2003, in the years where the median of $\mu_{\tilde{E}^*\left(\tilde{m}_{x,t}\right)}\left(m_{x,t}\right)$ is under 0.5, we cannot reject the null hypothesis. On the other hand, in the years 2001, 2007, 2010 and 2011, there are statistical evidences that the median is above 0.5.

Due to the interest in actuarial analyses in life expectancy both at birth and at retirement, we test the quality of the prediction by $\tilde{E}^*\left(\tilde{e}^*_{x,t}\right)$ in years [0,1) an [65,69). The results are also collected in Table 7. We can check that the median and mean membership levels of observed life expectancies in the period 2001-2012 are consistently above 0.5. However, the Wilcoxon rank test does not reject in both age groups that $\mu_{\tilde{E}^*\left(\tilde{e}^*_{x,t}\right)}\left(e_{x,t}\right) = 0.5$.

6. Empirical assessment of the Fuzzy Random Lee-Carter model in eight Western European countries[1]

6.1. *Methodological considerations*

In this section we make a comparative assessment on the prediction capability of our proposed fuzzy-random extension of the LC model (FRLC) with both the basic LC (BLC) in [26] and the pure fuzzy LC version in [22] (FKSLC). Let us remark that BLC only considers random uncertainty of coefficients k_t. On the other hand, FKSLC introduces fuzzy uncertainty in all the coefficients of the LC model by means of symmetrical TFNs. Likewise, FKSLC handles uncertainty with the weakest t-norm instead of the commonly used minimum operator.

To carry out the analysis, we use central mortality rates collected separately for men and women in eight Western Europe countries (i.e. we use 16 databases) from [38] (http://www.mortality.org). As we made in sections 4 and 5 for the case of Spanish male population, we fit the model parameters by using central mortality rates in the period 1970-2000 and we test models out-of-sample performance during 2001-2012. Ages are again grouped in 5 year intervals, except for ages lower than 1 year, for ages from 1 to 5 years and for ages greater or equal to 110 years.

We assess two aspects regarding the fitting quality of the models:

Item 1. We measure and compare models' performance to make point predictions on central mortality rates and life expectancies. This is made by using the conventional error measures: Root Mean Squared Error (RMSE), Normalised Mean Squared Error (NMSE) and Mean Absolut Error (MAE). We consider these point predictions: the expectation for BLC, the core of the fuzzy expectation for FRLC and, finally, the core of the fuzzy prediction for FKSLC. Notice that point predictions by BLC and FRLC are the same by definition. So, in fact, we are making a comparison of a couple of predictive methods: BLC/FRLC versus FKSLC. Following [32], this pairwise comparison between techniques is made with both a sign test (Wins/Losses) and a Wilcoxon rank test.

Item 2. We evaluate the capability of BLC, FRLC and FKSLC to predict future values of $m_{x,t,}$ and $e_{x,t}$ by means of confidence intervals. In this second item, we measure the accuracy of a method as the rate of right predictions on $m_{x,t}$ or $e_{x,t}$ through confidence intervals estimates provided by the methods.

In this regard, let us make the following remarks:

- BLC only considers random uncertainty of k_t. So, after establishing a significance level ε, that in our numerical assessment will be 10%, BLC predicts life variables as a $1-\varepsilon$ confidence interval like in (14).

- FRLC estimates the lower and upper bounds of the $1-\varepsilon$ confidence interval by means of two TFNs. To obtain standard confidence interval predictions, we transform these estimates into a conventional confidence interval that comes from the convex hull, $C(\cdot)$, of the expected intervals (4b) corresponding to $\frac{\varepsilon}{2}$ and $1-\frac{\varepsilon}{2}$ percentiles of fuzzy predictions (5% and 95% in our numerical application). So, for $m_{x,t}$, the $1-\varepsilon$ confidence interval prediction is $C\left(e_I\left(\tilde{m}_{x,t}^{*,\frac{\varepsilon}{2}}\right) \cup e_I\left(\tilde{m}_{x,t}^{*,1-\frac{\varepsilon}{2}}\right)\right)$. Analogously, the $1-\varepsilon$ confidence interval prediction of $e_{x,t}$ is $C\left(e_I\left(\tilde{e}_{x,t}^{*,\frac{\varepsilon}{2}}\right) \cup e_I\left(\tilde{e}_{x,t}^{*,1-\frac{\varepsilon}{2}}\right)\right)$. For example, the life expectancy at birth of a Spanish man born in 2004 for $\varepsilon = 10\%$ is built up from $\tilde{e}^*_{[0,1),2004}{}^{5\%} = (75.82, 1.18, 1.28)$ and $\tilde{e}^*_{[0,1),2004}{}^{95\%} = (77.85, 1.23, 1.30)$ (see Table 5). We easily find that $e_I\left(\tilde{e}^*_{[0,1),2004}{}^{5\%}\right) = [75.23, 76.46]$ and $e_I\left(\tilde{e}^*_{[0,1),2004}{}^{95\%}\right) = [77.23, 78{,}50]$. Then, the 90% confidence interval prediction for $e_{[0,1),2004}$ is (in years):

$$C([75.23, 76.46] \cup [77.23, 78{,}50]) = [75.23, 78.50]$$

- FKSLC directly predicts mortality variables as FNs. The expected interval of the FN obtained from this method is taken as its confidence interval.

The analysis of both questions is developed in two levels:

a) In each population, we independently assess the predictive capability of each method. For a given population we must predict 24 variables for each of the 12 years that testing period 2001-2012 comprises. In each year we find the mean value of the accuracy measures and so, for each population, we have 12 available mean values of accuracy (one per year). The results that we find in this case are exclusive to the population studied.

[1] This section is especially benefited by the helpful suggestions of one anonymous referee.

b) We will use the mean results of the accuracy predictions within the whole period 2001-20012 of all populations to make an inter-population assessment. It may lead to extract more general conclusions about the method performance. In this case, we will work with a sample of 16 different goodness of fit measures and we will extract more general conclusions.

