REVSTAT – Statistical Journal
Volume 13, Number 3, November 2015, 193–206

FORECASTING MORTALITY RATE BY SINGULAR SPECTRUM ANALYSIS

Authors: Rahim Mahmoudvand
– Statistical Research and Training Center (SRTC), Tehran, Iran
r.mahmoudvand@basu.ac.ir

Fatemeh Alehosseini
– Department of Statistics, Shahid Beheshti University, Iran
f_alehosseini@yahoo.com

Paulo Canas Rodrigues
– Department of Statistics, Federal University of Bahia, Salvador, BA, Brazil
and
CAST — Center for Applied Statistics and Data Analytics, University of Tampere, Finland
paulocanas@gmail.com

Received: September 2013 Revised: March 2014 Accepted: April 2014

Abstract:

- Singular spectrum analysis (SSA) is a relatively new and powerful non-parametric time series analysis technique that has demonstrated its capability in forecasting different time series in various disciplines. In this paper, we study the feasibility of using the SSA to perform mortality forecasts. Comparisons are made with the Hyndman–Ullah model, which is a new powerful tool in the field of mortality forecasting, and will be considered as a benchmark to evaluate the performance of the SSA for mortality forecasting. We use both SSA and Hyndman–Ullah models to obtain 10 forecasts for the period 2000–2009 in nine European countries including Belgium, Denmark, Finland, France, Italy, The Netherlands, Norway, Sweden and Switzerland. Computational results show a superior accuracy of the SSA forecasting algorithms, when compared with the Hyndman–Ullah approach.

Key-Words:

- *mortality rate; Singular Spectrum Analysis; Hyndman–Ullah model.*

AMS Subject Classification:

- 37M10, 15A18, 62M15.

1. INTRODUCTION

With the continuing increase in life expectancy, the mortality forecasting plays a major role to advice government policy and planning, and in decision making for pension and insurance industries. Lee and Carter (1992) proposed a new method which uses singular value decomposition to represent the logs of mortality rate in terms of two age-dependent factors and a single time-dependent factor. The time-dependent factor can be extracted and modelled using conventional time series methods so that forecasts could be made. 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 exhibited a good performance of mortality rate forecasts. However producing more accurate forecasts can help, both pension and insurance companies and governments, to make better decisions.

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). SSA has the ability to decompose the original time 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 (e.g. 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) and Golyandina and Zhigljavsky (2013).

Mahmoudvand *et al.* (2013) compared the ability of SSA with the Hyndman–Ullah model for mortality forecast in France. In this paper we extend that study to nine European countries (Belgium, Denmark, Finland, France, Italy, The Netherlands, Norway, Sweden and Switzerland); consider two forecasting algorithms for SSA: Recurrent SSA (RSSA, Danilov, 1997a, b) and Vector SSA (VSSA, Nekrutkin, 1999); and consider the time series until 2009, in a new approach.

Since the proposal of Hyndman and Ullah (2007) can be seen as a benchmark because it achieves more accurate mortality forecasts than many other approaches, it will be used to compare with SSA forecasting results and, therefore, to evaluate SSA as a plausible alternative for mortality forecasting.

The rest of the paper is structured as follows: in Section 2 we give a brief description of Hyndman and Ullah (2007) model, and in Section 3 present the generic SSA methodology. The application is presented Section 4 and Section 5 gives some concluding remarks.

2. HYNDMAN–ULLAH APPROACH

The Hyndman–Ullah approach can be expressed using the equation (Hyndman and Ullah, 2007)

$$\log m_t(x) \;=\; a(x) + \sum_{j=1}^{K} k_{t,j}\, b_j(x) + e_t(x) + \sigma_t(x)\, \epsilon_t(x) \;, \tag{2.1}$$

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. the difference between the spline curves and the fitted curves from the model. By comparison, the Lee–Carter model (Lee and Carter, 1992)

$$\log m_t(x) \;=\; a(x) + k_t\, b(x) + \epsilon_t(x) \;, \tag{2.2}$$

has one set of $(k_t, b(x))$, while the Hyndman–Ullah model includes more than one set of components. This extension presented by Hyndman and Ullah (2007) gives more flexibility to the model because the additional components capture non-random patterns, which are not explained by the first principal component. Other extensions of the Lee–Carter model are discussed in Booth *et al.* (2006) and Shang *et al.* (2011).

