

Feasibility of Singular Spectrum Analysis in the Field of Forecasting Mortality Rate

Rahim Mahmoudvand^{1*}, Fatemeh Alehosseini² and Mohammad Zokaei²

¹*Statistical Research and Training Center (SRTC)*

and ²*Shahid Beheshti University*

Abstract:

The Lee-Carter model and its extensions are the most popular methods in the field of forecasting mortality rate. But, in spite of introducing several different methods in forecasting mortality rate so far, there is no general method applicable to all situations.

Singular Spectrum Analysis (SSA) is a relatively new, powerful and non-parametric time series analysis that its capability in forecasting different time series has been proven in the various sciences. In this paper, we investigate the feasibility of using the SSA to construct mortality forecasts. We use the Hyndman-Ullah model, which is a new extension of Lee-Carter model, as a benchmark to evaluate the performance of the SSA for mortality forecasts in France data sets.

Key words: Hyndman-Ullah model, Lee-Carter model, Mortality, Singular Spectrum Analysis.

1. Introduction

Mortality rate is of the most prominent demographic indicators that national statistical offices and United Nations routinely consider in their reports. Mortality rates have many applications in governmental agencies and insurance companies. For example, insurance companies and social security systems use the projections of future mortality rates in order to construct life tables and determine annuity prices.

Of course, we notice that the human mortality rate has declined substantially over the last century by progress in the health systems which cause to additional stress in support-systems for the elderly, such as pension provision. This change also affects on the population size and structure, as well as social security systems.

*Corresponding author.

Therefore, modelling and projecting of disease incidence and mortality rates is a problem of fundamental importance in the insurance and pensions industry and has been attracting the interest of researchers in the last decade.

Over the past century, researchers have introduced various methods to model and forecast mortality rate. The methods being used by demographers and actuaries up to the early 1990s were found to be deficient in that they systematically underestimated this downward trend (Haberman, 2010). The reason for this underestimation is that the estimations are computed for a specific time period and do not account for the decreasing trend in mortality. The systematic underestimation of mortality rates for pricing and reserving, particularly with respect to guaranteed annuity and pension benefits, may lead to serious financial consequences for insurance companies and pension plans in the long term.

In 1992, Lee and Carter proposed a new method which used the technique of singular value decomposition to represent the log of mortality rates in terms of two age-dependent factors and a single time-dependent one. The time-dependent factor could then be extracted and modelled using conventional time series methods so that forecasts could be made. This model is typically intended for making long-term predictions of aggregate mortality indicators like future life expectancies or annuity costs. The popular method of Lee and Carter (1992) to model and forecast mortality rate has undergone various extensions and modifications. For a review and recent developments, see Hyndman and Ullah (2007), Hyndman *et al.* (2011) and references therein. These methods have been improved the performance of mortality rate forecasts, however, there is no general method applicable to all situations.

Recently, Hyndman and Ullah have introduced a modified version of Lee-Carter model and have shown that the performance of their method is more accurate than other extensions of Lee-Carter model.

Singular Spectrum Analysis (SSA) is a relatively new non-parametric approach for analysing time series data which incorporates elements of classical time series analysis, multivariate statistics, multivariate geometry, dynamical systems and signal processing (Golyandina *et al.*, 2001). The aim of SSA is to decompose the original series into the sum of a small number of independent and interpretable components such as a slowly varying trend, oscillatory components and a structureless noise. The literature review on SSA shows that there are more than hundred papers on the application of SSA in the different areas and in the majority of them, superiority of SSA compared to other time series analysis techniques has been demonstrated (see, for example, Hassani *et al.*, 2009; Hassani and Thomakos, 2010 and references therein). Most recent developments in the theory and methodology of SSA can be found in Zhigljavsky, 2010. In this paper, we investigate the feasibility of using SSA to construct mortality forecasts.

As we mentioned, the method of Hyndman and Ullah (2007) achieves better forecasting results to mortality forecasting than other approaches. Therefore, we just compare our results with the method of Hyndman and Ullah (2007).

The rest of the paper is structured as follows: in Section 2, we give a brief description of Hyndman and Ullah model and go over the generic SSA methodology in Section 3. Empirical results are presented in Section 4. Finally, in Section 5, we offer some concluding remarks and directions for future research.

2. Hyndman-Ullah Approach

The Hyndman-Ullah approach can be expressed using the following equation:

$$\log m_t(x) = a(x) + \sum_{j=1}^K k_{t,j} b_j(x) + e_t(x) + \sigma_t(x) \epsilon_t(x), \quad (1)$$

where $m_t(x)$ denotes the mortality rate for age x at time t , $a(x)$ is the average pattern of mortality by age across years, $b_j(x)$ is a basis function and $k_{t,j}$ is a time series coefficient. The error term $\sigma_t(x) \epsilon_t(x)$ accounts for observational error that varies with age; i.e., it is the difference between the observed rates and the spline curves. The error term $e_t(x)$ is modelling error; i.e., it is the difference between the spline curves and the fitted curves from the model. It is worth mentioning that in the Lee-Carter model:

$$\log m_t(x) = a(x) + k_t b(x) + \epsilon_t(x), \quad (2)$$

we have one set of $(k_t, b(x))$, whereas in the Hyndman-Ullah model more than one set of components are used. This extension gives more flexibility to model; because these additional components capture non-random patterns, which are not explained by the first principal component. Other extension that Hyndman-Ullah model has been added to the original Lee-Carter model are discussed in Booth *et al.* (2006) and Shang *et al.* (2011).

