City Research Online

City St George's, University of London

Citation: Russolillo, M., Giordano, G. & Haberman, S. (2011). Extending the Lee Carter Model: a Three-way Decomposition. Scandinavian Actuarial Journal, 2011(2), pp. 96-117. doi: 10.1080/03461231003611933

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/4034/

Link to published version: https://doi.org/10.1080/03461231003611933

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to City Research Online policy.

City Research Online: http://openaccess.city.ac.uk/ publications@citystgeorges.ac.uk

Extending the Lee Carter Model: a Three-Way Decomposition

Maria Russolillo[1], Giuseppe Giordano
Department of Statistics and Economics
University of Salerno, Italy
mrussolillo@unisa.it; ggiordan@unisa.it

&

Steven Haberman
Faculty of Actuarial Science and Insurance, Cass Business School,
City University, London, UK
s.haberman@city.ac.uk

[1] Corresponding Author: Maria Russolillo, Dip. Scienze Economiche e Statistiche, Università di Salerno, 84084 – Fisciano (Salerno), Italy. Tel. +39 089 963457 Fax. +39 089 962049. E-mail: mrussolillo@unisa.it

Maria Russolillo, Giuseppe Giordano and Steven Haberman

Extending the Lee Carter Model: a Three-Way Decomposition

Scandinavian Actuarial Journal

Abstract
In this paper, we focus on a Multidimensional Data Analysis approach to the Lee-Carter (LC) model of mortality trends . In particular, we extend the bilinear LC model and specify a new model based on a *three-way* structure, which incorporates a further component in the decomposition of the log-mortality rates. A multi-way component analysis is performed using the Tucker 3 model. The suggested methodology allows us to obtain combined estimates for the three modes: i) time, ii) age-groups and iii) different populations. From the results obtained by the Tucker 3 decomposition, we can jointly compare, in both a numerical and graphical way, the relationships among all three modes and obtain a time series component as a leading indicator of the mortality trend for a group of populations. Further, we carry out a correlation analysis of the estimated trends in order to assess the reliability of the results of the three-way decomposition. The model's goodness of fit is assessed using an analysis of the residuals. Finally, we discuss how the synthesised mortality index can be used to build concise projected life tables for a group of populations. An application which compares ten European countries is used to illustrate the approach and provide a deeper insight into the model and its implementation.

Keywords: Biplot, Mortality Forecasting, Singular Value Decomposition, Tucker3 model.

JEL Classification: C49, G20.

1. Introduction and Basic Notation

The Lee-Carter model (Lee and Carter, 1992) represents an extrapolative method based on a multiplicative two-way model which is used with time series analysis in forecasting mortality. Tabeau et al. (2001) have highlighted the statistical features of this method as an association model for a two-way data matrix. The model fitting can be addressed, from both an algebraic and a statistical point of view, by means of the Singular Value Decomposition (SVD – Eckart and Young, 1936). The model used for mortality data is[2]:

$$\ln(y_{ij}) = \alpha_j + \kappa_i \beta_j + \varepsilon_{ij} \qquad i = 1, \cdots, I;\ j = 1, \cdots, J \tag{1}$$

which represents the logarithm of death rates y_{ij} measured over I years and J age-classes, as the sum of an age-specific component α_j, that is independent of time and another component that is the product of a time-varying parameter κ_i, reflecting the trend in the level of mortality, and an age-specific component β_j, that represents how, rapidly or slowly, mortality at each age varies when the general level of mortality changes. The ε_{ij} component denotes the error term, which is assumed to be homoscedastic and normally distributed.

[2] The notation used here is different from the original proposed by Lee and Carter for sake of comparability with the terminology used in the three-way data analysis literature (see Kiers, 2000).

The LC model (1) is usually expressed by considering the mean centred log-mortality rates: (see for example: Lee and Carter, 1992; Renshaw and Haberman, 2003a):

$$\tilde{y}_{ij} = \ln\left(y_{ij}\right) - \alpha_j = \kappa_i \beta_j + \varepsilon_{ij} \qquad i = 1, \cdots, I;\ j = 1, \cdots, J \tag{2}$$

The values of α_j are computed by averaging the log-mortality rates across the I years, for each age-class $j = 1, ..., J$.

The dependent term $\tilde{y}_{ij}$ can be arranged in the rectangular matrix $\tilde{\mathbf{Y}}$ (I×J),

$$\tilde{\mathbf{Y}} = \begin{pmatrix} \tilde{y}_{11} & \tilde{y}_{12} & \cdots & \tilde{y}_{1J} \\ \tilde{y}_{21} & \tilde{y}_{22} & \cdots & \tilde{y}_{2J} \\ \vdots & \vdots & \tilde{y}_{ij} & \vdots \\ \tilde{y}_{I1} & \tilde{y}_{I2} & \cdots & \tilde{y}_{IJ} \end{pmatrix} \tag{3}$$

The SVD of this matrix is stated as follows:

$$\tilde{\boldsymbol{Y}}_{[I \times J]} = \mathbf{U}_{[I \times h]} \mathbf{S}_{[h \times h]} \mathbf{V}'_{[h \times J]} \qquad h \leq \min\{I, J\} \tag{4}$$

In the square brackets there is the size of each array. The scalar h is the rank of $\tilde{\mathbf{Y}}$, $\mathbf{S}$ is a diagonal matrix holding the positive singular values of $\tilde{\mathbf{Y}}$. The matrices $\mathbf{U}$ and $\mathbf{V}$ hold, respectively, the left and right singular vectors $\mathbf{u}_i$ and $\mathbf{v}_j$ forming orthogonal bases for the I rows and the J columns of the matrix $\tilde{\mathbf{Y}}$.

The estimates of κ_i and β_j, in equation (2), are obtained as:

$$\hat{\kappa}_i = u_{i1}; \quad \hat{\beta}_j = \lambda_1 v'_{1j} \tag{5}$$

and they minimize:

$$\sum_i \sum_j \left[\tilde{y}_{ij} - \kappa_i \beta_j\right]^2 \tag{6}$$

The aim of this contribution is to extend the two-way model introduced by Lee and Carter to include the case of a further criterion, so that mortality data can be shaped in a three-dimensional array. When mortality data are derived from a more extensive dataset, there can be auxiliary variables of interest which can be used as aggregating criteria, for example: gender, countries, ethnic groups, causes of death.

In the literature, there are several papers which have considered the modelling and forecasting of population mortality using the LC framework for a number of European Countries such as Belgium (Brouhns et al., 2002), Austria (Carter and Prkawetz, 2001), England and Wales (Renshaw and Haberman, 2003a,b), and Spain (Guillen and Vidiella-i-Anguera, 2005). It is noteworthy, however, that none of these papers take into account more than one country simultaneously.

Whilst the demographic and actuarial literature has emphasised mainly the analytical role of SVD as an estimation method, in a previous work, Giordano et al. (2008), have exploited the descriptive features of the SVD in the light of Explorative Data Analysis (EDA, Tukey, 1977). In particular, they show how the Principal Component Analysis (PCA, Hotelling, 1933) can be applied

to mortality rates, providing a useful synthesis and enhancing the model interpretation. A recent paper that has used a PCA extension to the lee-Carter model is Hatzopoulos and Haberman (2009).

In this paper, a new specification of the LC demographic model is presented. We introduce a three-way model structure which looks at the mortality data aggregated according to three criteria: time, age-class and country. Specifically, we refer to the Tucker 3 decomposition method (Kroonenberg, 1983). The proposed approach allows us to simplify the data structure and to obtain a rank reduced representation.

In the proposed model, the resulting time series component κ_i no longer concerns a single country, but provides an aggregate estimate of a set of countries. It could be used as a benchmark for the individual time series of each country and to build concise projected life tables for the set of countries under consideration.