3. SINGULAR SPECTRUM ANALYSIS

The basic SSA method consists of three complementary stages: decomposition, reconstruction and forecasting. In the first stage the time series is decomposed, in the second stage the noise free time series is reconstructed and in the third stage the reconstructed time series is used for forecasting new data points. A short description of the SSA technique is given below. More information can be found in Golyandina *et al.* (2001), Hassani (2007) and Golyandina and Zhigljavsky (2013).

3.1. Basic SSA

First Stage: Decomposition

1st step: Embedding. Let $x_1, ..., x_N$ be a time series of length N. Considering a window length L the result of this step is a $L \times K$ matrix $\mathbf{X} = [X_1 : ... : X_K]$, where $K = N - L + 1$ and $X_i = (x_i, ..., x_{i+L-1})^T,\ 1 \leq i \leq K$.

2nd step: Singular Value Decomposition (SVD). In this step, matrix $\mathbf{X}$ will be decomposed using SVD as $\mathbf{X} = \mathbf{X}_1 + \cdots + \mathbf{X}_d$, where $\mathbf{X}_i = \sqrt{\lambda_i} U_i V_i^T$ and $V_i = \mathbf{X}^T U_i / \sqrt{\lambda_i}$ with $\lambda_1, ..., \lambda_L$, the eigenvalues of $\mathbf{S} = \mathbf{X}\mathbf{X}^T$ and $U_1, ..., U_L$, the corresponding eigenvectors.

Second Stage: Reconstruction

3rd step: Grouping. The grouping step corresponds to splitting the elementary matrices into m disjunct subsets $I_1, ..., I_m$, and summing the matrices within each group. In our application we have $m = 2$, i.e. only two groups. $I_1 = \{1, ..., r\}$ and $I_2 = \{r + 1, ..., L\}$ are related to the single and noise components, respectively.

4th step: Diagonal averaging. The purpose of diagonal averaging is to transform each matrix $\mathbf{X}_{I_j}$ into a new series of length N. Using diagonal averaging we have that $\mathbf{X} = \widetilde{\mathbf{X}}_{I_1} + \cdots + \widetilde{\mathbf{X}}_{I_m}$, where $\widetilde{\mathbf{X}}_{I_j}$ is the hankelized form of $\mathbf{X}_{I_j}$, $j = 1, ..., m$. Considering $\tilde{x}_{m,n}^{(I_j)}$ the $(m, n)^{th}$ entry of the estimated matrix $\widetilde{\mathbf{X}}_{I_j}$ and denoting by $\{\tilde{y}_{j1}, ..., \tilde{y}_{jT}\}$ the reconstructed components in the matrix $\widetilde{\mathbf{X}}_{I_j}$, $j = 1, ..., m$, applying diagonal averaging follows that

$$\tilde{y}_{jl} = \begin{cases} \dfrac{1}{s-1} \displaystyle\sum_{n=1}^{s-1} \tilde{x}_{n,s-n}^{(I_j)} & 2 \leq s \leq L - 1 \ , \\ \dfrac{1}{L} \displaystyle\sum_{n=1}^{L} \tilde{x}_{n,s-n}^{(I_j)} & L \leq s \leq K + 1 \ , \\ \dfrac{1}{K + L - s + 1} \displaystyle\sum_{n=n-K}^{L} \tilde{x}_{n,s-n}^{(I_j)} & K + 2 \leq s \leq K + L \ . \end{cases}$$

Third Stage. Forecasting

The basic requirement to make SSA forecasting is that the time series satisfies a linear recurrent formula (LRF). A time series $Y_T = (y_1, ..., y_T)$ satisfies

LRF of order d if:

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + \cdots + a_d y_{t-d}, \qquad t = d+1, ..., T. \tag{3.1}$$

Although there are several versions of univariate SSA forecasting algorithms we consider here two of the mostly widely used: Recurrent SSA (RSSA, Danilov, 1997a, b) and Vector SSA (VSSA, Nekrutkin, 1999). In what follows, we give a brief description of these algorithms. Further details can be found in Golyandina *et al.* (2001).