3. SSA

The basic SSA method consists of three complementary stages: decomposition, reconstruction and forecasting. At the first stage, we decompose the series and at the second stage, the noise free series would be reconstructed and the reconstructed series would be used for forecasting new data points in the final stage. A short description of the SSA technique is given below (for more information, see; Golyandina *et al.*, 2001 and Hassani, 2007).

Stage I. Decomposition

1st step: Embedding. Embedding can be considered as a mapping that transfers a one-dimensional time series $Y_T = (y_1, \dots, y_T)$ into the multi-dimensional series X_1, \dots, X_K with vectors $X_i = (y_i, \dots, y_{i+L-1})^T \in \mathbf{R}^L$, where L ($2 \leq L \leq T - 1$) is window length and $K = T - L + 1$. The result of this step is called trajectory matrix:

$$\mathbf{X} = [X_1, \dots, X_K] = (x_{ij})_{i,j=1}^{L,K}. \quad (3)$$

Note that the trajectory matrix \mathbf{X} is a Hankel matrix, which means that all elements along the diagonal $i + j = \text{const}$ are equal. Let us show the results of this step for $Y = [1, 2, 3, 4, 5, 6]$ and $L = 3$. Trajectory matrix for this example is given by:

$$\mathbf{X} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \end{bmatrix}.$$

2nd step: Singular Value Decomposition (SVD). In this step, we perform the SVD of \mathbf{X} . Denote by $\lambda_1, \dots, \lambda_L$ the eigenvalues of $\mathbf{X}\mathbf{X}^T$ arranged in the decreasing order and by U_1, \dots, U_L the corresponding eigenvectors. The SVD of \mathbf{X} can be written as $\mathbf{X} = \mathbf{X}_1 + \dots + \mathbf{X}_L$, where $\mathbf{X}_i = \sqrt{\lambda_i} U_i V_i^T$ and $V_i = \mathbf{X}^T U_i / \sqrt{\lambda_i}$ (if $\lambda_i = 0$ we set $\mathbf{X}_i = 0$). Applying SVD on trajectory matrix of the previous example provides bellow decomposition:

$$\begin{aligned} \mathbf{X} &= \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \end{bmatrix} = \mathbf{X}_1 + \mathbf{X}_2 \\ &= \begin{bmatrix} 1.54 & 2.25 & 2.95 & 3.66 \\ 2.08 & 3.04 & 3.99 & 4.95 \\ 2.62 & 3.83 & 5.03 & 6.24 \end{bmatrix} + \begin{bmatrix} -0.54 & -0.25 & 0.05 & 0.34 \\ -0.08 & -0.04 & 0.01 & 0.05 \\ 0.38 & 0.17 & -0.03 & -0.24 \end{bmatrix}. \end{aligned}$$

Stage II. Reconstruction

1st Step: Grouping. The grouping step corresponds to splitting the elementary matrices into several groups and summing the matrices within each group. Let $I = \{i_1, \dots, i_p\}$ be a group of indices i_1, \dots, i_p . Then the matrix \mathbf{X}_I corresponding to the group I is defined as $\mathbf{X}_I = \mathbf{X}_{i_1} + \dots + \mathbf{X}_{i_p}$. The split of the set of indices $\{1, \dots, L\}$ into disjoint subsets I_1, \dots, I_m corresponds to the representation $\mathbf{X} = \mathbf{X}_{I_1} + \dots + \mathbf{X}_{I_m}$. The procedure of choosing the sets I_1, \dots, I_m is called the grouping. Note that in the previous example we have only two non zero components and therefore we can define at most two groups: $I_1 = \{1\}, I_2 = \{2\}$.

2nd Step: Diagonal averaging (Hankelization). The purpose of diagonal averaging is to transform a matrix to the form of a Hankel matrix which can

be subsequently converted to a time series. Using diagonal averaging, we have: $\mathbf{X} = \widetilde{\mathbf{X}}_{I_1} + \dots + \widetilde{\mathbf{X}}_{I_m}$, where $\widetilde{\mathbf{X}}_{I_j}$ is the hankelized form of \mathbf{X}_{I_j} for $j = 1, \dots, m$. Denoting $\{\tilde{y}_{j_1}, \dots, \tilde{y}_{j_T}\}$, the series correspond to the matrix $\widetilde{\mathbf{X}}_{I_j}$ for $j = 1, \dots, m$ show the reconstructed components of the original series. If $x_{m,n}^{(I_j)}$ is the $(m, n)^{\text{th}}$ entry of the estimated matrix \mathbf{X}_{I_j} then applying diagonal averaging formula follows that:

$$\tilde{y}_{j_i} = \frac{1}{s_2 - s_1 + 1} \sum_{n=s_1}^{s_2} x_{n,i+1-n}^{(I_j)}, \tag{4}$$

where $s_1 = \max\{1, i + 1 - K\}$, $s_2 = \min\{L, i\}$.