The paper is organised as follows. Section 2 presents the three-way extension of the Lee Carter model. Section 3 contains an application both to a two-way and a three-way dataset. The three-way dataset involves mortality rates divided by age-group for ten European countries in the period 1950-2000. Section 4 provides some concluding remarks.

2 . Three-way LC Model

Frequently, the crude mortality data can be disaggregated according to several criteria or according to different socio-demographic attributes, for example: gender, countries, ethnic groups, regions, causes of death and so on. We consider the case of a three-way structure.

It would be interesting to assess whether such data can be described as combinations of a reduced number of latent factors. Since the data are classified according to three different criteria, we could analyse them by carrying out as many two-way analyses as the number of elements of the third criterion. Furthermore, we would like to analyse the variability across the third criterion. Thus, we introduce a new specification of the LC model, which considers the new effect as follows:

$$\tilde{y}_{ijl} = \ln\left(y_{ijl}\right) - \alpha_{jl} = \kappa_i \beta_j \gamma_l + \varepsilon_{ijl} \qquad i = 1,\cdots,I;\ j = 1,\cdots,J;\ l = 1,\ldots,L \tag{7}$$

where α_j, β_j and κ_i have the same interpretation as before (see eq. 1 and 2), while γ_l ($l = 1,...,L$) represents the term associated to the generic element of the third criterion (which, henceforth, we refer to as representing different countries).

We note that all of the parameters share information common to the three ways, that is the β_i estimates are influenced both by time and country effects. The left hand term can be arranged in a three way array, for example: (*Time* × *Age-group* × *Country*). The dependent term can be expressed as a deviation from the mean effects computed either for one or two criteria.

The singular value decomposition needs to be reformulated to take into account the new data structure. In order to address this decomposition issue, several algorithms leading to different statistical methods *(i.e. Multiple Factorial Analysis, Generalised Canonical Analysis, PARAFAC model, Tucker's Methods, etc.)* have been proposed in the literature (Carroll 1968, Escofier and Pagès 1990, Kiers 2000 and Kroonenberg 1983). The natural extension of the SVD for the three-way data structure is the *Tucker 3 model* (Tucker, 1966). This method, which is also known as *Three-Mode Principal Component Analysis*, can be seen as a generalisation of SVD.

In the two-way SVD, the singular values are arranged in a diagonal matrix (see eq. 4). In the *Tucker 3* model these values form a three-way array - the *core array* - (see fig.1). However, the relationships with the singular values of the various two-way analysis are not so immediate. In Fig. 1 is shown the three-way decomposition according to the Tucker 3 model. The three dimensional array $\tilde{\mathbf{Y}}$ is decomposed in a three-way core array $\mathbf{\Lambda}$ and in three matrices (**K**, **G**, **B**).

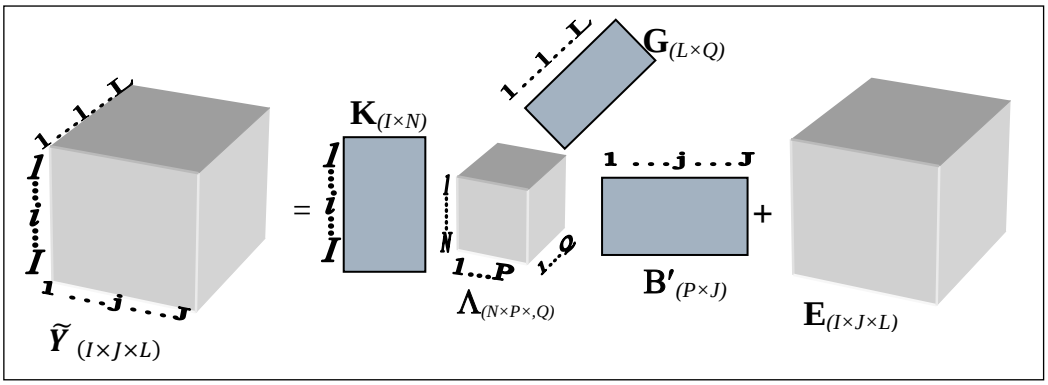


Fig. 1 – The three-way model

Referring back to equation (8), the matrix **K** (*I*×*N*) holds the generic element κ_{in}; the matrix **B** (*J*×*P*) has generic element β_{jp}; and the matrix **G** (*L*×*Q*) has generic entry γ_{lq}. The **K, G** and **B** are column-wise ortho-normal matrices and their columns hold the components associated to each mode.

The *Tucker 3* model factorises the three-way array $\tilde{\mathbf{Y}}$ such that[3]:

$$ln(\tilde{y}_{ijl}) = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{l=1}^{L}\lambda_{npq}\beta_{jp}\kappa_{in}\gamma_{lq} \qquad n=1,...,N;\ p=1,...P;\ q=1,...,Q \qquad (8)$$

The left hand side of equation (8) holds the log-mortality rates expressed as deviations from the average of each age-group *j*. The terms *N*, *P* and *Q* denote the numbers of elements in each of the three modes, respectively years, age-groups, and countries; *I*, *J*, *K* are the numbers of components selected in the approximation. The coefficients $\beta_{jp}, \kappa_{in}, \gamma_{lq}$, are the entries of the component matrices for the first, second and third mode. The λ_{npq} are the elements of the three-way core matrix. By setting *n*, *p*, *q* equal to 1, we obtain the estimates for the model stated in (7) assuming the remaining components as residual term. The patterns of the SVD residuals can be analysed to confirm whether further components may contain useful information in terms of goodness of fit. By retaining further components, we may also obtain two-dimensional maps displaying the relationships among the three modes in the array $\tilde{\mathbf{Y}}$.

As described in detail by Kiers (2000), several possibilities exist for graphical presentation of the model: *i)* to plot the entities of one single mode, *ii)* to plot combinations of two modes and superimpose the entities of the third mode, or *iii)* to plot entities of two modes jointly, for each of the entities of the third mode (the so called "joint plots", see Kroonenberg, 1983, p.164-165). Different tools are available to interpret the results of a three-way decomposition (see Tucker, 1966; Kroonenberg, 1983 for a thorough discussion).

The squared singular values in the core matrix $\mathbf{\Lambda}$ represent the amount of variation accounted for by each component in each mode. These can be summed in order to provide the global variability. When divided by the total variability, the core elements indicate the proportional fit of a component. Moreover, it can be shown (Kroonenberg, 1983) that the total sum of squares can be split into a fitted part and a residual part for each mode. This can be useful, for instance, for indicating how well the model fits the data.

Further insight into the interpretation phase is provided both by the output components and by the core matrix elements as shown by the example in the following section.

[3] The algorithms for estimating the model in eq. (7) are described in *Kroonenberg & De Leeuw* (1980).

3. The Procedure at Work: Exploring Mortality Rates in Ten European Countries

An in-depth illustration of our procedure at work is shown by using the mortality rates from 1950 to 2000 provided by the *Human Mortality Database*[4] for ten different European countries: *Austria* (*Aus), Belgium (Bel), Denmark (Den), Finland (Fin), France (Fra), Italy (Ita), Netherland (Net), Spain (Spa), Sweden (Swe), United Kingdom (UK).* This choice is motivated by the requirement of dealing with a homogeneous data-set (western Europe). The available data for ten countries show no missing values in the considered period.

The data are arranged according to the notation given in section 2. The three-way data array holds the mean centred log-mortality rates for each country. The rows represent the 51 years, the columns are the 21 age-groups: *"0", "1-4", "5-9",..., "95-99"*. The third mode is given by the ten countries. The data can be interpreted as follows: for each country, year and age group, a positive value indicates a mortality rate greater than the average value of the correspondent age-group. Notice that data are centred across observations in the first mode (*years*). Thus, we can compare the results for the individual countries, given by a two-way analysis, with the three-way synthesis.