Following [17] and [18], an adequate non-parametrical test to carry out this kind of analysis is the Friedman rank test (Friedman $\chi 2$ and Iman-Davenport F statistics) that may be completed by the pairwise comparisons that allow using Friedman ranks (Z-score). Likewise, given that FRLC and FKSLC are extensions of BLC, we will implement the multiple sign test described in [17] where the control technique is BLC.

Table 8a. Mean RMSE, NMSE and MAE (per years) of central mortality rates point predictions for Spanish men by the evaluated methods (Item 1).

	RMSE		NMSE		MAE	
Year	BLC/FRLC	FKSLC	BLC/FRLC	FKSLC	BLC/FRLC	FKSLC
2001	0.006	0.002	0.0039	0.0009	0.0031	0.0003
2002	0.007	0.011	0.0044	0.0051	0.0040	0.0097
2003	0.020	0.011	0.0108	0.0058	0.0280	0.0088
2004	0.012	0.023	0.0070	0.0091	0.0110	0.0392
2005	0.018	0.018	0.0101	0.0082	0.0240	0.0247
2006	0.011	0.020	0.0066	0.0111	0.0103	0.0297
2007	0.015	0.022	0.0084	0.0095	0.0168	0.0360
2008	0.016	0.022	0.0089	0.0115	0.0204	0.0361
2009	0.014	0.018	0.0078	0.0097	0.0151	0.0240
2010	0.013	0.016	0.0074	0.0089	0.0139	0.0202
2011	0.011	0.015	0.0066	0.0085	0.0110	0.0175
2012	0.022	0.026	0.0119	0.0129	0.0407	0.0498
	Wins/Losses	10/2**	Wins/Losses	9/3*	Wins/Losses	10/3**
	W	8**	W	11**	W	15*

Notes: (1) "Wins/Losses" stands for the number of cases in which BLC and FRLC point predictions are better/worse than FKSLC. (2) W stands for the value of the Wilcoxon rank test statistic. (3) "*", "**" and "***" stand for the rejection of the null hypothesis with a significance level of 10%, 5% and 1% respectively.

Table 8b. Mean RMSE, NMSE and MAE (per years) of life expectancy point predictions for Spanish men by the evaluated methods (Item 1).

	RMSE		NMSE		MAE	
Year	BLC/FRLC	FKSLC	BLC/FRLC	FKSLC	BLC/FRLC	FKSLC
2001	0.152	0.246	1.51E-04	2.45E-04	0.122	0.160
2002	0.194	0.232	1.91E-04	2.30E-04	0.149	0.155
2003	0.327	0.218	3.21E-04	2.16E-04	0.287	0.202
2004	0.190	0.400	1.83E-04	3.90E-04	0.173	0.287
2005	0.270	0.344	2.61E-04	3.35E-04	0.235	0.264
2006	0.348	0.647	3.30E-04	6.19E-04	0.235	0.484
2007	0.306	0.571	2.89E-04	5.44E-04	0.251	0.404
2008	0.443	0.754	4.14E-04	7.38E-04	0.308	0.561
2009	0.560	0.889	5.18E-04	8.30E-04	0.008	0.010
2010	0.712	1.060	6.51E-04	9.79E-04	0.007	0.009
2011	0.769	1.127	6.98E-04	1.03E-03	0.007	0.008
2012	0.703	1.056	6.36E-04	9.75E-04	0.012	0.013
	Wins/Losses	11/1*	Wins/Losses	11/1**	Wins/Losses	11/1**
	W	9**	W	9**	W	5**

Notes: (1) "Wins/Losses" stands for the number of cases in which BLC and FRLC point predictions are better/worse than FKSLC (2) W stands for the value of the Wilcoxon rank test statistic. (3) "*", "**" and "***" stand for the rejection of the null hypothesis with a significance level of 10%, 5% and 1% respectively.

Table 8c. Mean proportion of successful predictions on central mortality rates of BLC, FRLC and FKSLC with confidence intervals (Item 2).

	Proportion of successful predictions			Test Results
Year	BLC	FRLC	FKSLC	**Global comparison**
2001	0.625	0.875	0.708	Friedman $\chi2$ = 20.667***
2002	0.625	0.875	0.750	Iman-Davenport F Statistic = 31***
2003	0.542	0.667	0.458	**Pairwise comparisons**
2004	0.625	0.833	0.458	
2005	0.542	0.667	0.292	FRLC versus BLC
2006	0.625	0.792	0.375	Z score = 2.858
2007	0.583	0.792	0.500	p-values: (a) 4.27E-03; (b) 0.013; (c) 0.009
2008	0.583	0.708	0.333	
2009	0.583	0.750	0.250	FKSLC versus FRLC
2010	0.583	0.750	0.292	Z score = -4.491
2011	0.583	0.708	0.333	p-values: (a) 7.10E-06; (b) 2.13E-05; (c) 2.13E-05
2012	0.542	0.583	0.167	
Multiple sign test (the control method is BLC)				FKSLC versus BLC
Wins/Losses of FRLC against BLC: 12/0 +				Z score = -1.633
Wins/Losses of FKSLC against BLC: 2/10 -				p-values: (a) 0.102; (b) 0.307; (c) 0.102

Notes: (1) Friedman $\chi2$ follows a Squared-Chi with 2 grades of freedom and Iman-Davenport F follows a Snedecor F with 2(24) grades of freedom. (2) "*", "**" and "***" stand for the rejection of the null hypothesis with a significance level of 10%, 5% and 1% respectively. (3) (a) indicates standard p-value, and (b) and (c) Nemenyi and Holm p-value corrections for multiple pairwise comparisons. (4) "+" indicates that the evaluated method outperforms the control method with at least at 10% significance level whereas "-" indicates that the evaluated method underperforms the control method with at least at 10% significance level.