Diagonal averaging on the components of previous example are given by:

$$\begin{aligned} \mathbf{X} &= \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \end{bmatrix} = \widetilde{\mathbf{X}}_1 + \widetilde{\mathbf{X}}_2 \\ &= \begin{bmatrix} 1.54 & 2.17 & 2.87 & 3.83 \\ 2.17 & 2.87 & 3.83 & 4.99 \\ 2.87 & 3.83 & 4.99 & 6.24 \end{bmatrix} + \begin{bmatrix} -0.54 & -0.17 & 0.13 & 0.17 \\ -0.17 & 0.13 & 0.17 & 0.01 \\ 0.13 & 0.17 & 0.01 & -0.24 \end{bmatrix}. \end{aligned}$$

Accordingly, reconstructed components are as follow:

$$\mathbf{Y} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 1.54 \\ 2.17 \\ 2.87 \\ 3.83 \\ 4.99 \\ 6.24 \end{bmatrix} + \begin{bmatrix} -0.54 \\ -0.17 \\ 0.13 \\ 0.17 \\ 0.01 \\ -0.24 \end{bmatrix}.$$

Stage III. Forecasting

The basic requirement for having SSA forecasting is that the series satisfies in a linear recurrent formula (LRF). Recall that series $Y_T = [y_1, \dots, y_T]$ satisfies LRF of order d if:

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + \dots + a_d y_{t-d}, \quad t = d + 1, \dots, T. \tag{5}$$

The series governed by LRFs admits natural recurrent continuation since each term of such a series is equal to a linear combination of several preceding terms. Of course, the coefficients of this linear combination can be used for the continuation as well.

An essential result for SSA forecasting was obtained in Danilov (1997a, 1997b) and developed in Nekrutkin (1999). There are several versions of univariate SSA forecasting algorithm, however, we consider the main version called Recurrent SSA (Introduced by Danilov, 1997a, 1997b). Here, we use abbreviation RSSA for this approach. In what follows, we give a brief description of this method, for more information see Golyandina *et al.* (2001).

Assume U_j^∇ denotes the vector of the first $L-1$ components of the eigenvector U_j and π_j is the last component of U_j ($j = 1, \dots, r$). Define coefficient vector \mathfrak{R} as follow:

$$\mathfrak{R} = \frac{1}{1-v^2} \sum_{j=1}^r \pi_j U_j^\nabla,$$

where $v^2 = \sum_{j=1}^r \pi_j^2$.

Considering the above notations, the RSSA forecasts $(\hat{y}_{T+1}, \dots, \hat{y}_{T+M})$ can be obtained by the following formula:

$$\hat{y}_i = \begin{cases} \tilde{y}_i, & i = 1, \dots, T, \\ \mathfrak{R}^T Z_i, & i = T+1, \dots, T+M, \end{cases} \quad (6)$$

where $Z_i = [\hat{y}_{i-L+1}, \dots, \hat{y}_{i-1}]^T$ and $\tilde{y}_1, \dots, \tilde{y}_T$ are the values of reconstructed series and can be obtained by (4) with $\mathbf{X}_{I_j} = \sum_{i=1}^r \mathbf{X}_i$. For the considered numerical example, let $r = 1$. Then we have obtained $\mathfrak{R} = [0.60, 0.81]^T$ which produces the first ahead forecasts by $\hat{y}_7 = 0.81\hat{y}_6 + 0.60\hat{y}_5 = 0.81 \times 6.24 + 0.6 \times 4.99 = 7.99$.

3.1 Forecast Accuracy

To evaluate the accuracy and reliability of the forecasts, one can use a suitable combination of the following three approaches:

- (a) Construction of confidence intervals;
- (b) Assessment of retrospective forecasts;
- (c) Checking the stability of forecasts.

Despite we do not dismiss the approaches (a) and (c), in the present paper we only concentrate on the approach (b); that is, on the assessing the retrospective forecasts. For more information on approach (a), we refer to Golyandina *et al.* (2001). Moreover, for approach (c), Pepelyshev *et al.* (2010) formalized the problem of assessing the stability of long-horizon forecasting of a given family of forecasting techniques and argued that SSA could be a suitable family of techniques to consider in some applications.

Retrospective forecasts (approach (b)) are performed by truncating the series and forecasting values at the points temporarily removed. These forecasts can then be compared with the observed values of the time series for making an assessment of the quality of the forecasts. Let $e_{T,h}(x) = y_{T+h}(x) - \hat{y}_{T,h}(x)$ denote the forecast error, where $\hat{y}_{T,h}(x)$ shows the forecasts of $y_{T+h}(x)$ using SSA ($h = 1, \dots, M$). Then, we define the Integrated Squared Forecast Error as below:

$$ISFE_{T,h} = \sum_x e_{T,h}^2(x), \quad (7)$$

and use as a measure of accuracy.

3.2 SSA Parameter Selection

The SSA steps show that it depends upon two basic, but very important, parameters:

- the window length, L ,
- the number of singular values, r , that need to be selected for filtering time series.

Choosing improper values of parameters L or r yields incomplete reconstruction and misleading results in forecasting. In spite of the importance of the choosing parameters, no theoretical solution has yet been proposed to this problem. Of course, there are worthwhile efforts and various techniques for selecting the appropriate value of L (see, for example, Golyanidina, 2010; Hassani *et al.*, 2011; Mahmoudvand and Zokaei, 2012 and Mahmoudvand *et al.*, 2013). Considering theoretical results for the structure of the trajectory matrix and separability, it seems mostly suitable for decomposition stage to propose L as close as one-half of the time series length but it can not guaranteed you the best predictions (See for example, Mahmoudvand *et al.*, 2013).

Accordingly, as mentioned above, there are not enough algebraic and theoretical materials for choosing optimal L and r . Certainly, the choice of parameters depends on the available data and the analysis that we want to perform. Here, we would like to use SSA as a method of forecasting. Therefore, one of the best criteria is based on the forecasting errors.