Let us assume that U_j^∇ is the vector of the first $L-1$ components of the eigenvector U_j and π_j is the last component of U_j $(j = 1, ..., r)$. Denoting $\upsilon^2 = \sum_{j=1}^r \pi_j^2$ we define the coefficient vector $\mathfrak{R}$ as:

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

Recurrent SSA

Considering the above notation, the RSSA forecasts $(\hat{y}_{T+1}, ..., \hat{y}_{T+M})$ can be obtained by

$$\hat{y}_i = \begin{cases} \tilde{y}_i, & i = 1, ..., T, \\ \mathfrak{R}^T Z_i, & i = T+1, ..., T+M, \end{cases} \tag{3.2}$$

where, $Z_i = [\hat{y}_{i-L+1}, ..., \hat{y}_{i-1}]^T$ and $\tilde{y}_1, ..., \tilde{y}_T$, are the values for the reconstructed time series and can be obtained from 4th step in above.

Vector SSA

Define linear operator:

$$\mathcal{P}^{(v)} Y = \begin{pmatrix} \mathbf{\Pi} Y_\Delta \\ \mathfrak{R}^T Y_\Delta \end{pmatrix}, \qquad Y \in \operatorname{span}\{U_1, ..., U_r\}, \tag{3.3}$$

where $\mathbf{\Pi} = \mathbf{U}^\nabla \mathbf{U}^{\nabla T} + (1-\upsilon^2)\mathfrak{R}\mathfrak{R}^T$ and Y_Δ denotes the last $L-1$ elements of Y. Suppose the vector Z_j is defined as follows

$$Z_j = \begin{cases} \widetilde{X}_j & \text{for } j = 1, ..., K, \\ \mathcal{P}^{(v)} Z_{j-1} & \text{for } j = K+1, ..., K+M+L-1, \end{cases} \tag{3.4}$$

where $\widetilde{X}_j$ are the j^{th} reconstructed columns of the trajectory matrix of the time series after grouping and discarding noise components. Now, by constructing the matrix $\mathbf{Z} = [Z_1, ..., Z_{K+M+L-1}]$ and performing diagonal averaging, we obtain a new time series $\hat{y}_1, ..., \hat{y}_{T+M+L-1}$, where $\hat{y}_{T+1}, ..., \hat{y}_{T+M}$ form the M terms of the VSSA forecast.

3.2. Forecast accuracy

To evaluate the accuracy and reliability of forecasts in time series, one can use a suitable combination of the following three approaches: (i) construction of confidence intervals; (ii) assessment of retrospective forecasts; and (iii) checking the stability of forecasts. Although the three represent important approaches, in the present paper we will be interested only in (ii) assessment of retrospective forecasts. Further information about approaches (i) and (iii) can be found in Golyandina *et al.* (2001) and Pepelyshev *et al.* (2010), respectively.

Retrospective forecasts are usually performed by truncating the time series and by obtaining forecasts for points temporarily removed. These forecasts can then be compared with the observed values of the time series to asses their quality and reliability. Let $e_{T,h}(x) = y_{T+h}(x) - \hat{y}_{T,h}(x)$ denote the forecast error, where $\hat{y}_{T,h}(x)$ are the forecasts for $y_{T+h}(x)$ using RSSA or VSSA ($h = 1, ..., M$). Then, a measure of accuracy such as the Integrated Squared Error of forecast can be written as

$$ISE_{T,h} = \sum_x e^2_{T,h}(x) \ . \tag{3.5}$$

3.3. SSA parameter selection

The SSA calibration depends upon two basic, but very important, parameters: the window length L, and the number of eigentriples used for reconstruction r. The choice of improper values for the parameters L or r yield incomplete reconstruction and the forecasting results might be misleading. Despite the importance in choosing proper values for these parameters, no theoretical solution has been proposed to solve this problem. Some of the techniques to choose the appropriate value of L can be found in Golyandina (2010), Hassani *et al.* (2011), Mahmoudvand and Zokaei (2012) and Mahmoudvand *et al.* (2013). An overall agreeable suggestion to choose the window length is to have it close to the middle of the series and proportional to the number of observations per period (e.g. to 12 for monthly time series, to four for quarterly time series, etc.). However, this choice does not guarantee the best predictions (e.g. Mahmoudvand, et al, 2013).