In the next two subsections, firstly, the PCA is carried out for the two-way Italian mortality dataset. Afterwards, we apply the *Tucker3* method to the extended LC model for our three-dimensional array. In this case, we show several features of the proposed methodology and provide some interpretative comments on the relationships between the age-groups and the different years and countries. The PCA of each two-way dataset allows the estimation and visualisation of the mortality indices (κ_i) and the age-specific component (β_j) for the different countries. The comparison is made easier by the graphical display of the first two components.

As an illustration, we firstly analyse the Italian dataset (total population). This analysis, just applied to one country, is useful to show the peculiar interpretation of the PCA to the LC model (Giordano et al, 2008). Then, the extension to the three-way analysis provides a unique mortality index as a synthesis of the mortality trends in the ten countries. The principal results are reported in the following subsections.

3.1 Two-Way Data Analysis: Italian Data

We carried out the Principal Component Analysis on the variance-covariance matrix obtained for the Italian mortality data. The first eigenvalue accounts for a large amount of inertia (95.74%) showing the high correlation among the age-groups. We also consider the second component which accounts for the 3.13% of the variability. Thus, the first two components in the factorial plan explain 98.88% of the total variation (see Table 1).

Table 1 – The first three eigenvalues: Italian data.

Number	Eigenvalue	Percentage	Cumulated Percentage
1	2.9055	95.74	95.74
2	0.0951	3.13	98.88
3	0.0094	0.31	99.19

The relationships among the age-group variables are showed in Table 2. For each age-group we determine the factor loading. On the first axis, the largest contributions are given by the age groups *"0", "1-4", "5-9", "10-14"* and *"15-19"*. As an indication of the significance of the first component in the LC model, we observe how the factor loadings on the second axis are much smaller than the first axis.

A thorough description of the aforementioned relationship can be observed in the biplot representation (see Fig. 2), which has been obtained by means of the software SPAD®. In this figure, each year is represented as a point, whereas the age-groups are represented as vectors. The factorial map, provided by the first two components, shows how the first component represents the

[4] Data Source: *Human Mortality Database*. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at www.mortality.org or www.humanmortality.de (data downloaded on 2007).

common temporal trend for all age-groups. Indeed, the year-points lie from right to left, ordered from the year 1950 to the year 2000, showing that all age-groups share highly correlated trends.

The second component characterises a diversification of the mortality indices due to the effect of specific age-groups. In particular, on the upper side are located the oldest age groups. On the lower side, we find the central age-groups *30-34, 35-39.*

Thus, for the Italian data, the choice of the first principal component as a composite indicator of all of the age-groups seems to be well supported by the empirical observations.

Table 2 – Factor loadings for the first two components: Italian data.

	Loadings	
Age group	Axis 1	Axis 2
0	*0.82*	*0.07*
1-4	*0.91*	*-0.12*
5-9	*0.60*	*-0.01*
10-14	*0.46*	*0.01*
15-19	*0.25*	*0.00*
20-24	*0.23*	*-0.06*
25-29	*0.25*	*-0.12*
30-34	*0.25*	*-0.13*
35-39	*0.27*	*-0.05*
40-44	*0.27*	*0.03*
45-49	*0.26*	*0.06*
50-54	*0.24*	*0.08*
55-59	*0.21*	*0.08*
60-64	*0.20*	*0.07*
65-69	*0.20*	*0.06*
70-74	*0.21*	*0.06*
75-79	*0.22*	*0.06*
80-84	*0.21*	*0.05*
85-89	*0.18*	*0.05*
90-94	*0.14*	*0.04*
95-99	*0.11*	*0.03*

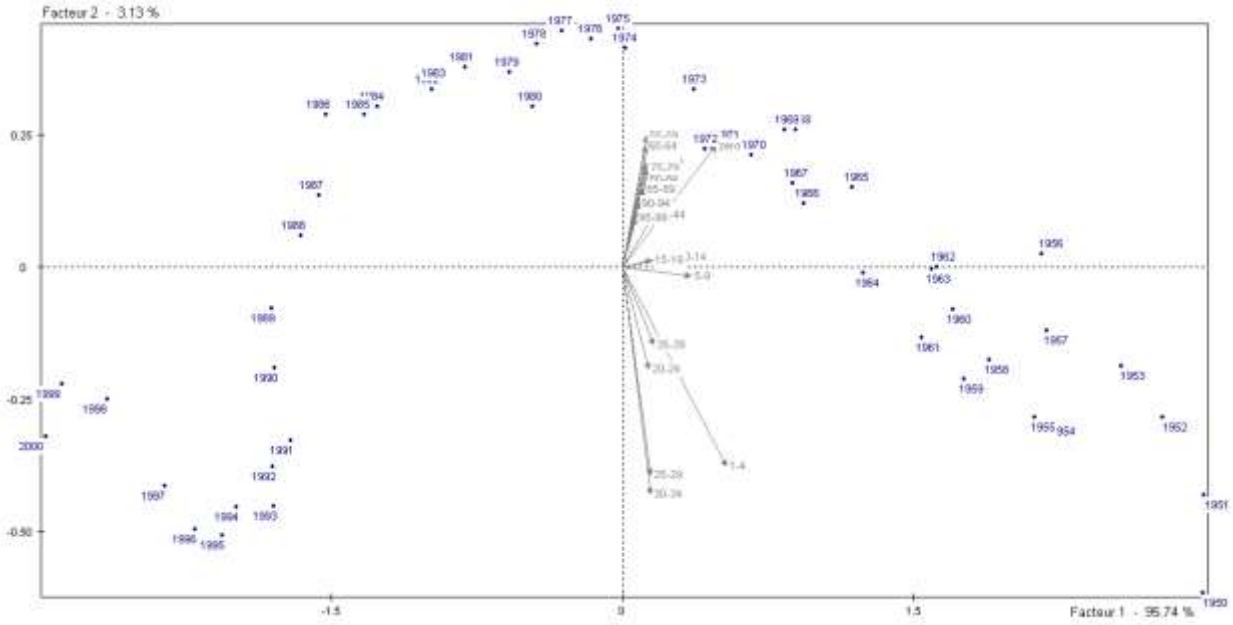


Fig.2 – The first two components of the PCA decomposition: Italian data

3.2. Three-Way Data Analysis: European Data

The three-way data analysis has been carried out by means of the *Matlab*® toolbox *Tucker3* (written by Henk A.L. Kiers, April 2001). We find that the model with the selection of (2×2×2) components gives a good fit, explaining 84.75% of the total variation.

In Table 3, we present the rotated Core Matrix values. The value -31.16 indicates the higher contribution made by the first SVD components for the three modes. In Table 4, the explained variation of each core entry is analysed.

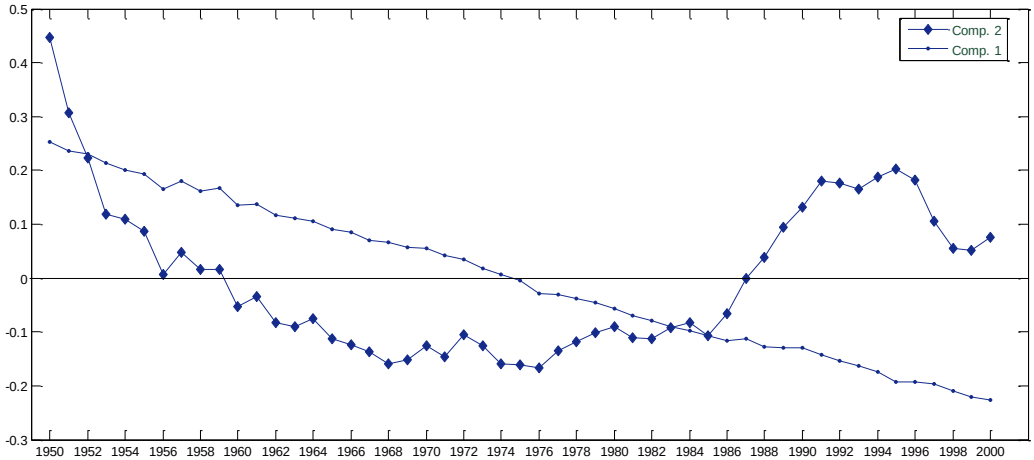
Table 3 – The core matrix elements

	Age 1 × Country 1	*Age 2 × Country 1*	*Age 1 × Country 2*	*Age 2 × Country 2*
Years1	**16.71**	-.001	-0.031	-0.29
Years2	-.00	**3.48**	0.47	-2.13