Table 8d. Mean proportion of successful predictions on life expectancies of BLC, FRLC and FKSLC with confidence intervals (Item 2).

	Proportion of successful predictions			Test Results
Year	BLC	FRLC	FKSLC	**Global comparison**
2001	0.792	0.958	0.958	Friedman $\chi2$ = 8.0417***
2002	0.833	0.958	0.958	Iman-Davenport F Statistic = 5.5431**
2003	0.667	0.792	0.667	**Pairwise comparisons**
2004	0.833	0.958	0.625	
2005	0.750	0.833	0.458	FRLC versus BLC
2006	0.875	0.958	0.375	Z score = 2.756
2007	0.833	0.917	0.583	p-values: (a) 0.006; (b) 0.018; (c) 0.012
2008	0.833	0.875	0.292	
2009	0.875	0.917	0.292	FKSLC versus FRLC
2010	0.875	0.958	0.250	Z score = -3.878
2011	0.833	0.958	0.250	p-values: (a) 6.88E-05; (b) 2.06E-04; (c) 1.38E-04
2012	0.833	0.875	0.167	
Multiple sign test (the control method is BLC)				FKSLC versus BLC
Wins/losses of FRLC against BLC: 12/0 +				Z score = -1.123
Wins/losses of FKSLC against BLC: 3/9				p-values: (a) 0.262; (b) 0.785; (c) 0.262

Notes: (1) Friedman $\chi2$ follows a Squared-Chi with 2 grades of freedom and Iman-Davenport F follows a Snedecor F with 2(24) grades of freedom (2) "*", "**" and "***" stand for the rejection of the null hypothesis with a significance level of 10%, 5% and 1% respectively. (3) (a) indicates standard p-value, and (b) and (c) Nemenyi and Holm p-value corrections for multiple pairwise comparisons. (4) "+" indicates that the evaluated method outperforms the control method with at least at 10% significance level whereas "-" indicates that the evaluated method underperforms the control method with at least at 10% significance level.

6.2. *Comparison of BLC, FRLC and FKSLC for each population*

We now show the adequacy of the three LC methods evaluated in 16 populations. We present in a more detailed way the results corresponding to Spanish male population (Table 8a-8d) and a summary table for all the analysed countries (Tables (9a-9e)).

Regarding item 1, we can check in Tables 8a and 8b that for Spanish male population, BLC/FRLC point predictions of $m_{x,t}$. and $e_{x,t}$ are, generally, more accurate than those by FKSLC and this best adjustment has a consistent statistical significance. Furthermore, Table 9a shows that in the studied populations, as in the case of Spanish men, point predictions of $m_{x,t}$ from BLC/FRLC are normally better than those from FKSLC and this fact has also statistical significance. We can appreciate three exceptions: Belgian male population, where FKSLC beats BLC/FRLC with a consistent statistical level and UK and Netherlands fe-

male populations where we do not appreciate any significant better method. Table 9c shows that in the prediction of $e_{x,t}$, it is less clear that BLC/FRLC predictions are better than those by FKSLC. BLC/FRLC beats FKSLC with a clear statistical significance in eight populations but in five populations FKSLC works clearly better. Likewise, in three populations the possible superior performance of a given method has no statistical significance.

In regards to item 2, in Spanish male population, we can check in Tables 8c and 8d that Friedman rank test rejects the homogeneity in the accuracy of the predictions over analysed life variables by the three assessed methods. Pairwise comparisons lead us to conclude that FRLC makes better interval predictions than BLC and FKSLC. However, despite the fact that we can detect that BLC beats FKSLC, this superior performance has no statistical significance. In this sense, multiple sign test shows that our method clearly beats the control method and, on the other hand, the control method seems to be superior to FKSLC but without statistical significance. Tables 9c-9d show that those facts are common to all studied populations. So, Friedman $\chi 2$ and Iman-Davenport statistics always reject the homogeneity of the prediction capability by the three methods. This fact applies for $m_{x,t,}$ and for $e_{x,t}$. Pairwise Friedman ranks tests show that the prediction on central mortality rates by FRLC beats significantly those obtained by BLC and FKSLC. Likewise, we can also check that BLC usually makes more accurate interval predictions than FKSLC but that better performance, except for the case of French women, has not statistical significance.

In the analysis of life expectancy predictions, pairwise Friedman ranks tests (see Table 9d) reveal that FRLC predicts confidence intervals consistently better than other methods in most populations. In any case, it is also true that in French and Italy female populations and Portugal male population (Netherlands male population) the greater accuracy of FRLC over BLC (FKSLC over FRLC) is not statistically significant. We can also check that in most cases BLC includes more percentage of observed values of $e_{x,t}$ than FKSLC but it is only statistically relevant in five populations. However, in the case of Netherlands male population, FKSLC model predicts life expectancies better than BLC with a clear significance level. Results of multiple sign tests in Table 9e show that our method improves significantly BLC (the control method) whereas this clearly does not follow with FKSLC method.

Table 9a. Results of sign and Wilcoxon tests on the difference between the accuracy of point estimates on central mortality rates by BLC/FRLC and FKSLC in the period 2001-2012 (Item 1).