However, note that different approaches can be considered in order to find the error of forecasts. We propose to use a typical Mean Squared Forecast Error, defined as:

$$MSE_{T,h}(x) = \frac{1}{M_{T,h}} \sum_{j=1}^{M_{T,h}} e_{T+j-1,h}^2(x), \quad (8)$$

where $M_{T,h}$ is the number of h -step forecasts that can be obtained by using T

observations of mortality series for age x . To choose the parameters of the SSA, we will apply the values of the SSA parameters that minimize this measure, i.e:

$$(L, r) = \underset{r < L < T-1}{\operatorname{argmin}} \operatorname{MSE}_{T,h}(x). \quad (9)$$

Therefore, we may obtain different parameters for every year and age.

4. Results

We demonstrate the feasibility of SSA for forecasting mortality rates using one application involving demographic data age specific mortality rates. In this case, we have $y_t(x) = \log(m_t(x))$ where $m_t(x)$ denotes the mortality rate for age x in year t .

4.1 Empirical Results: The Case of France

Annual French mortality rates (1899-2001) for single years of age were obtained from the Human Mortality Database (<http://www.mortality.org/>). These rates are simply the ratio of death counts to population exposure in the relevant interval of age and time. Figure 1 shows the variant pattern of log mortality rates for several ages and years in France.

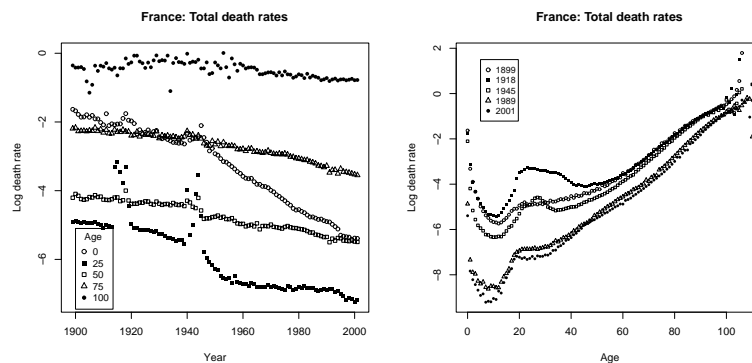


Figure 1: Changes in the total log death rates with respect to both age and year over the period 1899-2001

Depicted plots in Figure 1 show that from 1900 to 2000, there was a general pattern of decline in mortality rates for all ages. Comparison of curves in left plot shows that the severity of decline in infant mortality is higher than adults. Moreover, the effect of two world war I and II on increasing the death rates have appeared in the form of outliers in series 25 years old in the left plot. But this effect is not meaningful in other series. Structural changes can be seen from the shape of the log death rates for the younger ages; whereas deviation from linear

models decreases by age. This means that we may need to use more components in SSA to capture non-linearity of the series of younger ages. Right plot in Figure 1 shows the difference among overall death rates in 1899, 1918, 1945, 1989 and 2001. This plot indicates that the overall death rates have declined by years over these periods. Relationship between death rate and age is non-linear from age 0 to about 55 and the models seem to be close to the linear forms for the other ages.

A more complete description and comparison of France mortality over the past century can be found for example in Caselli *et al.* (1987).

Comparison

We compare our results with those obtained using the method of Hyndman and Ullah (2007). So, we consider the data and cases that Hyndman and Ullah have used in their comparisons. The methods are applied to the French mortality data for years $1899 - m$ and we forecast years $m + 1, \dots, \min(2001, m + 10)$. In order to do a comprehensive comparison, we have considered three different cases, as follows:

- (i) Death rates for all ages, over period 1899- m , are considered and the results of forecasting by both Hyndman-Ullah model and SSA have compared.
- (ii) Death rates for each ages, over period 1899- m , are considered as an individual time series and the results of forecasting by both Hyndman-Ullah model and SSA for several ages have compared individually.
- (iii) A ten-ahead forecasts for period 1992-2001 are obtained by both methods for several ages and compared.

In the cases (i) and (ii), we have done analysis by considering $m = f, f + 1, \dots, 2000$ for $f = 1959, 1969, 1979$ and 1989. Calculations of the Hyndman-Ullah model have done using package demography for R and for SSA we provided several codes in R. The forecasts are compared with actual values and we average the MISE, which is the mean of squared errors that integrated by age, (on the log scale) over $m = f, f + 1, \dots, 2000$ and can be computed by the following formula:

$$MISE_{T,h} = \sum_{x=0}^{\max x} MSE_{T,h}(x). \quad (10)$$

The way of obtaining $MSE_{T,h}(x)$ is explained by an example in Table 1. Note that, the last considered year for this study is 2001. Therefore, only forecasts till this year were obtained. Notation n.a in Table 1 is an abbreviation of not applicable.