For better results, the parameter choice should be made accordingly to available data and intended analysis.

In practice it is relatively rare that the number of singular values r, needed to be selected to reconstruct noise free series from a noisy time series, is known a priori. Among several ways to determine r described in the literature, the easiest way is done by checking breaks in the eigenvalues spectra. As a rule of thumb, a pure noise series produces a slowly decreasing sequences of singular values. Another useful insight is provided by considering separability between signal and noise components, which is a fundamental concept in studying SSA properties, by using w-correlations (Golyandina *et al.*, 2001) between two vectors $Y^{(1)} = [y_1^{(1)}, ..., y_T^{(1)}]^T$ and $Y^{(2)} = [y_1^{(2)}, ..., y_T^{(2)}]^T$:

$$\rho_w = \frac{\sum_{j=1}^{T} w_j^{L,T} y_j^{(1)} y_j^{(2)}}{\sqrt{\sum_{j=1}^{T} w_j^{L,T} \left(y_j^{(1)}\right)^2 \times \sum_{j=1}^{T} w_j^{L,T} \left(y_j^{(2)}\right)^2}}\ , \tag{3.6}$$

where, $w_j^{L,T} = \min\{j, L, T-j+1\}$ and $2 \leq L \leq T-1$. According to this measure, two series are separable if the absolute value of their w-correlation is small. Therefore, we determine r in such a way that the reconstructed series and residual have a small w-correlation. Another way to determine r is by examining the forecast accuracy, i.e. r is determined in such a way that the minimum error in forecasting will be obtained. Considering L fixed, the choice of r can be done as

$$r = \underset{r<L<T-1}{\operatorname{argmin}}\ ISE_{T,h}(x)\ . \tag{3.7}$$

In this study, we considered $L = 10$ and employed equation (3.7) to determine the number of singular values used for reconstruction, r.

4. RESULTS

Following a preliminary study by Mahmoudvand *et al.* (2013), we intend to demonstrate the feasibility of SSA for forecasting mortality rates using age-specific mortality rates from nine European countries: Belgium, Denmark, Finland, France, Italy, The Netherlands, Norway, Sweden and Switzerland. 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 nine European countries

Annual mortality rates of nine European countries for single years of age were obtained from the Human Mortality Database (http://www.mortality.org/). These mortality rates are the ratios of death counts to population exposure in the relevant interval of age and time. Figure 1 shows the typical patterns of log mortality rates for several ages and years in the considered countries.