	RMSE		NMSE		MAE	
	Wins/Losses	*W*	Wins/Losses	*W*	Wins/Losses	*W*
Austria (Men)	12/0**	0***	12/0**	0***	12/0**	0***
Austria (Women)	12/1**	0***	12/0**	0***	12/0**	0***
Belgium (Men)	3/9*	23	5/7	25	2/10**	18
Belgium (Women)	12/0**	0***	12/0**	0***	12/0**	0***
France (Men)	11/1**	12**	11/1**	11**	11/1**	11**
France (Women)	11/1**	11**	10/2**	22	11/1**	11**
Italy (Men)	12/0**	0***	12/0**	0***	12/0**	0***
Italy (Women)	11/1**	8**	8/4	34	11/1**	3***
Netherlands (Men)	9/3*	15*	10/2**	11**	9/3*	16*
Netherlands (Women)	8/4	31	8/4	28	8/4	30
Portugal (Men)	10/2**	22	9/3*	24	10/2**	16*
Portugal (Women)	4/8	23	3/9*	14**	4/8	22
Spain (Men)	10/2**	8**	9/3*	11**	10/2**	15*
Spain (Women)	12/0**	0***	12/0**	0***	12/0**	0***
UK(Men)	11/1**	11**	12/0**	0***	11/1**	12**
UK (Women)	5/7	27	6/6	38	5/7	31

Notes: (1) "Wins/Losses" are accounted from the perspective of BLC/FRLC. (2) "*", "**" and "***" stand for the rejection of the null hypothesis with a significance level of 10%, 5% and 1% respectively.

Table 9b. Results of sign and Wilcoxon tests on the difference between the accuracy of point estimates on life expectancies by BLC/FRLC and FKSLC in the period 2001-2012 (Item 1).

	RMSE		NMSE		MAE	
	Wins/Losses	W	Wins/Losses	W	Wins/Losses	W
Austria (Men)	12/0**	0***	12/0**	0***	12/0**	0***
Austria (Women)	0/12**	0***	0/12**	0***	0/12**	0***
Belgium (Men)	0/12**	0***	0/12**	0***	0/12**	0***
Belgium (Women)	12/0**	0***	12/0**	0***	12/0**	0***
France (Men)	11/1**	12**	11/1**	12**	11/1**	12**
France (Women)	9/3*	11**	9/3*	29	9/3*	28
Italy (Men)	12/0**	0***	12/0**	0***	12/0**	0***
Italy (Women)	4/8	30	4/8	30	3/9*	26
Netherlands (Men)	0/12**	0***	0/12**	0***	0/12**	0***
Netherlands (Women)	4/8	18	4/8	18	4/8	18
Portugal (Men)	12/0**	0***	12/0**	0***	12/0**	0***
Portugal (Women)	12/0**	0***	8/4	37	8/4	37
Spain (Men)	11/1**	9**	11/1**	9**	11/1**	9**
Spain (Women)	1/11**	12**	1/11**	12**	1/11**	12**
UK(Men)	12/0**	11**	12/0**	0***	12/0**	0***
UK (Women)	0/12**	0***	0/12**	0***	0/12**	0***

Notes: (1) "Wins/Losses" are accounted from the perspective of BLC/FRLC. (2) "*", "**" and "***" stand for the rejection of the null hypothesis with a significance level of 10%, 5% and 1% respectively.

Table 9c. Results of Friedman rank tests and pairwise Friedman rank tests for the accuracy of the confidence interval predictions on central mortality rates by BLC, FRLC and FKSLC in sample populations in the period 2001-2012 (Item 2).

	Pairwise Z scores from Friedman ranks			Friedman test	
	FRLC vs BLC	FKSLC vs FRLC	FKSLC vs BLC	Friedman $\chi 2$	Iman-Davenport F
Austria (Men)	3.164***	-3.776***	-0.612	8.542**	6.079***
Austria (Women)	3.062***	-4.082***	-1.021	14.083***	15.621***
Belgium (Men)	2.654**	-4.491***	-1.837	18.375***	35.933***
Belgium (Women)	2.654**	-4.695***	-2.041	22.167***	133.026***
France (Men)	2.858**	-4.491***	-1.633	20.660***	68.042***
France (Women)	2.449**	-4.695***	-2.245*	20.000***	55.000***
Italy (Men)	2.654**	-4.491***	-1.837	16.420***	23.828***
Italy (Women)	2.654**	-4.695***	-2.041	22.167***	133.026***
Netherlands (Men)	2.858**	-4.491***	-1.633	20.667***	68.208***
Netherlands (Women)	2.654**	-4.287***	-1.633	10.830***	9.046***
Portugal (Men)	2.654**	-4.695***	-2.041	22.167***	133.026***
Portugal (Women)	3.062***	-3.674***	-0.612	15.500***	20.059***
Spain (Men)	2.858**	-4.491***	-1.633	20.667***	68.208***
Spain (Women)	3.062***	-4.082***	-1.021	14.083***	15.621***
UK(Males)	2.654**	-4.082***	-1.429	5.417*	3.207*
UK (Women)	2.654**	-4.491***	-1.837	16.417***	23.815***

Notes: (1) "*", "**" and "***" stand for the rejection of the null hypothesis with a significance level of 10%, 5% and 1% respectively. (2) Friedman $\chi 2$ follows a Squared-Chi with 2 grades of freedom and Iman-Davenport F follows a Snedecor F with 2(24) grades of freedom.

Table 9d. Results of Friedman rank tests and pairwise Friedman rank tests for the accuracy of the confidence interval predictions on life expectancies by BLC, FRLC and FKSLC in sample populations in the period 2001-2012 (Item 2).