Table 1: The way of obtaining $MSE_{T,h}(x)$ for $T = 1989$

Step	m										$MSE_{T,h}(x)$		
	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998		1999	2000
1	$\hat{y}_{1989,1}(x)$	$\hat{y}_{1990,1}(x)$	$\hat{y}_{1991,1}(x)$	$\hat{y}_{1992,1}(x)$	$\hat{y}_{1993,1}(x)$	$\hat{y}_{1994,1}(x)$	$\hat{y}_{1995,1}(x)$	$\hat{y}_{1996,1}(x)$	$\hat{y}_{1997,1}(x)$	$\hat{y}_{1998,1}(x)$	$\hat{y}_{1999,1}(x)$	$\hat{y}_{2000,1}(x)$	$\frac{1}{12} \sum_{T=1989}^{2000} e_{T,1}^2(x)$
2	$\hat{y}_{1989,2}(x)$	$\hat{y}_{1990,2}(x)$	$\hat{y}_{1991,2}(x)$	$\hat{y}_{1992,2}(x)$	$\hat{y}_{1993,2}(x)$	$\hat{y}_{1994,2}(x)$	$\hat{y}_{1995,2}(x)$	$\hat{y}_{1996,2}(x)$	$\hat{y}_{1997,2}(x)$	$\hat{y}_{1998,2}(x)$	$\hat{y}_{1999,2}(x)$	n.a	$\frac{1}{11} \sum_{T=1989}^{1999} e_{T,2}^2(x)$
3	$\hat{y}_{1989,3}(x)$	$\hat{y}_{1990,3}(x)$	$\hat{y}_{1991,3}(x)$	$\hat{y}_{1992,3}(x)$	$\hat{y}_{1993,3}(x)$	$\hat{y}_{1994,3}(x)$	$\hat{y}_{1995,3}(x)$	$\hat{y}_{1996,3}(x)$	$\hat{y}_{1997,3}(x)$	$\hat{y}_{1998,3}(x)$	n.a	n.a	$\frac{1}{10} \sum_{T=1989}^{1998} e_{T,3}^2(x)$
4	$\hat{y}_{1989,4}(x)$	$\hat{y}_{1990,4}(x)$	$\hat{y}_{1991,4}(x)$	$\hat{y}_{1992,4}(x)$	$\hat{y}_{1993,4}(x)$	$\hat{y}_{1994,4}(x)$	$\hat{y}_{1995,4}(x)$	$\hat{y}_{1996,4}(x)$	$\hat{y}_{1997,4}(x)$	n.a	n.a	n.a	$\frac{1}{9} \sum_{T=1989}^{1997} e_{T,4}^2(x)$
5	$\hat{y}_{1989,5}(x)$	$\hat{y}_{1990,5}(x)$	$\hat{y}_{1991,5}(x)$	$\hat{y}_{1992,5}(x)$	$\hat{y}_{1993,5}(x)$	$\hat{y}_{1994,5}(x)$	$\hat{y}_{1995,5}(x)$	$\hat{y}_{1996,5}(x)$	n.a	n.a	n.a	n.a	$\frac{1}{8} \sum_{T=1989}^{1996} e_{T,5}^2(x)$
6	$\hat{y}_{1989,6}(x)$	$\hat{y}_{1990,6}(x)$	$\hat{y}_{1991,6}(x)$	$\hat{y}_{1992,6}(x)$	$\hat{y}_{1993,6}(x)$	$\hat{y}_{1994,6}(x)$	$\hat{y}_{1995,6}(x)$	n.a	n.a	n.a	n.a	n.a	$\frac{1}{7} \sum_{T=1989}^{1995} e_{T,6}^2(x)$
7	$\hat{y}_{1989,7}(x)$	$\hat{y}_{1990,7}(x)$	$\hat{y}_{1991,7}(x)$	$\hat{y}_{1992,7}(x)$	$\hat{y}_{1993,7}(x)$	$\hat{y}_{1994,7}(x)$	n.a	n.a	n.a	n.a	n.a	n.a	$\frac{1}{6} \sum_{T=1989}^{1994} e_{T,7}^2(x)$
8	$\hat{y}_{1989,8}(x)$	$\hat{y}_{1990,8}(x)$	$\hat{y}_{1991,8}(x)$	$\hat{y}_{1992,8}(x)$	$\hat{y}_{1993,8}(x)$	n.a	n.a	n.a	n.a	n.a	n.a	n.a	$\frac{1}{5} \sum_{T=1989}^{1993} e_{T,8}^2(x)$
9	$\hat{y}_{1989,9}(x)$	$\hat{y}_{1990,9}(x)$	$\hat{y}_{1991,9}(x)$	$\hat{y}_{1992,9}(x)$	n.a	n.a	n.a	n.a	n.a	n.a	n.a	n.a	$\frac{1}{4} \sum_{T=1989}^{1992} e_{T,9}^2(x)$
10	$\hat{y}_{1989,10}(x)$	$\hat{y}_{1990,10}(x)$	$\hat{y}_{1991,10}(x)$	n.a	n.a	n.a	n.a	n.a	n.a	n.a	n.a	n.a	$\frac{1}{3} \sum_{T=1989}^{1991} e_{T,10}^2(x)$
MSE	$\frac{\sum_{k=1}^{10} e_{1989,k}^2(x)}{10}$	$\frac{\sum_{k=1}^{10} e_{1990,k}^2(x)}{10}$	$\frac{\sum_{k=1}^{10} e_{1991,k}^2(x)}{10}$	$\frac{\sum_{k=1}^9 e_{1992,k}^2(x)}{9}$	$\frac{\sum_{k=1}^8 e_{1993,k}^2(x)}{8}$	$\frac{\sum_{k=1}^7 e_{1994,k}^2(x)}{7}$	$\frac{\sum_{k=1}^6 e_{1995,k}^2(x)}{6}$	$\frac{\sum_{k=1}^5 e_{1996,k}^2(x)}{5}$	$\frac{\sum_{k=1}^4 e_{1997,k}^2(x)}{4}$	$\frac{\sum_{k=1}^3 e_{1998,k}^2(x)}{3}$	$\frac{\sum_{k=1}^2 e_{1999,k}^2(x)}{2}$	0	$\frac{1}{3} \sum_{T=1989}^{1991} e_{T,10}^2(x)$

The results for the first case are shown in Figure 2. This figure shows that our approach are more accurate than Hyndman-Ullah approach in all considered periods for both short and long term forecasting. A similar results can be found in Figure 3 for case (ii). In this case, comparisons have been done for ages 0, 25, 50, 75 and 100. However, we did this analysis on the other ages and observed similar results in almost all ages. These ages are just representatives for all ages. It should be mentioned that, $MSE_{T,h}(x)$ have obtained by using the method of Table 1 for case (ii).