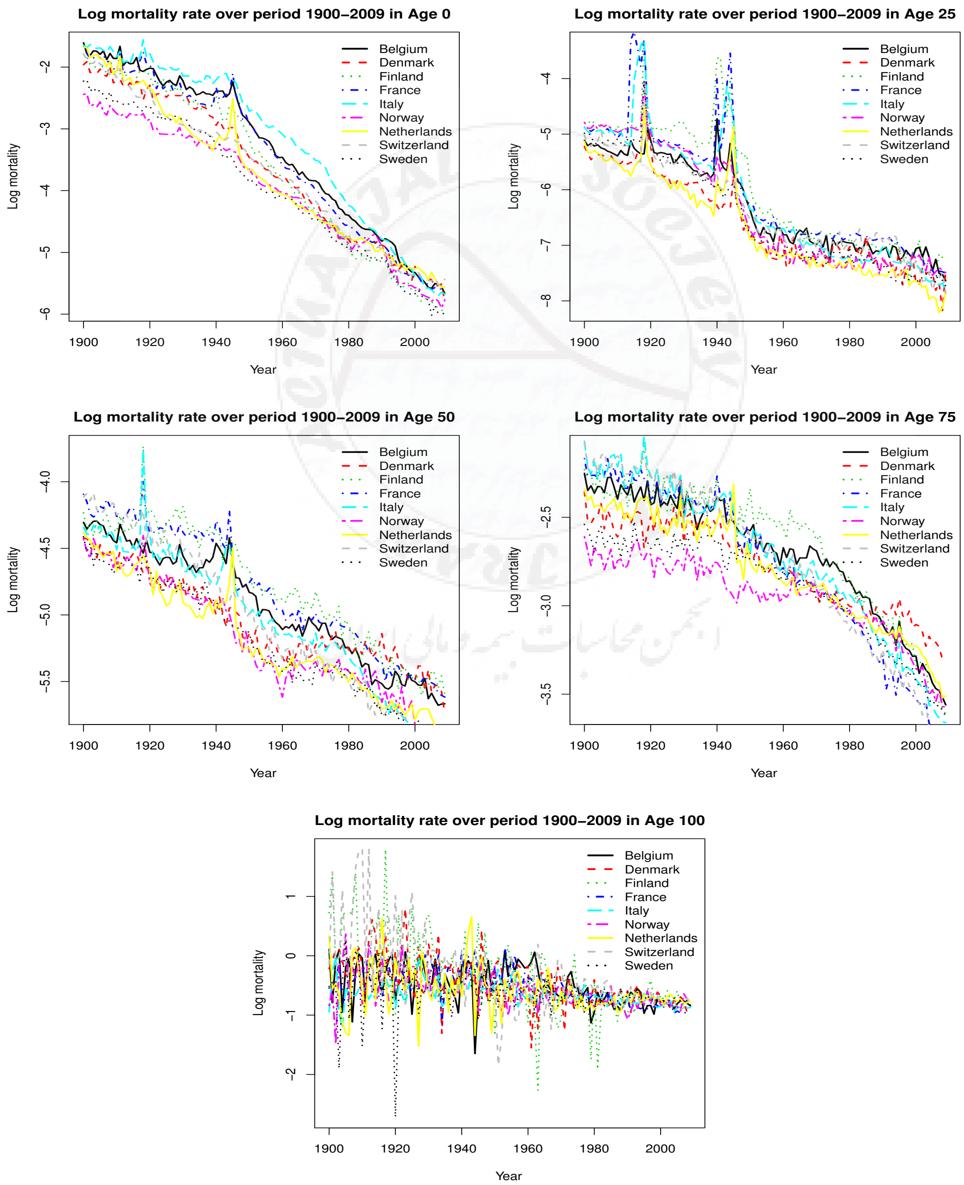

Figure 1: Changes in the total log mortality rates with respect to both age and year, over the period 1900–2009.

The plots depicted in Figure 1 show that, from 1900 to 2009, there was a general pattern of decline in mortality rates for all ages and all countries, as reported by Mahmoudvand *et al.* (2013) for the case of France. By analysing these plots, it can be seem that the decline for infant mortality is steeper than for adult mortality. The effects of the World War I (1914–1918) and World War II (1939–1945) are clearly visible in the top right plot of Figure 1, for the age of 25, being more dramatic for France and Italy, as expected. For the other ages the same effect is also visible but it not as extreme. Since the number of people with 100 years old is small, the bottom plot of Figure 1 shows a less clear pattern but a decrease is visible in terms of mortality rate and variability, with the time, for all countries.

Comparison

The results of our proposal were compared with the results obtained from the method of Hyndman and Ullah (2007). The comparison was made by considering the European mortality data between 1900 and 1999, with forecast for the years 2000–2009. Calculations for the Hyndman–Ullah model were made with the R package "demography" and for SSA we developed our own R code (available upon request). The forecasts were compared with the observed values of the time series, using the integrated squared error (3.5), where the squared errors that integrated by age, on the log scale.

Forecasts of log mortality rate for the period 2000–2009, using the time series 1900–1999, for all ages between 0 and 100 years, for both SSA and Hyndman–Ullah approaches, were computed and compared. This is, for each age, from 0 to 100 years old, and for each of the nine countries, the time series between 1900 and 1999 is used to forecast the next 10 values between 2000 and 2009, which result in the ten ISE of forecasting reported in Table 1. According to the ISE values, results in the mortality forecasts by RSSA and VSSA are significantly better than the results for the Hyndman–Ullah method. Ratios of ISE in the second and third rows for each country of Table 1 show that SSA provides more than 90% improvement in log mortality forecast for some country–year combinations. This confirms the superiority of SSA over the Hyndman–Ullah method. Moreover, the VSSA forecasting procedure is slightly better than the RSSA forecasting procedure, particularly for the long term forecasts. By comparing the results in Table 1 and the plots in Figure 1, it is clear that, because of its construction, HU procedure produces good results when the time series are smoother. However both RSSA and VSSA produce better results when forecasting the most of the mortality rates in these time series.