	Pairwise Z scores from Friedman ranks			Friedman test	
	FRLC vs BLC	FKSLC vs FRLC	FKSLC vs BLC	Friedman $\chi 2$	Iman-Davenport F
Austria (Men)	3.164***	-4.185***	-1.021	19.042***	42.247***
Austria (Women)	2.449**	-3.266***	-0.816	7.750**	5.246**
Belgium (Men)	2.143*	-4.287***	-2.143*	18.370***	35.892***
Belgium (Women)	2.245*	-4.491***	-2.245*	20.100***	56.692***
France (Men)	2.347*	-3.470***	-1.123	12.540***	12.037***
France (Women)	1.123	-3.572***	-2.449**	9.375***	7.051***
Italy (Men)	2.449**	-4.899***	-2.449**	24.000***	∞***
Italy (Women)	2.041	-2.449**	-0.408	6.500**	4.086**
Netherlands (Men)	4.695***	-2.041	2.654**	22.167***	133.026***

Netherlands (Women)	2.347*	-3.470***	-1.123	12.542***	12.041***
Portugal (Men)	2.245*	-3.572***	-1.327	13.040***	13.088***
Portugal (Women)	1.429	-3.878***	-2.449**	11.420***	9.986***
Spain (Men)	2.756**	-3.980***	-1.225	16.250***	23.065***
Spain (Women)	2.960***	-2.449**	0.510	14.083***	15.622***
UK(Males)	2.552**	-4.491***	-1.939	20.000***	55.000***
UK (Women)	2.654**	-4.695***	-2.041	22.167***	133.026***

Notes: (1) "*", "**" and "***" stand for the rejection of the null hypothesis with a significance level of 10%, 5% and 1% respectively. (2) Friedman $\chi 2$ follows a Squared-Chi with 2 grades of freedom and Iman-Davenport F follows a Snedecor F with 2(24) grades of freedom.

Table 9e. Results of multiple sign test for the confidence interval predictions of FRLC and FKSLC with a control method (BLC) (Item 2).

	Central mortality rates predictions					
	FRLC vs BLC			FKSLC vs BLC		
	Wins	Losses	r_j	Wins	Losses	r_j
Austria (Men)	12	0	0 **	3	8	3
Austria (Women)	10	0	0 **	3	8	3
Belgium (Men)	10	0	0 **	2	10	2 **
Belgium (Women)	10	0	0 **	1	11	1 **
France (Men)	11	0	0 **	2	9	2 **
France (Women)	6	0	0 **	0	9	0 **
Italy (Men)	12	0	0 **	0	12	0 **
Italy (Women)	8	0	0 **	0	8	0 **
Netherlands (Men)	12	0	0 **	12	0	0 **
Netherlands (Women)	10	0	0 **	3	9	3
Portugal (Men)	11	0	0 **	2	10	2 **
Portugal (Women)	6	0	0 **	1	10	1 **
Spain (Men)	12	0	0 **	2	9	2 **
Spain (Women)	11	0	0 **	6	5	5
UK(Men)	11	0	0 **	1	10	1 **
UK (Women)	12	0	0 **	1	11	1 **
	Life expectancy predictions					
	FRLC vs BLC			FKSLC vs BLC		
	Wins	Losses	r_j	Wins	Losses	r_j
Austria (Men)	12	0	0 **	3	8	3
Austria (Women)	10	0	0 **	3	8	3
Belgium (Men)	10	0	0 **	0	10	0 **
Belgium (Women)	10	0	0 **	1	11	1 **
France (Men)	11	0	0 **	2	9	2 **
France (Women)	6	1	1 **	0	9	0 **
Italy (Men)	12	0	0 **	0	12	0 **
Italy (Women)	8	0	0 **	2	6	2 **
Netherlands (Men)	12	0	0 **	12	0	0 **
Netherlands (Women)	10	0	0 **	3	9	3
Portugal (Men)	11	0	0 **	2	10	2 **
Portugal (Women)	6	0	0 **	1	10	1 **
Spain (Men)	12	0	0 **	2	9	2 **
Spain (Women)	11	0	0 **	6	5	5
UK(Men)	11	0	0 **	1	10	1 **
UK (Women)	12	0	0 **	1	11	1 **

Notes: (1) "**" stands for the rejection of the null hypothesis with a significance level of at least 5%. (2) "Wins/Losses" stands for the number of wins/losses of the evaluated method over the control method. (3) r_j stands for the minimum between number of wins and losses of the evaluated method.

6.3. *A global comparison of BLC, FRLC and FKSLC*

In this section we show the results of testing BLC, FRLC and FKSLC from a sample composed by the mean values of accuracy prediction measures within 2001-2012 of the 16 populations considered in this paper. They are summarized in Tables 10a-10d. Regarding item 1, when evaluating predictions about central mortality rates, Table 10a shows that BLC and FRLC have greater accuracy than FKSLC method and it is significant. From Table 10b we can also indicate that

point predictions on life expectancy by BLC/FRLC are more accurate than those by FKSLC but this better performance has not enough statistical significance.

Tables 10c and 10d reveal that Friedman rank test undoubtedly rejects that the three evaluated methods provide interval confidence predictions with homogenous accuracy. Likewise, we can observe in these tables that from the interval confidence prediction perspective, our method improves BLC and FKSLC. Also, that BLC provides better predictions than FKSLC. In this sense, multiple sign tests reveal that whereas FRLC improves BLC significantly, FKSLC performs poorer than the control method.

Table 10a. Mean value of RMSE, NMSE and MAE of the predictions on central mortality rates in 2001-2012 in sample populations (Item 1).

	RMSE		NMSE		MAE	
	BLC/FRLC	FKSLC	BLC/FRLC	FKSLC	BLC/FRLC	FKSLC
Austria (Men)	0.0462	0.0538	0.0711	0.0698	0.0249	0.0295
Austria (Women)	0.0037	0.0024	0.0003	0.0002	0.0017	0.0029
Belgium (Men)	0.0207	0.0151	0.0295	0.0150	0.0096	0.0074
Belgium (Women)	0.0150	0.0425	0.0208	0.1658	0.0076	0.0213
France (Men)	0.0059	0.0178	0.0029	0.0245	0.0033	0.0096
France (Women)	0.0056	0.0114	0.0043	0.0135	0.0030	0.0053
Italy (Men)	0.0125	0.0204	0.0147	0.0327	0.0062	0.0105
Italy (Women)	0.0083	0.0108	0.0086	0.0118	0.0042	0.0050
Netherlands (Men)	0.0141	0.0150	0.0129	0.0146	0.0073	0.0086
Netherlands (Women)	0.0080	0.0124	0.0067	0.0219	0.0043	0.0060
Portugal (Men)	0.0117	0.0176	0.0130	0.0210	0.0063	0.0089
Portugal (Women)	0.0124	0.0115	0.0075	0.0061	0.0126	0.0086
Spain (Men)	0.0138	0.0170	0.0165	0.0247	0.0078	0.0084
Spain (Women)	0.0055	0.0080	0.0032	0.0061	0.0029	0.0037
UK(Men)	0.0101	0.0143	0.0080	0.0149	0.0049	0.0082
UK (Women)	0.0104	0.0093	0.0114	0.0086	0.0052	0.0047
	Wins/Losses	12/4**	Wins/Losses	11/5	Wins/Losses	13/3**
	W	50	W	62	W	30**