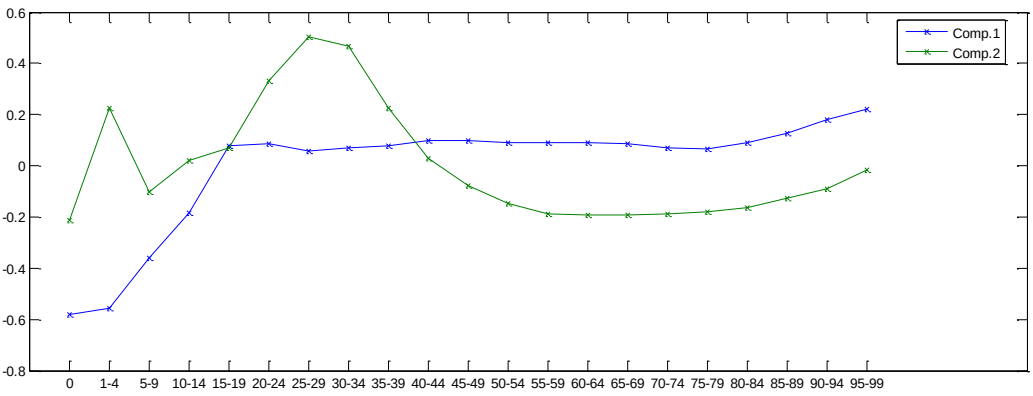
Table 4 – Core Component Analysis: % of importance

	Age 1 × Country 1	*Age 2 × Country 1*	*Age 1 × Country 2*	*Age 2 × Country 2*
Years1	**79.90**	0.00	0.00	0.02
Years2	0.00	**3.47**	0.06	1.30

The first two components are plotted in Fig. 3, respectively, for $\kappa_{i,n=1,2}$ (mode 1) and $\beta_{j,p=1,2}$ (mode 2).



Mode1: Years



Mode2: Age groups

Fig. 3 - Plots of the components for years (mode 1) and age-groups (mode 2).

In the top part of Fig. 3, we note the linear trend of the first component which couldn be used in a subsequent forecasting phase. In the bottom part of Fig. 3, we notice that the first component shows a regular pattern for the age-groups going from 15-19 to 80-84. For this component, we note the different contributions given by the younger age-groups.

After exploring the component scores, it is interesting to analyse in more detail the relationships between the trend (years) and age profiles. Usually, this is performed conditionally on each component of the third mode. In Fig.4, we show the joint plot of *years* and *age-groups* conditional on the first component of *country*. In this representation, the *age-group* components should be thought of as vector-points (loading scores) and the *years* play the role of "statistical units". By considering each age-group as variable, we may project along its direction all of the years-points and analyse the particular trend for that age-group.

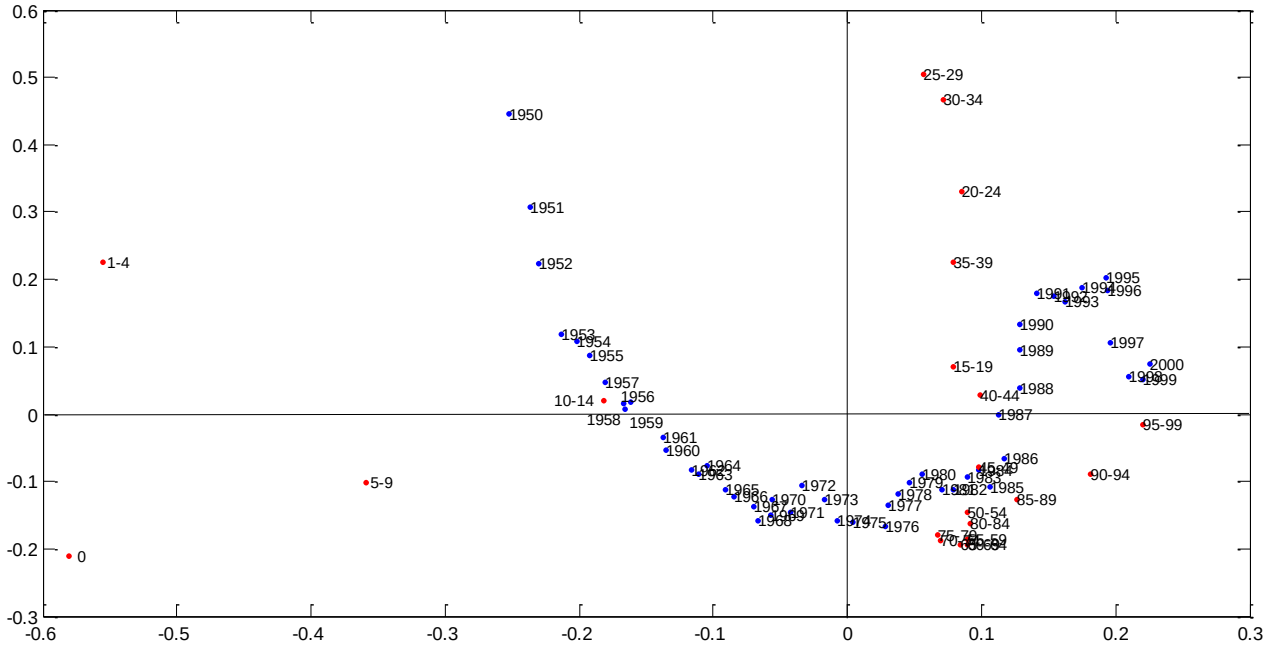


Fig.4 - The joint plot of the three-way mortality data: years and age-groups plotted conditionally to the first component of country.

In Fig. 4, we have plotted the age-groups as a vector-variable. We note that this interpretation is still valid because of the preliminary choice of centring the log-mortality rates across the years component for each age-group. A different pre-treatment would have given different results and hence different interpretations.

The age-groups ranging from "0" to "10-14" seem to explain the greater part of the variability along the horizontal axis. Along the vertical axis, we note a division between the age-groups ranging from 20 to 40 years old and the oldest age-groups from 45 to 94 years old. Let us notice that this result is similar to the two-way analysis shown for the Italian data in Fig.2, but with an axis rotation.