Table 1: ISE of forecasts for the considered countries.

Country	Model	Year									
		2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Belgium	HU[1]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	RSSA	0.65	0.65	0.38	0.60	0.51	0.82	1.44	0.49	0.38	0.99
	VSSA	0.90	0.59	0.37	0.79	0.39	0.53	1.00	0.38	0.52	1.08
Denmark	HU	4.70	1.68	3.31	2.92	2.29	5.48	3.57	7.80	6.72	6.96
	$\frac{\text{RSSA}}{\text{HU}}$	0.28	0.54	0.28	0.30	0.38	0.24	0.45	0.33	3.20	0.30
	$\frac{\text{VSSA}}{\text{HU}}$	0.30	0.43	0.39	0.24	0.45	0.23	0.39	0.28	3.33	0.26
Finland	HU	4.15	4.95	5.19	4.55	6.08	5.60	5.92	6.17	5.22	8.27
	$\frac{\text{RSSA}}{\text{HU}}$	0.21	0.11	0.20	0.20	0.44	0.89	0.15	0.51	0.41	0.32
	$\frac{\text{VSSA}}{\text{HU}}$	0.25	0.17	0.21	0.21	0.28	0.70	0.19	0.38	0.43	0.30
France	HU	0.50	0.74	0.61	0.93	1.83	1.61	2.13	2.44	2.50	2.15
	$\frac{\text{RSSA}}{\text{HU}}$	0.17	0.21	0.22	0.27	0.17	0.19	0.34	0.34	0.27	0.35
	$\frac{\text{VSSA}}{\text{HU}}$	0.26	0.21	0.23	0.40	0.14	0.16	0.28	0.26	0.21	0.33
Italy	HU	0.70	1.09	1.68	1.35	2.48	2.49	3.37	2.93	2.84	3.39
	$\frac{\text{RSSA}}{\text{HU}}$	0.16	0.10	0.13	0.14	0.21	0.17	0.27	0.29	0.19	0.27
	$\frac{\text{VSSA}}{\text{HU}}$	0.16	0.13	0.12	0.11	0.14	0.18	0.17	0.27	0.15	0.21
Netherlands	HU	1.00	0.73	1.05	1.10	1.83	2.44	2.78	3.41	4.39	3.94
	$\frac{\text{RSSA}}{\text{HU}}$	0.24	0.36	0.23	0.23	0.15	0.26	0.22	0.27	0.30	0.26
	$\frac{\text{VSSA}}{\text{HU}}$	0.23	0.30	0.10	0.32	0.14	0.23	0.19	0.23	0.29	0.22
Norway	HU	3.23	5.76	2.71	2.52	4.86	5.07	10.24	12.55	6.11	8.51
	$\frac{\text{RSSA}}{\text{HU}}$	0.21	0.09	0.58	0.35	0.16	0.21	0.43	0.64	0.20	0.32
	$\frac{\text{VSSA}}{\text{HU}}$	0.27	0.08	0.35	0.32	0.17	0.20	0.34	0.57	0.24	0.37
Sweden	HU	4.52	3.17	4.79	2.82	3.38	4.17	3.85	4.37	8.20	6.07
	$\frac{\text{RSSA}}{\text{HU}}$	0.30	0.17	0.16	0.36	0.23	0.28	0.32	0.27	0.43	0.40
	$\frac{\text{VSSA}}{\text{HU}}$	0.29	0.17	0.26	0.27	0.18	0.17	0.19	0.13	0.43	0.25
Switzerland	HU	1.54	6.46	4.94	3.33	2.72	4.05	5.01	4.19	5.48	5.46
	$\frac{\text{RSSA}}{\text{HU}}$	0.23	0.43	0.52	0.32	0.45	0.32	0.37	0.32	0.29	0.40
	$\frac{\text{VSSA}}{\text{HU}}$	0.31	0.51	0.46	0.31	0.40	0.27	0.40	0.27	0.29	0.41

[1]Due to a small amount of missing values the HU values were not possible to obtain. Data imputation techniques (e.g. Rodrigues and de Carvalho, 2013) can be used to fill in the missing values.