Notes: (1) "Wins/Losses" stands for the number of cases in which BLC and FRLC point predictions are better/worse than FKSLC. (2) W stands for the value of the Wilcoxon rank test statistic. (3) "*", "**" and "***" stand for the rejection of the null hypothesis with a significance level of 10%, 5% and 1% respectively.

Table 10b. Mean value of RMSE, NMSE and MAE of the predictions on life expectancies in 2001-2012 in sample populations (Item 1).

	RMSE		NMSE		MAE	
	BLC/FRLC	FKSLC	BLC/FRLC	FKSLC	BLC/FRLC	FKSLC
Austria (Men)	1.284	1.429	1.26E-03	1.42E-03	0.999	1.122
Austria (Women)	0.3710	0.2432	2.97E-04	1.95E-04	0.3293	0.2158
Belgium (Men)	0.6309	0.5018	6.31E-04	5.02E-04	0.5594	0.4443
Belgium (Women)	0.1325	0.3647	1.06E-04	2.91E-04	0.1169	0.3298
France (Men)	0.3806	0.5661	3.61E-04	5.44E-04	0.3155	0.4776
France (Women)	0.1896	0.2127	1.41E-04	1.58E-04	0.1627	0.1818
Italy (Men)	0.5940	0.9189	5.50E-04	8.65E-04	0.4933	0.7639
Italy (Women)	0.1770	0.1739	1.34E-04	1.32E-04	0.1595	0.1558
Netherlands (Men)	1.3555	1.2628	1.34E-03	1.25E-03	1.1806	1.1047
Netherlands (Women)	0.3707	0.2834	3.00E-04	2.31E-04	0.3242	0.2526
Portugal (Men)	0.6177	0.7797	6.27E-04	8.03E-04	0.4653	0.5970
Portugal (Women)	1.2565	1.2897	2.34E-03	2.51E-03	0.5543	0.5793
Spain (Men)	0.4146	0.6287	3.87E-04	5.95E-04	0.1495	0.2130
Spain (Women)	0.2247	0.1671	1.68E-04	1.26E-04	0.1944	0.1433
UK(Men)	0.9079	0.9679	8.82E-04	9.48E-04	0.7989	0.8514
UK (Women)	0.5779	0.5256	4.74E-04	4.32E-04	0.4970	0.4516
	Wins/Losses	9/7	Wins/Losses	9/7	Wins/Losses	9/7
	W	61	W	55	W	65

Notes: (1) "Wins/Losses" stands for the number of cases in which BLC and FRLC point predictions are better/worse than FKSLC. (2) W stands for the value of the Wilcoxon rank test statistic. (3) "*", "**" and "***" stand for the rejection of the null hypothesis with a significance level of 10%, 5% and 1% respectively.

Table 10c. Mean proportion of successful predictions on central mortality rates in 2001-2012 by BLC, FRLC and FKSLC in sample populations (Item 2).

	Proportion of successful predictions			**Test Results**
	BLC	FRLC	FKSLC	**Global comparison**
Austria (Men)	0.326	0.729	0.295	Friedman $\chi2 = 32$***
Austria (Women)	0.524	0.847	0.431	Iman-Davenport F Statistic = ∞***
Belgium (Men)	0.434	0.576	0.236	**Pairwise comparisons**
Belgium (Women)	0.576	0.760	0.354	
France (Men)	0.618	0.733	0.340	FRLC versus BLC
France (Women)	0.587	0.767	0.396	Z score = 2.858
Italy (Men)	0.514	0.656	0.306	p-values: (a) 0.005; (b) 0.014; (c) 0.005
Italy (Women)	0.681	0.802	0.476	
Netherlands (Men)	0.385	0.590	0.316	FKSLC versus FRLC
Netherlands (Women)	0.618	0.806	0.483	Z score = -5.567
Portugal (Men)	0.566	0.705	0.257	p-values:(a) 0.000; (b) 0.000; (c) 0.000
Portugal (Women)	0.583	0.778	0.451	
Spain (Men)	0.587	0.750	0.410	FKSLC versus BLC
Spain (Women)	0.556	0.792	0.486	Z score = -2.828
UK(Men)	0.458	0.688	0.358	p-values: (a) 0.005; (b) 0.014; (c) 0.009
UK (Women)	0.698	0.858	0.451	
Multiple sign test (the control method is BLC)				
Wins/Losses of FRLC against BLC: 16/0 + Wins/Losses of FKSLC against BLC: 0/16 -				

Notes: (1) Friedman $\chi2$ follows a Squared-Chi with 2 grades of freedom and Iman-Davenport F follows a Snedecor F with 2(32) grades of freedom. (2) "*", "**" and "***" stand for the rejection of the null hypothesis with a significance level of 10%, 5% and 1% respectively. (3) (a) indicates standard p-value, and (b) and (c) Nemenyi and Holm p-value corrections for multiple pairwise comparisons. (4) "+" indicates that the evaluated method outperforms the control method with at least at 10% significance level whereas "-" indicates that the evaluated method underperforms the control method with at least at 10% significance level.