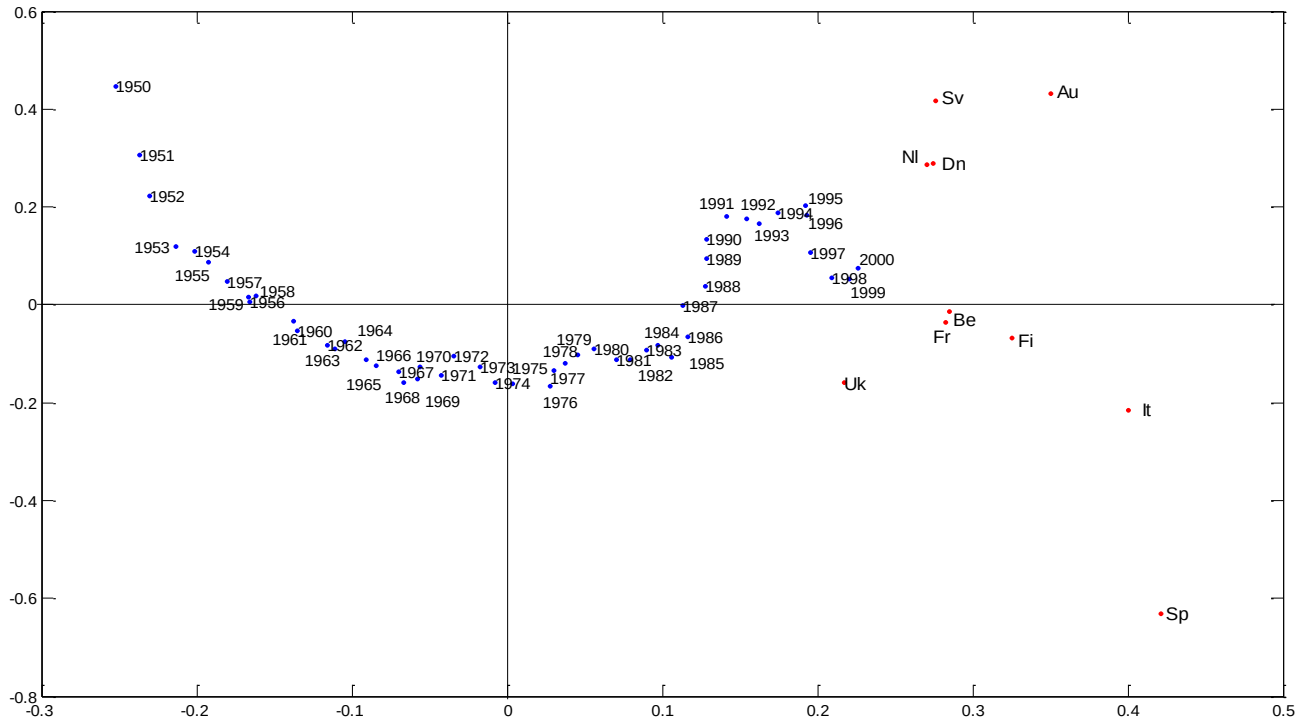


Fig.5 - The joint plot of the three-way mortality data: years and country plotted conditionally to the first component of age-group.

Similarly, Fig.5 shows the bi-plot of *years* and *country* conditional on the first component of the *age-group* mode. We notice that the ten countries show highly correlated patterns. However the estimated trends for Belgium, Finland, France, Italy and UK seem to be better represented on the first axis of the bi-plot. The second axis separates Spain (on the bottom side) from Austria, Sweden, the Netherlands and Denmark (on the top side).

Finally, we show a comparison among the ten European Countries by plotting the κ_{in} $(i = 1950, \dots, 2000; n = 1, \dots, 10)$ estimated for each single country. In Fig.6, the ten trends κ_{in} are compared with the joint estimated trend obtained through the three-way decomposition (shown in Fig. 3).

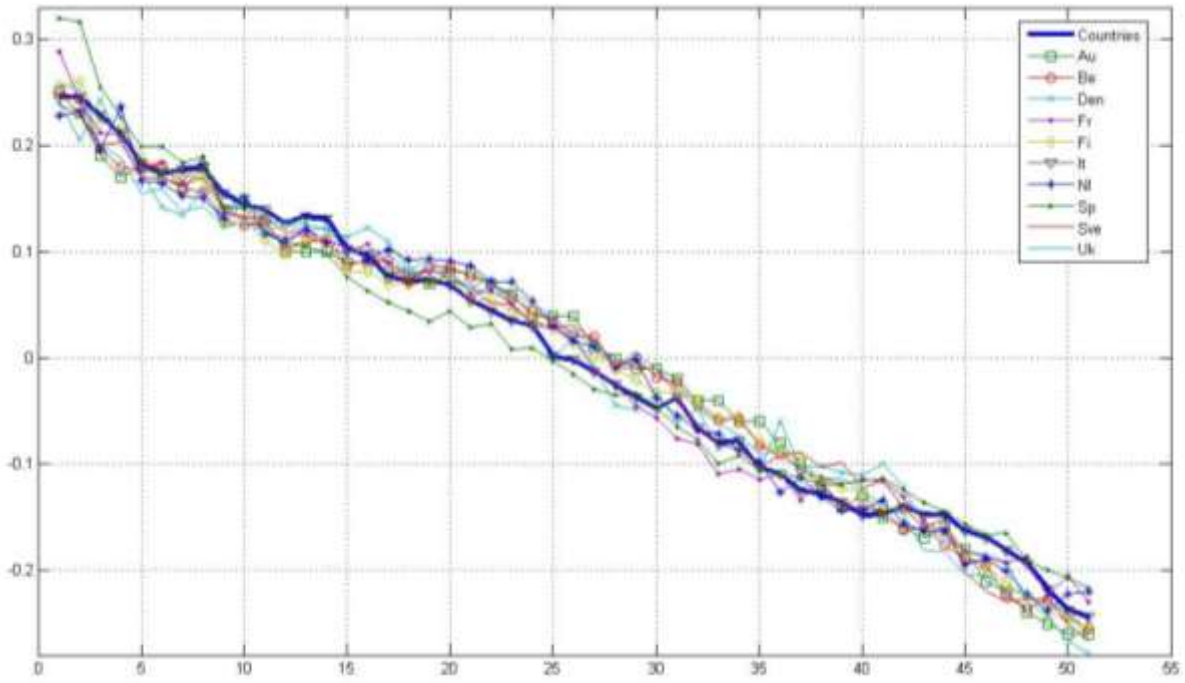


Fig. 6 – The ten κ_i trends and the three-way synthesis (bold line)

From Fig. 6, it is evident that there is a strong trend similarity among the ten series. This enforces the idea that the synthesised κ_i (shown in bold) is a good approximation for the whole series, as confirmed by the correlation matrix in Table 5.

Table 5 – Correlation matrix among the ten estimates κ_i and the three-way synthesis (Countries)

	Countries	**Au**	**Be**	**Dn**	**Fi**	**Fr**	**It**	**Nl**	**Sp**	**Sv**	**Uk**
Countries	1.000										
Au	0.993	1.000									
Be	0.996	0.997	1.000								
Dn	0.991	0.983	0.985	1.000							
Fi	0.995	0.980	0.985	0.986	1.000						
Fr	0.998	0.994	0.996	0.987	0.988	1.000					
It	0.996	0.985	0.988	0.987	0.992	0.992	1.000				
Nl	0.994	0.989	0.993	0.985	0.989	0.991	0.990	1.000			
Sp	0.985	0.962	0.968	0.966	0.985	0.980	0.985	0.968	1.000		
Sv	0.995	0.992	0.994	0.990	0.988	0.994	0.989	0.991	0.970	1.000	
Uk	0.995	0.994	0.997	0.988	0.986	0.994	0.986	0.992	0.966	0.995	1.000

3.2.1 Goodness of Fit Analysis

As regards the assessment of goodness of fit, the sole use of the explained variation is not a satisfactory diagnostic indicator but, as noted by Brouhns et al (2002) and by Renshaw and Haberman (2003a), it has been widely used in the demographic literature for the explorative purpose for choosing the number of relevant components (Tuljapurkar et al., 2000). An useful tool is the boxplot of the residuals, for each age group, against Years and Country. For illustrative purposes, in Fig. 7, we plot the three-way residual patterns for the full set of countries. The box-plots show the residual patterns of (*51x10)* observations for each age-groups. We notice that, for each age-group and each year, the standardised residuals show symmetric and regular patterns, confirming that we have a good fit for the three-way LC method.