Figure 2 shows the results of a sensitivity analysis to choose the model parameters, window length L for SSA and K for the Hyndman–Ullah model, where the mean ISE (MISE) of forecasts over the period 2000–2009 is presented. Recall that MISE is provided by:

$$MISE = \frac{1}{M}\sum_{h=1}^{M} ISE_{T,h} \tag{4.1}$$

where M denotes the number of forecasts.

Although the model parameters, $L$for SSA and K for HU, are not directly comparable, it can be seen in the plots of Figure 2 that, for most of the cases, the results of SSA, both RSSA and VSSA, are better than those of Hyndman–Ullah, in terms of MISE. This confirms the ability of SSA for mortality forecasting, being the RSSA slightly better than the VSSA, as visible in Table 1.

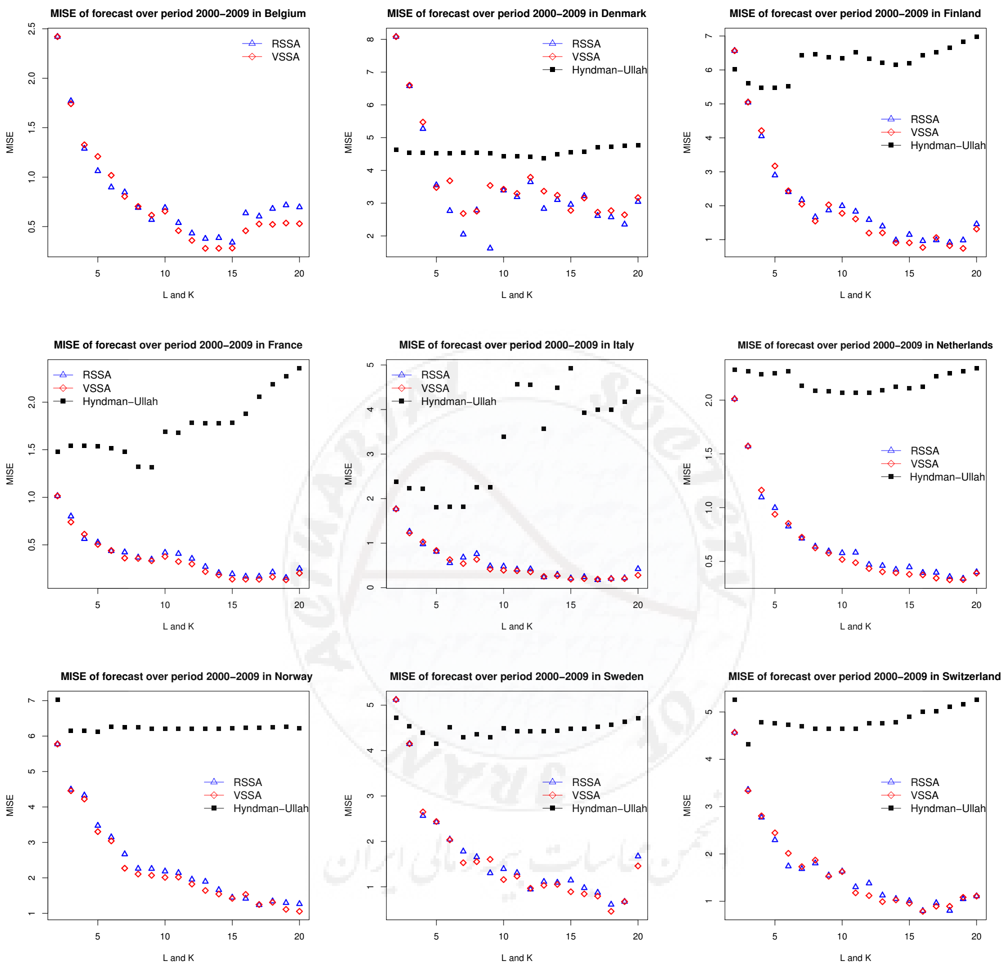

Figure 2: Mean integrated squared error (MISE) of total log mortality rates forecast by RSSA and Hyndman–Ullah model over the period 2000–2009 for different SSA and HU parameters.