Table 10d. Mean proportion of successful predictions on life expectancies in 2001-2012 by BLC, FRLC and FKSLC in sample populations (Item 2).

	Proportion of successful predictions			**Test Results**
	BLC	**FRLC**	**FKSLC**	**Global comparison**
Austria (Men)	0.326	0.729	0.295	Friedman $\chi2 = 30.125$***
Austria (Women)	0.524	0.847	0.431	Iman-Davenport F Statistic = 241***
Belgium (Men)	0.434	0.576	0.236	**Pairwise comparisons**
Belgium (Women)	0.576	0.760	0.354	
France (Men)	0.618	0.733	0.340	FRLC versus BLC
France (Women)	0.587	0.767	0.396	Z score = 3.005
Italy (Men)	0.514	0.656	0.306	p-values: (a) 0.003; (b) 0.008; (c) 0.005
Italy (Women)	0.681	0.802	0.476	
Netherlands (Men)	0.385	0.590	0.316	FKSLC versus FRLC
Netherlands (Women)	0.618	0.806	0.483	Z score = -5.480
Portugal (Men)	0.566	0.705	0.257	p-values: (a) 0.000; (b) 0.000; (c) 0.000
Portugal (Women)	0.583	0.778	0.451	
Spain (Men)	0.587	0.750	0.410	FKSLC versus BLC
Spain (Women)	0.556	0.792	0.486	Z score = -2.475
UK(Men)	0.458	0.688	0.358	p-values: (a) 0.013; (b) 0.004; (c) 0.013
UK (Women)	0.698	0.858	0.451	
Multiple sign test (the control method is BLC)				
Wins/Losses of FRLC against BLC: 16/0 + Wins/Losses of FKSLC against BLC: 1/15 -				

Notes: (1) Friedman $\chi2$ follows a Squared-Chi with 2 grades of freedom and Iman-Davenport F follows a Snedecor F with 2(32) grades of freedom. (2) "*", "**" and "***" stand for the rejection of the null hypothesis with a significance level of 10%, 5% and 1% respectively. (3) (a) indicates standard p-value, and (b) and (c) Nemenyi and Holm p-value corrections for multiple pairwise comparisons. (4) "+" indicates that the evaluated method outperforms the control method with at least at 10% significance level whereas "-" indicates that the evaluated method underperforms the control method with at least at 10% significance level.

7. Conclusions and further extensions

This paper proposes a fuzzy-random approach of the Lee-Carter (LC) model. A fuzzy version of the LC model was firstly proposed by [22], who considered two different formulations. In the first one, which was refined in the works [14,24], the authors introduced fuzziness in all the parameters of the model by using TFNs. Nevertheless, the model developed in this paper assumes, as it was done in the seminal paper [26] and its subsequent extensions, that the trend of mortality across time is captured with an ARIMA model.

This fuzzy-random approach of the LC model can also be used to derivate variables linked to central mortality rates as probabilities of death or survival and life expectancies. From these variables, it is possible to price life annuities or insurance contracts. It can be done by using directly fitted fuzzy probabilities, as in the framework exposed in [34] or, alternatively, by reducing these fuzzy probabilities to a crisp value with the use of a defuzzifiying method.

When applying this new model to Spanish male population within the period 1970-2012, it is found that the model is satisfactory when it comes to its capability of fitting outcomes in the estimation sample (1970-2000) and forecasting central mortality rates over a time horizon of more than 10 years (2001-2012).

Moreover, we have made a comparative assessment of our fuzzy random methodology with seminal LC method [26] and fuzzy version of LC [22] and we have checked that, from interval confidence prediction perspective, our proposed methodology improves the models of these papers.

Acknowledgments

Authors thank helpful comments of two anonymous referees.

References

[1] Andrés-Sánchez, J. de (2012). Claim reserving with fuzzy regression and the two ways of ANOVA. *Applied Soft Computing* 12(8), 2435-2441.

[2] Andrés-Sánchez, J. de (2016) Fuzzy Regression Analysis: An Actuarial Perspective. In: Kahraman, C. and Cengiz, K. (eds.) *Fuzzy statistical making: Theory and Applications*, 175-201. Springer International Publishing.

[3] Andrés-Sánchez, J. de (2017). An empirical assessment of fuzzy Black and Scholes pricing option model in Spanish stock option market. *Journal of Intelligent & Fuzzy Systems* 33(4), 2509-2521.

[4] Andrés-Sánchez, J. de (2018) Pricing European options with triangular fuzzy parameters: assessing alternative triangular approximations in the Spanish stock option market. *International Journal of Fuzzy Systems* 20 (5), 1624-1643.

[5] Andrés-Sánchez, J. de and González-Vila Puchades, L. (2012) Using fuzzy random variables in life annuities pricing. *Fuzzy Sets and Systems* 188, 27-44.

[6] Andrés-Sánchez, J. de and González-Vila, L. (2017) The valuation of life contingencies: A symmetrical triangular fuzzy approximation". *Insurance: Mathematics and Economics* 72, 83-94.

[7] Booth, H. y Tickle, L. (2003). The future aged: new projections of Australia's elderly population. *Population Studies*, 22(4), 38-44.

[8] Brouhns, N., Denuit, M. and Vermunt, J. (2002) A Poisson log-bilinear regression approach to the construction of projected lifetables. *Insurance: Mathematics and Economics*, 31(3), 373-393.

[9] Buckley, J.J. and Qu, Y. (1990) On using α-cuts to evaluate fuzzy equations. *Fuzzy Sets and Systems* 38, 309-312.

[10] Carter, L. and Prkawetz, A. (2001) Examining structural shifts in mortality using the Lee-Carter method. Mpidr wp 2001-007, Center for Demography and Ecology Information, University of Wisconsin-Madison.