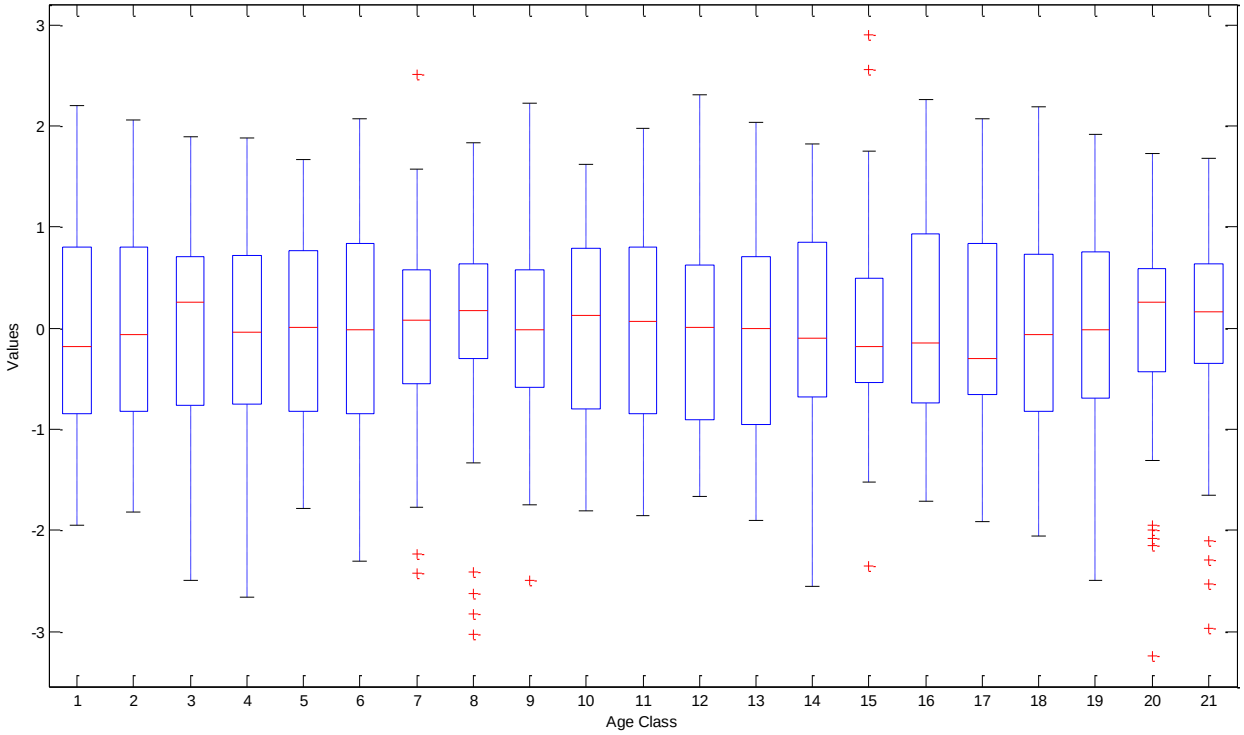


Fig. 7 – Box-plots of the Tucker3 model residuals for each age group from 0, 1-4,..., to 95-99. Each box represents the variability of 510 observations (51 years x *10 countries).*

3.2.2 Discussion on forecasting by the Three-Way Model

The advantages of the Lee-Carter model are essentially twofold: 1. it is a simple model which is combined with time-series analysis for describing the secular changes in mortality, it has been successfully used for forecasting, and the results compete favourably with those from other extrapolative procedures; 2. it allows the discussion of mortality uncertainty through the uncertainty of the time-varying variable k_i (see, for example, Renshaw and Haberman, 2009). The attractive and convenient linear feature of the time-varying variable k_i is purely statistical and varies for different countries and different periods; the reasons behind the statistical features of the parameters vary too. Therefore, care would be needed when the model is used for explanatory purposes. In this subsection, we provide indications around the way that the Three-Way Model extension might be of practical use in terms of the traditional Lee-Carter model application, with the intention of extending this topic in further studies.

As with the traditional LC modelling approach, we could apply a time series analysis to the estimated k_i and produce projected life tables. In our case, we would be interested both in each country and in highlighting the opportunity to use the synthesised $k_{i,}$ derived by the three-way model, in order to generate aggregated life tables. An advantage of computing aggregated life tables would be that such tables would take into account the moderating effect of being based on a synthesised k_i. This effect could be used, for example, for presenting projected life expectancy at birth for the ten countries and for Europe, in aggregate. In future research, we aim to compute period and cohort based life expectancies from forecast mortality rates for each single country by using the particular time series k_i of that country. In this paper, our aim is to extrapolate the three-way LC model in order to obtain a synthesised k_i that could be used for forecasting.

In the actuarial literature, it is well known that mortality improvements have a significant effect on the pricing of and reserving for life annuities. More generally, mortality trends affect any insurance cover that provides some kind of "living benefits". In order to protect the insurance company from the effect of mortality improvements, actuaries have to employ life tables that include a forecast of the future trends of mortality rates (the so-called projected tables). For this

reason, by applying a time series approach to the synthesised k_i derived by our procedure, the idea could be to generate a projected aggregate life table for the ten countries of Europe and a set of projected life tables for the individual countries. These could then be used to extrapolate life expectancies and annuity values aggregated for all ten European countries, as well as individually.

We would expect that the life expectancies and annuity values forecast under the synthesised k_i are intermediate between the ones generated by using the k_i series related to each single country, reflecting the different model structure that generates the respective time series k_i. This possibility remains to be verified.

4. Concluding Remarks

In this paper, we deal with the traditional formulation of the LC model and extend it by introducing a third component in the model specification. We introduce a three-way structure which looks at the mortality data aggregated according to three criteria: time, age-class and country.

In order to deal with the three-way structure, we refer to the Tucker 3 decomposition method. This factorial decomposition is interpreted in the light of a three-way data analysis and allows us to simplify the data structure and obtain a rank reduced representation.

Through the results derived by the Tucker3 decomposition we have compared, in a numerical and graphical way, the relationships among all three modes and extrapolated a time series component as a leading indicator of the mortality for the whole group of countries. By means of our procedure, we have compared the results of the individual countries, on the basis of a two-way analysis, with the three-way synthesis. The extension to the three-way analysis provides a unique mortality index κ_i which no longer concerns a single country, but provides an aggregate estimate for a set of countries.

In the discussion, we suggest how the model proposed in this paper may be used to deal with forecasting. The resulting time series component could be used as a benchmark for the individual time series of each country and to build concise projected life tables for the whole group of countries.

We acknowledge that we have returned to the original version of the LC model and extended the SVD formulation. More recent developments of the model have dealt with other aspects – for example, the introduction of the more natural Poisson error structure (Brouhns et al, 2002; Renshaw and Haberman, 2003a, 2003b).

We believe that the exploratory tools of multivariate data analysis, which we have proposed, give a new perspective to the demographic analysis supporting the analytical results with a geometrical interpretation and a graphical representation.

Acknowledgements

The authors thank an anonymous referee whose suggestions improved the original manuscript.

Appendix. The Matlab Procedure.

In the following is reported the Matlab output to reproduce the three-way analysis and the related plot shown in this paper. It makes use of the Matlab source code written by Henk Kiers in 2001 as reported in his web page http://www.ppsw.rug.nl/~kiers/ where the whole toolbox can be downloaded.
The Tucker3.m must run provided that a Data Matrix **X** is shaped as specified in the note 2 below.
Here is reported the interactive parameters choice in bold **"< >"** symbols.

****

WELCOME to the interactive TUCKER3 analysis program
(written by Henk A.L. Kiers, April 2001)
The program works fully interactively.
Simply answer the questions, or press the Enter key for the default option.