5. CONCLUSION

In this paper, the usefulness and ability of Singular Spectrum Analysis (SSA) to forecast mortality rates was studied. The results of SSA based forecasting procedures were compared with those of Hyndman and Ullah method, which can be seen as a benchmark for mortality forecasting. As in the preliminary study presented by Mahmoudvand *et al.* (2013) in this field of research, we can also conclude that the forecasting accuracy of SSA is higher than the forecasting accuracy of the Hyndman and Ullah method, for most of the cases. Within the two SSA based approaches, the RSSA shows slightly better results than the VSSA.

It should be noticed that our proposal does not take into consideration the correlations among ages, which certainly can add useful informations to the analyses and improve the forecast accuracy. Multivariate versions of SSA would be a valid alternative to deal with such correlations and should be considered in further studies. Other alternatives for further improvement of mortality forecasting might be achieved when considering other SSA based forecasting algorithms.

REFERENCES

[1] Booth, H.; Hyndman, R.; Tickle, L. and Jong, P.D. (2006). Lee-Carter mortality forecasting: a multi-country comparison of variants and extensions, *Demographic Research*, **15**, 289–310.

[2] Caselli, G.; Vallin, J.; Vaupel, J.W. and Yashin, A.I. (1987). Age-specific mortality trends in France and Italy since 1900: period and cohort effects, *European Journal of Population*, **3**, 33–60.

[3] Danilov, D. (1997a). Principal components in time series forecast, *Journal of Computational and Graphical Statistics*, **6**, 112–121.

[4] Danilov, D. (1997b). The Caterpillar method for time series forecasting. In "Principal Components of Time Series: the Caterpillar method" (D. Danilov and A. Zhigljavsky, Eds.), pages 73–104, University of St. Petersburg, St. Petersburg. (In Russian).

[5] Hassani, H. (2007). Singular Spectrum Analysis: Methodology and Comparison, *Journal of Data Science*, **5**, 239–257.

[6] Hassani, H.; Heravi, S. and Zhigljavsky, A. (2009). Forecasting European Industrial Production with Singular Spectrum Analysis, *International Journal of Forecasting*, **25**, 103–118.

[7] Hassani, H.; Mahmoudvand, R. and Zokaei, M. (2011). Separability and Window Length in Singular Spectrum Analysis, *Comptes Rendus Mathematique*, **349**, 987–990.

[8] 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.

[9] Hyndman, R.J.; Booth, H. and Yasmeen, F. (2011). *Coherent mortality forecasting: the product-ratio method with functional time series models*, Working paper 01/11, Monash University, Australia.

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

[11] Golyandina, N.; Nekrutkin, V. and Zhigljavsky, A. (2001). *Analysis of Time Series Structure: SSA and related techniques*, Chapman & Hall/CRC, New York – London.

[12] Golyandina, N. and Zhigljavsky, A. (2013). *Singular Spectrum Analysis for Time Series*, Springer, New York – London.

[13] Lee, R.D. and Carter, L.R. (1992). Modeling and forecasting U.S. mortality (discussion 671–675), *Journal of the American Statistical Association*, **87**, 659–671.

[14] Mahmoudvand, R.; Alehosseini, F. and Zokaei, M. (2013). Feasibility of Singular Spectrum Analysis in the Field of Forecasting Mortality Rate, *Journal of Data Science*, **11**, 851–866.

[15] Mahmoudvand, R.; Najari, N. and Zokaei, M. (2013). On the Parameters for Reconstruction and Forecasting in the Singular Spectrum Analysis, *Communication in Statistics: Simulations and Computations*, **42**, 860–870.

[16] 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.

[17] Nekrutkin, V.V. (1999). Approximation spaces and continuation of time series. In "Statistical Models with Applications in Econometrics and Neibouring Fields" (S.M. Ermakovand and Yu.N. Kashtanov, Eds.), pages 3–32, University of St. Petersburg, St. Petersburg. (In Russian).

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

[19] Rodrigues, P.C. and de Carvalho, M. (2013). Spectral modeling of time series with missing data, *Applied Mathematical Modelling*, **37**, 4676–4684.

[20] Shang, H.L.; Booth, H. and Hyndman, R. (2011). Point and interval forecasts of mortality rates and life expectancy: A comparison of ten principal component methods, *Demographic Research*, **25**, 173–214.

[21] Zhigljavsky, A. (2010). Singular Spectrum Analysis for time series: Introduction to this special issue, *Statistics and Its Interface*, **3**, 255–258.