[11] Chen, F., Chen, Y., Zhou, J. and Liu, Y. (2016) Optimizing h value for fuzzy linear regression with asymmetric triangular fuzzy coefficients. *Engineering Applications of Artificial Intelligence* 47, 16-24.

[12] Currie, I., Kirkby, J., Durban, M. and Eilers, P. (2004) Smooth Lee-Carter models and beyond. In *Workshop on Lee-Carter Methods*. Edinburgh.

[13] Debón, A., Montes, F. and Puig, F. (2008) Modelling and forecasting mortality in Spain. *European Journal of Operational Research* 189 (3), 624-637

[14] Demirel, D. F. and Basak, M. (2015) A modified fuzzy Lee-Carter method for modeling human mortality. In: Rosa, A. Merelo, J.J., Dourado, A., Cadenas, J.M., Madani, K., Ruano, A. and Filipe, J. (eds.) *Proceedings of the 7th International Joint Conference on Computational Intelligence 2015*, (1), 17-24.

[15] Dickey, D.A. and Fuller, W.A. (1979) Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association* 74 (366a), 427-431.

[16] Dubois, D. and Prade, H. (1993) Fuzzy numbers: an overview. In: Dubois, D., Prade, H. and Yager, R.R. (eds.) *Fuzzy Sets for intelligent systems*, 113-148. Morgan Kaufmann Publishers.

[17] García, S., Fernández, A., Luengo, J. and Herrera, F. (2010) Advanced nonparametric tests for multiple comparisons in the design of experiments in computational intelligence and data mining: Experimental Analysis of Power. *Information Sciences* 180, 2044–2064.

[18] García, S. and Herrera, F. (2008) An Extension on "Statistical Comparisons of Classifiers over Multiple Data Sets" for all Pairwise Comparisons. *Journal of Machine Learning Research* 9, 2677-2694.

[19] Huang, T., Zhao, R. and Tang, W. (2009) Risk model with fuzzy random individual claim amount. *European Journal of Operational Research* 192, 879–890.

[20] Ishibuchi, H. and Nii, M. (2001) Fuzzy regression using asymmetric fuzzy coefficients and fuzzified neural networks. *Fuzzy Sets and Systems* 119, 273–290.

[21] Jong, P. de, Tickle, L. and Xu, J. (2016) Coherent modeling of male and female mortality using Lee-Carter in a complex number framework. *Insurance: Mathematics and Economics* 71, 130–137.

[22] Koissi, M.C and Shapiro, A. (2006) Fuzzy formulation of the Lee-Carter model for mortality forecasting. *Insurance: Mathematics and Economics* 39, 287-309.

[23] Koissi, M.C., Shapiro, A. and Högnäs, G. (2006) Evaluating and extending the Lee–Carter model for mortality forecasting: Bootstrap confidence interval. *Insurance: Mathematics and Economics* 38, 1–20.

[24] Kou, Y. (2016) Fuzzy Formulation of the Lee-Carter Model for the Mortality Forecasting with Age-Specific Enhancement. In: Cao, B.-Y., Wang, P.-Z., Liu, Z.-L. and Zhong, Y.-B. (eds.) *International Conference on Oriental Thinking and Fuzzy Logic*, 177-196. Springer International Publishing.

[25] Lee, R. and Nault, F. (1993) Modeling and forecasting provincial mortality in Canada. Montreal: *World Congress of the International Union for Scientific Study of Population*.

[26] Lee, R.D. and Carter, L.R. (1992) Modeling and forecasting US mortality. *Journal of the American Statistical Association* 419 (87), 659–675.

[27] Moskowitz, H. and Kim, K. (1993) On assessing the H value in fuzzy linear regression. *Fuzzy Sets and Systems* 58 (3), 303-327.

[28] Muzzioli, S. and De Baets, B. (2013). A comparative assessment of different fuzzy regression methods for volatility forecasting. *Fuzzy Optimization and Decision Making* 12(4), 433-450.

[29] Pascual, L., Romo, J. and Ruiz, E. (2004) Bootstrap predictive inference for ARIMA processes. *Journal of Time Series Analysis* 25 (4), 449-465.

[30] Renshaw, A. and Haberman, S. (2003) Lee-Carter mortality forecasting: a parallel generalized linear modelling approach for England and Wales mortality projections. *Journal of the Royal Statistical Society* 52, 119–137.

[31] Renshaw, A. and Haberman, S. (2006) A cohort-based extension to the Lee-Carter model for mortality reduction factors. *Insurance: Mathematics and Economics* 38(3), 556–570.

[32] SCI2S (Soft Computing and Intelligent Information Systems research group, University of Granada) *Statistical Inference in Computational Intelligence and Data Mining*. https://sci2s.ugr.es/sicidm.

[33] Shapiro, A. (2009) Fuzzy random variables. *Insurance: Mathematics and Economics* 44, 307-314.

[34] Shapiro, A. (2013) Modeling future lifetime as a fuzzy random variable. *Insurance: Mathematics and Economics* 53, 864-870.

[35] Tanaka, H. (1987) Fuzzy data analysis by possibility linear models. *Fuzzy Sets and Systems* 24, 363–375.

[36] Wilmoth, J.R. (1993). Computational methods for fitting and extrapolating the Lee-Carter model of mortality change. Technical report, Departament of Demography, University of California.

[37] Wilmoth, J.R. (1996) Health and Mortality among Elderly Populations. In: *Mortality projections for Japan: A comparison of four methods*, 266–287. Oxford University Press.

[38] Wilmoth, J.R., Andreev, K., Jdanov, D., et al. (2017). Methods Protocol for the Human Mortality Database. University of California, Berkeley, and Max Planck Institute for Demographic Research, Rostock, URL: http://mortality.org [version November 27, 2017].

[39] Yager, R.R. (1982) Fuzzy prediction based on regression models. *Information Sciences* 26, 45-63.