NOTES:
1. The program does not handle missing data:
you should make sure that all missing values are replaced by sensible values
2. The data should be available in the matrix X, which is set up as follows:
- rows correspond to A-mode entities
- columns correspond to combinations of B- and C-mode entities,
with B-mode entities nested within C-mode entities
For example, a 2 x 2 x 3 array is given as follows:
x_1_1_1 x_1_2_1 x_1_1_2 x_1_2_2 x_1_1_3 x_1_2_3
x_2_1_1 x_2_2_1 x_2_1_2 x_2_2_2 x_2_1_3 x_2_2_3
3. The program may change matrix X; keep a copy of X !
4. Tucker3.m works in the workspace, and makes (and hence overwrites) many matrices;
you are advised to BACKUP the contents of your workspace before analysis with Tucker3.m
5. You can specify labels for A- B- and C- mode entities (in laba, labb, labc);
otherwise standard labels will be made by the program, as laba, labb, labc.
6. All matrix output is tab-delimited, and can hence easily be converted to tables, after copying to Word
7. Faster analysis (for advanced users) can be carried out using splithalft3.m
Note: If you are interested in doing a botstrap analysis, it is assumed that the A-mode is the mode from which you want to resampling
If this is not the case, first reorder your data

START of program

Specify the number of A-mode entities: **<51>**
Specify the number of B-mode entities: **<21>**
Specify the number of C-mode entities: **<10>**
To see anova results, specify "1": **<1>**

Total ssq	=	1060.338983	
SS_gm	=	0.000000	(0.00)
SS_a	=	673.790430	(63.54)
SS_b	=	0.000000	(0.00)
SS_c	=	0.000000	(0.00)
SS_ab	=	289.212959	(27.28)
SS_ac	=	37.252262	(3.51)
SS_bc	=	0.000000	(0.00)
SS_abc	=	60.083333	(5.67)

How do you want to CENTER your array ?
1= across A-mode

2= across B-mode
3= across C-mode
12= across A-mode and across B-mode
13= across A-mode and across C-mode
23= across B-mode and across C-mode
Specify your choice**: <12>**
Data have been centered across A-mode and across B-mode
How do you want to NORMALIZE your array ?
1= within A-mode
2= within B-mode
3= within C-mode
Specify your choice: **<enter>**
NOTE:
The preprocessed data are now available in Xprep
Xprep can be used for analyses outside the Tucker3.m program

To see PCA of MEAN, specify "1": **<1>**
Over which mode do you want to compute means? Enter 1,2 or 3 (for A,B, or C)? **<3>**
A 51 by 21 matrix for modes A x B will be computed
Which mode will serve as "variables mode"? Enter 1 (=A) or 2 (=B): **<2>**
B will contain loadings; A will contain scores
If you want to standardize "variables", type "1": **<enter>**
Eigenvalues and cumulative percentages (Scree plot in Figure)

26.89	92.98
1.19	97.09
0.48	98.73
0.12	99.14
0.09	99.45
0.05	99.63
0.04	99.75
0.02	99.84
0.01	99.88
0.01	99.91
0.01	99.94
0.01	99.95
0.00	99.97
0.00	99.98
0.00	99.98
0.00	99.99
0.00	99.99
0.00	100.00
0.00	100.00
0.00	100.00
0.00	100.00

How many components do you want to maintain for rotation? **<2>**
A1, B1 and/or C1 contain solution based on varimax rotation of loadings

To study/plot the output you can now temporarily leave the interactive program
to get back, type "return".

You are now about to do a Tucker3 analysis.
You have to specify the dimensionalities to use.
(see Timmerman & Kiers, 2000, Kiers & der Kinderen, 2003)

If you want to do the PCASUPs, specify "1": **<return>**

You can now do a TUCKER3 ANALYSIS

How many A-mode components do you want to use? **<2>**
How many B-mode components do you want to use? **<2>**
How many C-mode components do you want to use? **<2>**

Tucker3 function value at start is 56.49327452
Final Tucker3 function value after iteration 7 is 53.26369179

Tucker3 analysis with 2x2x2 components gave a fit of 84.75 %

PLOTTING

No special plotting routines are incorporated, but you can use
any scatter-plotting routine you like (e.g., scatter, etc.,or "plotje.m")

If you want to make plots of entities of one mode, you can use:
matrices Aplot, Bplot, or Cplot, and use labels in laba, labb, labc.

If you want to make plots of entities of two modes,
with the third projected in it as axes, you can use:
(BAplot,C),(ACplot,B), or (CBplot,A)

It is sometimes useful to RENORMALIZE components, e.g., such that
the influence of the different A-mode components on the core is equalized
This will usually lead to correlated components in that mode, and
to a more simple structure for the components in the OTHER modes,

If you want to renormalize components, specify "1": **<enter>**

Find useful simple structure rotation of core and components
You can now carry out SIMPLE STRUCTURE rotations with varying weights;
if desired, you can specify a range of different weights for each mode.
To specify a range of values you type, e.g., 1:4 (then the values 1,2,3 and 4 are chosen),
or 1:5:25 (then the values 1,6,11,16 and 21 are chosen),
or [1 2 5 10 100] (then the values 1,2,5,10 and 100 are chosen).
You can also specify a single weight (e.g., just 1).
Analyses will be carried out with all combinations of relative weights.
Specify (range of) relative weight(s) for A: **<enter>**
Specify (range of) relative weight(s) for B: **<enter>**
Specify (range of) relative weight(s) for C: **<enter>**
Simple structure rotation with relative weights **0 0 0**

Results from 5 random starts

Run no. 1 f= 5.289 (core: 5.289; A: 0.000; B: 0.000; C: 0.000), 4 iters
Run no. 2 f= 5.289 (core: 5.289; A: 0.000; B: 0.000; C: 0.000), 4 iters
Run no. 3 f= 5.289 (core: 5.289; A: 0.000; B: 0.000; C: 0.000), 4 iters
Run no. 4 f= 5.289 (core: 5.289; A: 0.000; B: 0.000; C: 0.000), 4 iters
Run no. 5 f= 5.289 (core: 5.289; A: 0.000; B: 0.000; C: 0.000), 3 iters

Three-way orthomax function value for best solution is 5.2891

Varimax values of core and AS, BT and CU (all based on "natural weights")

Core	A	B	C
5.289	1.553	2.878	0.773

These simplicity values are based on "natural" weights and therefore comparable across matrices.
When multiplied by the relative weights, they give the contribution to the overall simplicity value
(they are I^2/p, J^2/q or K^2/r, resp., times the sum of the variances of squared values)

relative weights			simplicity values			
A	B	C	core	A	B	C
0.00	0.00	0.00	5.29	1.55	2.88	0.77

Give SIMPLE STRUCTURE rotated solution(s) in detail

Use last simple structure rotated solution ? type "return"
to specify a different simple structure rotation, specify "1" **<return>**

Backnormalize A,B,C, and H

Rotated Solution for A, B and C is in AS, BT and CU
Rotated core in K

Rotated B (BT)

-0.21	-0.58
0.23	-0.55
-0.10	-0.36
0.02	-0.18
0.07	0.08
0.33	0.09
0.50	0.06
0.47	0.07
0.23	0.08
0.03	0.10
-0.08	0.10
-0.15	0.09
-0.19	0.09
-0.19	0.09
-0.19	0.08
-0.19	0.07
-0.18	0.07
-0.16	0.09
-0.13	0.13
-0.09	0.18
-0.02	0.22

Rotated C (CU)

0.43	0.35
-0.01	0.29
0.29	0.27
-0.07	0.33
-0.04	0.28
-0.22	0.40
0.29	0.27
-0.63	0.42
0.42	0.28
-0.16	0.22

Rotated core

B1xC1	B2xC1	B1xC2	B2xC2
-2.13	0.47	3.48	0.00
-0.29	-0.03	-0.01	16.71

You can keep present as final solution, or study a different one
If you want to study one more solution, specify "1" : **<return>**