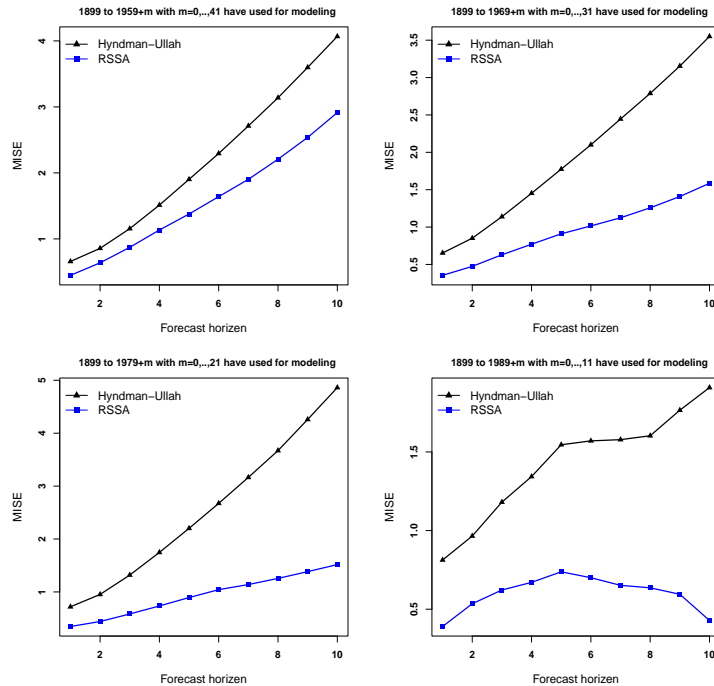


Figure 2: France total log mortality rates forecast accuracy

For the last case, forecasts of log mortality rate for period 1991-2001, using dataset 1899-1991, for both SSA and Hyndman-Ullah approaches, are computed and compared. Here, MSE was obtained by using the last row of Table 1. Figure 4 shows the forecast by both methods as well as real log mortality rates for ages 0, 25, 50, 75 and 100 over the period 1992-2001. This figure indicates that the RSSA is more accurate than Hyndman-Ullah. It also shows that the SSA can distinguishes structural changes of the log mortality series in almost all cases better than Hyndman-Ullah method. In other hand, MSE of forecast in the mentioned ages are computed and reported in Table 2. According to the MSE values, improvements in the mortality forecasts by RSSA are significantly in comparison to the results of Hyndman-Ullah method. Ratios of MSE in the