You can now manually PERMUTE and REFLECT columns/rows of solution

If you want to reflect/permute columns/rows, specify "1": **<1>**

Give vector for reflection of columns of A (e.g., [1 -1 -1 1 ..]) : **<return>**

Give vector with new order of columns of A (e.g., [3 1 4 2 ..]): **< [2 1]>**

Give vector for reflection of columns of B (e.g., [1 -1 -1 1 ..]) **: <enter>**

Give vector with new order of columns of B (e.g., [3 1 4 2 ..]): **< [2 1]>**

If you want to further reflect/permute columns/rows, specify "1": **<enter>**

Rotated B (BT)

B 1	-0.58	-0.21
B 2	-0.55	0.23
B 3	-0.36	-0.10
B 4	-0.18	0.02
B 5	0.08	0.07
B 6	0.09	0.33
B 7	0.06	0.50
B 8	0.07	0.47
B 9	0.08	0.23
B10	0.10	0.03
B11	0.10	-0.08
B12	0.09	-0.15
B13	0.09	-0.19
B14	0.09	-0.19
B15	0.08	-0.19
B16	0.07	-0.19
B17	0.07	-0.18
B18	0.09	-0.16
B19	0.13	-0.13
B20	0.18	-0.09
B21	0.22	-0.02

Rotated C (CU)

C 1	0.35	0.43
C 2	0.29	-0.01
C 3	0.27	0.29
C 4	0.33	-0.07
C 5	0.28	-0.04
C 6	0.40	-0.22
C 7	0.27	0.29
C 8	0.42	-0.63

C 9	0.28	0.42
C10	0.22	-0.16

Rotated core

B1xC1	B2xC1	B1xC2	B2xC2
16.71	-0.01	-0.03	-0.29
0.00	3.48	0.47	-2.13

You can produce a JOINT PLOT now
In joint plots, one mode is fixed, while the entities of the other two are plotted
If you want to produce a joint plot for the current solution, specify "1": **<1>**
Which mode do you want to keep fixed ("1", "2" or "3")? **<3>**
For which component do you want to see the joint plot? **<1>**

The plot displays 100.00 % of the information represented by the component of your choice

Do you want to produce another joint plot? If so, type "1": **<1>**

Which mode do you want to keep fixed ("1", "2" or "3")? **<2>**
For which component do you want to see the joint plot? **<1>**

The plot displays 100.00 % of the information represented by the component of your choice

If you want to carry out a BOOTSTRAP procedure for
computing confidence intervals for the current solution, specify "1": **<enter>**

If you want to carry out a STABILITY CHECK on current or different solution, specify "1": **<enter>**
Press any key to conclude analysis with FITPARTITIONING
Contribution to fit (in %s) for all combinations of components
These contributions can simply be added to get aggregate contributions of in case all components are orthogonal
If this is not the case, aggregate fit contributions are given separately, below

B1xC1		B2xC1		B1xC2	B2xC2
79.90	0.00	0.00	0.02		
0.00	3.47	0.06	1.30		

Relative fit of B-mode entities (in %)

B 1	95.5
B 2	94.5
B 3	83.0
B 4	69.8
B 5	22.5
B 6	49.7
B 7	66.2
B 8	67.2
B 9	62.6
B10	64.4
B11	69.6
B12	69.7
B13	73.4
B14	73.0
B15	73.6
B16	71.8

B17	70.7
B18	72.5
B19	76.8
B20	82.2
B21	83.4

Relative fit of C-mode entities (in %)

C 1	87.8
C 2	85.0
C 3	67.6
C 4	83.4
C 5	81.3
C 6	92.8
C 7	81.2
C 8	91.7
C 9	78.9
C10	81.8

NOTE: Residuals on "Res", can be used for further analyses

References

Andersson, C. A., Bro, R., 2000., The N-way Toolbox for MATLAB, Chemometrics & Intelligent Laboratory Systems. 52 (1):1-4. http://www.models.life.ku.dk/source/nwaytoolbox/

Booth, H., Maindonald, J., Smith, S. 2002. Applying Lee-Carter under conditions of varying mortality decline. *Population Studies* 56: 325-336.

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

Carroll, J.D., 1968. Generalization of canonical correlation analysis to three or more sets of variables. *Proceedings of the 76th Annual Convention of the American Psychological Association* 3: 227-228.

Carter, L., Prkawetz, A., 2001. Examining structural shifts in mortality using the Lee-Carter method. Mpidr wp 2001-2007. Center for Demography and Ecology Information, University of Wisconsin-Madison.

Eckart, C., Young, G., 1936. The approximation of one matrix by another of lower rank. *Psychometrika* 1: 211-218.

Escofier, B., Pagès, J., 1990. Multiple factor analysis. *Computational Statistics and Data Analysis* 18: 121-140.

Gabriel, K. R., 1971. The biplot-graphic display of matrices with application to principal component analysis. *Biometrika* 58: 453-467.

Giordano, G., Russolillo, M., Haberman, S., 2008. Comparing Mortality Trends via Lee Carter Method in the Framework of Multidimensional Data Analysis. *Mathematical and Statistical Methods in Insurance and Finance*, 131-138. Edited by C. Perna & M. Sibillo. Milan: Springer.

Girosi, F., King G., 2005. A Reassessment of the LC Mortality Forecasting Method. Online at http://gking.harvard.edu/files/abs/lc-abs.shtml.

Guillen, M., Vidiella-i-Anguera, A., 2005. Forecasting Spanish natural life expectancy. *Risk Analysis* 25 (5): 1161-1170.

Hatzopoulos, P. Haberman, S., 2009. A parameterized approach to modelling and forecasting mortality. *Insurance: Mathematics and Economics* 44: 103-123

Hotelling, H., 1933. Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.* 24: 417-441, 498-520

Hyndman, R.J., Ullah, Md.S., 2007. Robust forecasting of mortality and fertility rates: a functional data approach. *Computational Statistics and Data Analysis* 51: 4942-4956.

Kiers, H.A.L., 2000. Displaying results from three-way methods. *Journal of Chemometrics* 14: 151-170.

Kiers, H.A.L., van Mechelen, I., 2001. Three-way component analysis: Principles and illustrative application. *Psychological Methods* 6: 84-110.

Kroonenberg, P.M., 1983. Three mode Principal Component Analysis. Theory and Applications. Leiden: DSWO Press.

Kroonenberg, P.M., de Leeuw, 1980. Principal Component Analysis of three mode data by means of alternating least squares algorithms. *Psychometrika* 45: 69-97.

Lee, R.D., Carter, L.R., 1992. Modelling and Forecasting U.S. Mortality. *Journal of the American Statistical Association* 87: 659-671.

Li, N., Lee, R.D., 2005. Coherent mortality forecasts for a group of populations: an extension of the Lee-Carter method. *Demography* 42: 575-594.

Renshaw, A., Haberman, S., 2003a. Lee-Carter mortality forecasting: a parallel generalized linear modelling approach for England and Wales mortality projections. *Applied Statistics* 52: 1, 119-137.

Renshaw, A., Haberman, S., 2003b. On the forecasting of mortality reduction factors. *Insurance: Mathematics and Economics* 32: 379-401.

Renshaw, A., Haberman, S., 2009. On age-period-cohort mortality rate parametric projections. *Insurance: Mathematics and Economics* 45: 255-270.

Tabeau, E.; van den Berg Jeths, A.; Heathcote, C. eds. 2001. Forecasting mortality in developed countries. Dordrecht: Kluwer Academic.

Tucker, L.R. 1966. Some mathematical notes on three-mode factor analysis. *Psychometrika 31:* 279–311.

Tukey, J. W., 1977. Exploratory Data Analysis. Addison-Wesley.

Wilmoth, J.R., Valkonen, T. 2001. A parametric representation of mortality differentials over age and time. Paper presented at EAPS Seminar, "Trends in Differentials in Morbidity and Mortality: Analysis and Explanation." Pontignano, Italy.