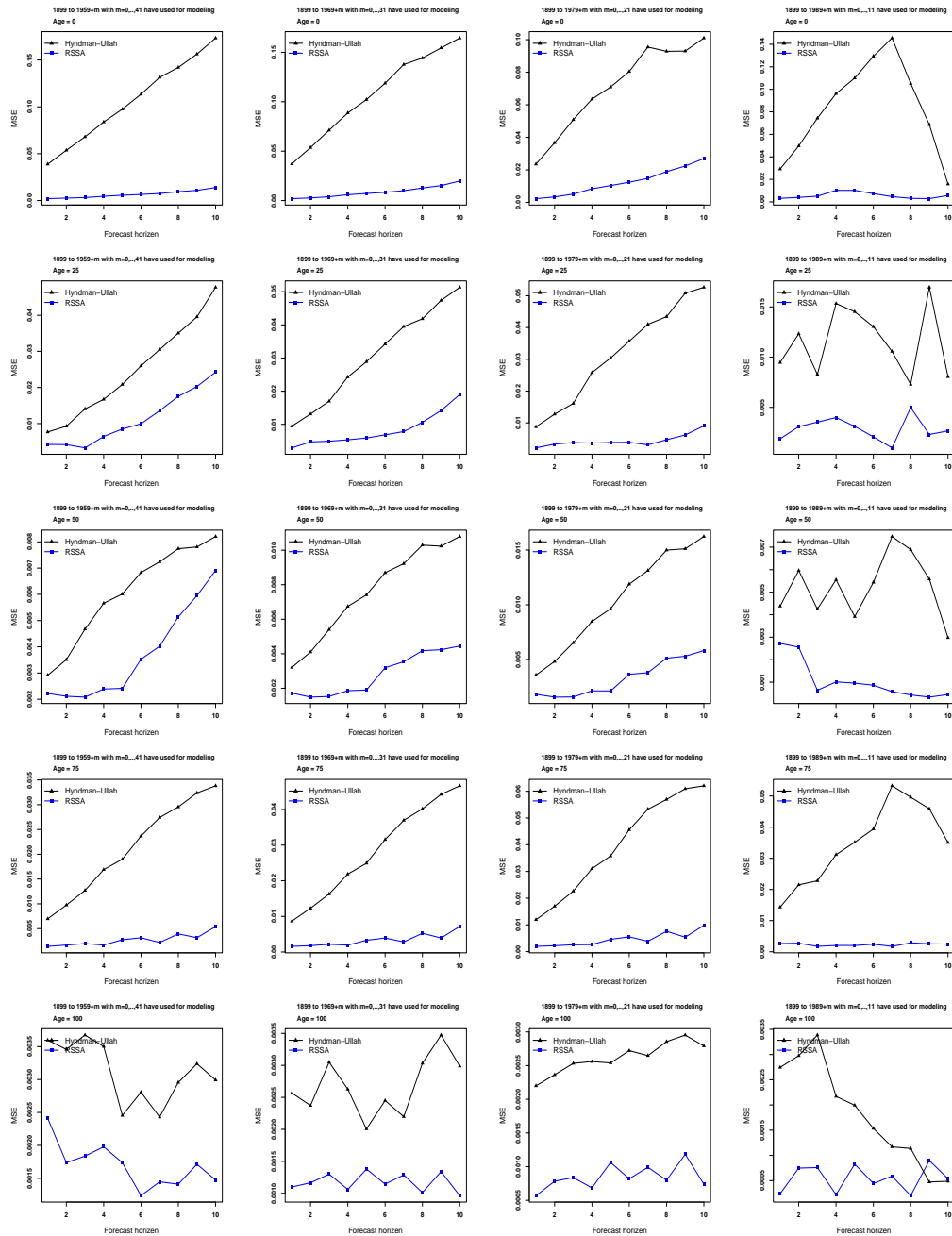


Figure 3: France total log mortality rates forecast accuracy for ages 0, 25, 50, 75 and 100

last row of table shows that RSSA provides more than 52% improvement in log mortality forecast in all cases, which confirms the superiority of SSA.

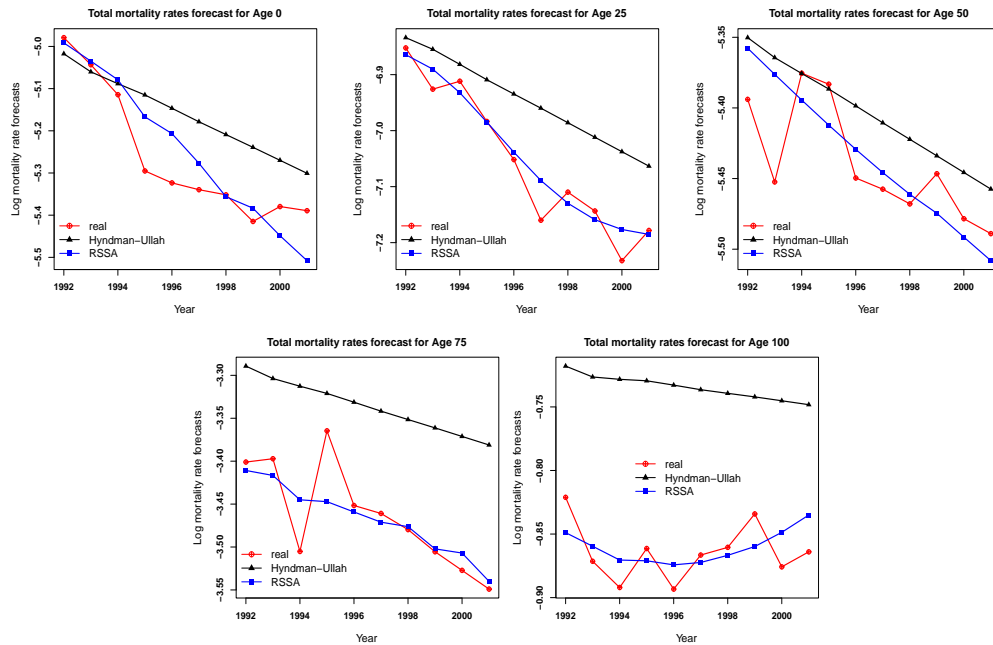


Figure 4: France total log mortality rates forecast accuracy for ages 0, 25, 50, 75 and 100

Table 2: MSE of forecasts for ages 0, 25, 50, 75 and 100 by SSA and Hyndman-Ullah methods

Age	0	25	50	75	100
HU	0.0120	0.0032	0.0019	0.0165	0.0026
RSSA	0.0049	0.0011	0.0009	0.0014	0.0004
RSSA/HU	0.4059	0.3354	0.4780	0.0858	0.1577

Finally, we have done sensitivity analysis with respect to the window length in SSA and K in Hyndman-Ullah model for comparison plan (iii). Results are displayed in Figure 5. As it can be seen from these plots, generally speaking, the sensitivity of results to the window length parameter in SSA is low which again add to the benefits of SSA as a method of mortality forecasting. Whereas, variation in the results of Hyndman-Ullah model is very high in most cases.

5. Conclusion

In this paper, the feasibility of Singular Spectrum Analysis (SSA) for forecasting mortality rates has been studied. The results of SSA have been compared with those of Hyndman and Ullah method. According to the considered data sets, we can conclude that the accuracy of SSA is more than Hyndman and Ullah method.

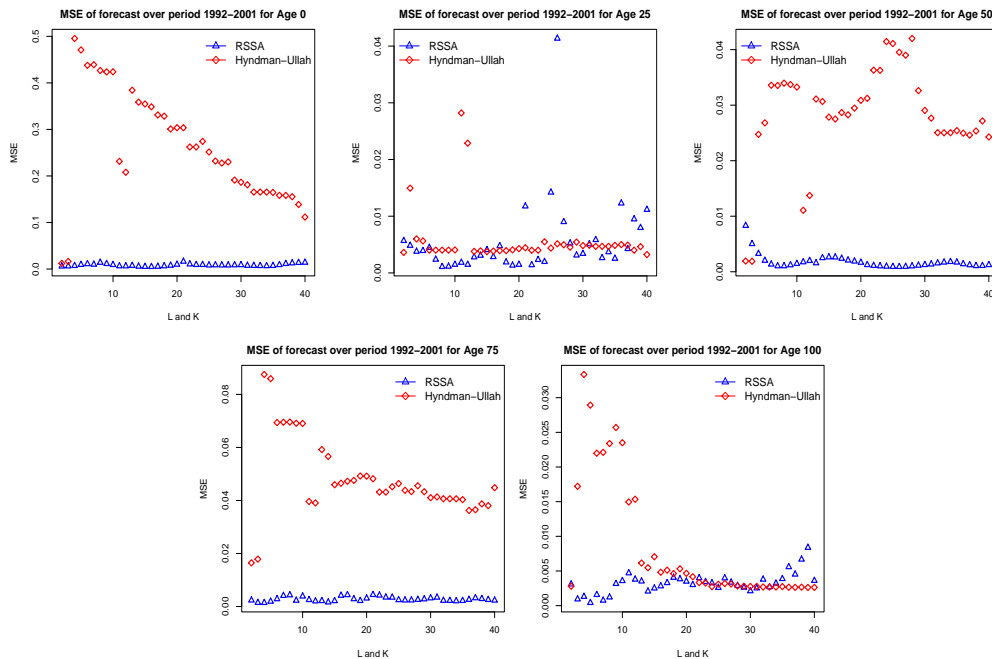


Figure 5: MSE of France total log mortality rates forecast by RSSA and Hyndman-Ullah model for ages 0, 25, 50, 75 and 100

Since the aim of this paper is just introducing a new method, we do not compare the proposed method with other competitive methods here. Hyndman-Ullah, as a relatively powerful method, is considered as our benchmark, only. Further comprehensive analysis is needed to cover all aspects of accuracy criteria and model accuracy statistics. In addition, several other versions of SSA-based forecasting method have been introduced in its literature that can be considered in the future works. Finally, note that our proposed method don't consider the correlations among ages which certainly can add useful information for analyses. Multivariate versions of SSA should be other option that can be applied in this cases. This idea also need to be studied in the future studies.

References

- Booth, H., Hyndman, R. J., Tickle, L. and Jong, P. D. (2006). Lee-Carter mortality forecasting: a multi-country comparison of variants and extensions. *Demographic Research* **15**, 289-310.
- Caselli, G., Vallin, J., Vaupel, J. W. and Yashin, A. (1987). Age-specific mortality trends in France and Italy since 1900: period and cohort effects. *European Journal of Population* **3**, 33-60.

-
- Danilov, D. L. (1997a). Principal components in time series forecast. *Journal of Computational and Graphical Statistics* **6**, 112-121.
- Danilov, D. L. (1997b). The Caterpillar method for time series forecasting. In *Principal Components of Time Series: The Caterpillar Method* (Edited by D. Danilov and A. Zhigljavsky), 73-104. Saint Petersburg State University, Saint Petersburg, Russia.
- Hassani, H. (2007). Singular spectrum analysis: methodology and comparison. *Journal of Data Science* **5**, 239-257.
- Hassani, H., Heravi, S. and Zhigljavsky, A. (2009). Forecasting European industrial production with singular spectrum analysis. *International Journal of Forecasting* **25**, 103-118.
- Hassani, H. and Thomakos, D. (2010). A review on singular spectrum analysis for economic and financial time series. *Statistics and Its Interface* **3**, 377-397.
- Hassani, H., Mahmoudvand, R. and Zokaei, M. (2011). Separability and window length in singular spectrum analysis. *Comptes Rendus Mathematique* **349**, 987-990.
- Hyndman, R. J. and Ullah, M. S. (2007). Robust forecasting of mortality and fertility rates: a functional data approach. *Computational Statistics and Data Analysis* **51**, 4942-4956.
- Hyndman, R. J., Booth, H. and Yasmineen, F. (2013). Coherent mortality forecasting: the product-ratio method with functional time series models. *Demography* **50**, 261-283.
- Golyandina, N., Nekrutkin, V. and Zhigljavsky, A. (2001). *Analysis of Time Series Structure: SSA and Related Techniques*. Chapman & Hall/CRC, New York.
- Lee, R. D. and Carter, L. R. (1992). Modeling and forecasting U. S. mortality. *Journal of the American Statistical Association* **87**, 659-671.
- Mahmoudvand, R. and Zokaei, M. (2012). On the singular values of the Hankel matrix with application in singular spectrum analysis. *Chilean Journal of Statistics* **3**, 43-56.
- Mahmoudvand, R., Najari, N. and Zokaei, M. (2013). On the optimal parameters for reconstruction and forecasting in the singular spectrum analysis. *Communication in Statistics - Simulations and Computations* **42**, 860-870.

Pepelyshev, A. and Zhigljavsky, A. (2010). Assessing the stability of long-horizon SSA forecasting. *Statistics and Its Interface* **3**, 321-327.

Shang, H. L., Booth, H. and Hyndman, R. J. (2011). Point and interval forecasts of mortality rates and life expectancy: a comparison of ten principal component methods. *Demographic Research* **25**, 173-214.

Zhigljavsky, A. (2010). Singular spectrum analysis for time series: introduction to this special issue. *Statistics and Its Interface* **3**, 255-258.

Received April 17, 2013; accepted June 10, 2013.

Rahim Mahmoudvand
Statistical Research and Training Center (SRTC)
No. 5, East 25th St., Asad Abady St., Yousef Abad, Tehran, Iran
r.mahmodvand@gmail.com

Fatemeh Alehosseini
Department of Statistics
Shahid Beheshti University
Evin, Tehran, 1983963113, Tehran, Iran
f_alehosseini@yahoo.com

Mohammad Zokaei
Department of Statistics
Shahid Beheshti University
Evin, Tehran, 1983963113, Tehran, Iran
zokaei@sbu.ac.ir