SUST Journal of Natural and Medical Sciences (JNMS) vol. 18 (2) 2017
ISSN (Print): 1858-6805 e-ISSN (Online): 1858-6813

عمــادة البحــث العلمــي
DEANSHIP OF SCIENTIFIC RESEARCH

SUST
Journal of Natural and Medical Sciences
Journal home page:
http://Scientific-journal.sustech.edu/

Using Lee-Carter Model to Fit and Forecast Age-Specific Mortality Rate for Colon Cancer

Arwa Elsiddig Elfaki Elhassan*[1] and Hamza Ibrahim Hamza[2]

1. Sudan University of Science and Technology , College of Science Email:arwaalfaki@gmail.com,
2. Alzaeem Alazhari University, College of Science
Corresponding Author: Email:Hamza75ib@gmail.com

Received: October 2017 Accepted: November 2017

Abstract:
This study was mainly focused on modeling and forecasting cancer mortality rates using standard Lee Carter Model, This model has become reference and leading statistical model for forecasting mortality. The model is applied to Colon cancer mortality data for Egyptian for both sexes ((male-female) aged from 5 to 74 years for the time period 2001-2014 and forecast mortality index by ARIMA Random Walk with Drift (0,1,0) from 2015 to 2020. The data obtained from World Health Organization (WHO).The model's parameters estimated by Singular Value Decomposition (SVD) and Maximum Likelihood Estimation (MLE), the comparison of the two methods (SVD, MLE) based on the mean error (ME) and mean square error (MSE). The results obtained by using different statistics packages like R and Iterative Lee Carter (ilc). The results showed MLE is better to estimate the parameters (ME=0.00506, MSE=0.11065) for male, while for female the SVD is better to estimate the parameters (ME=0.00835, MSE=0.07022). The error of forecasting age-specific mortality rate was (MPE=0.30682) for male and the error of forecast age-specific cancer mortality rate (MPE=0.04833) for female. The results showed the younger ages from 5 to 13, and from 30 to 39 had lower cancer mortality rate for male, and higher age-specific cancer mortality rate was in age-group (70-74) and it was 27.91 in year 2020 (per 100,000) for male .While for female had lower age-specific cancer mortality rate in younger ages from 5 to 24 and higher age-specific cancer mortality rate in old ages especially (60-64) and it was 21.55 in year 2020 (per 100,000). The study came out with a number of recommendation, the most important were to use Lee-Carter method to modeling and forecasting age-specific mortality rate , and to have care and accuracy when registering data of died people.

Keywords: ARIMA, WHO, SVD and MLE.

المستخلص:
في هذه الدراسة تم استخدام نموذج Lee-Carter للتنبؤ بمعدل الوفيات. والذي أصبح مرجعاً ورائد للنماذج الإحصائية للتنبؤ بمعدل الوفيات.وتم أستخدام النموذج على بيانات معدل الوفيات لسرطان القولون لسكان مصر (الذكور-الإناث) الذين تتراوح اعمارهم بين 5-74 في الفترة الزمنية 2001-2014 ومن ثم التنبؤ بدليل الوفاة بإستخدام نموذج المشي العشوائي بإنجراف (0,1,0). وكان مصدر هذة البيانات منظمة الصحة العالمية. وتم تقدير معالم النموذج بإستخدام طريقة القيمة المفردة وطريقة الإمكان الأعظم.وللمقارنة بين الطريقتين تم إستخدام متوسط الخطأ ومتوسط مربع الخطأ .وتم الحصول على النتائج بإستخدام مجموعة من الحزم الإحصائية أهمها R و ilc. أظهرت الدراسة ان النموذج قدر

المعالم بخطأ ومقداره (متوسط الخطأ=0.05377 و0.00835) للذكور والإناث على التوالي. والخطأ في التنبؤ بمعدلات الوفيات العمرية (متوسط الخطأ النسبي =,0.30682 و 0.04833) للذكور والإناث على التوالي.أن معدلات الوفيات العمرية منخفضة عند الأعمارالصغيرة بينما للأعمار الكبيرة مرتفعة بالنسبة للذكور لكل السنوات. وأعلى معدل وفاة عمرية لسرطان القولون كانت في الفئة العمرية (70−74) وكانت 27.91 في سنة 2020 (لكل 100.000)، بينما للإناث تكون منخفضة في الأعمار الصغيرة وتأخذ في الإرتفاع لبقية الأعمار . وأعلى معدل وفاة عمرية كانت 21.55 في سنة 2020 (لكل 100.000). خرجت الدراسة بعدد من التوصيات أهمها إستخدام طريقة Lee-Carter في التنبؤ بمعدل الوفيات العمرية . الإهتمام والدقة في تسجيل بيانات الموتى .

الكلمات المفتاحية :منظمة الصحة العالمية، طريقة القيمة المفردة وطريقة الإمكان الأعظم .

Introduction:

Mortality modeling has been used for long times, there were many models proposed since Gompertz published his law of mortality in 1825[1].The earliest models were simple and focused on producing mathematical functions to fit observed mortality rates. Nowadays, many methods have been proposed to model and forecast mortality rates, there is an extensive list of mortality forecasting models. Lee-Carter Model is one of the most popular methodologies for forecasting mortality rates. Lee and Carter applied their model on U.S. mortality data from the time period 1933-1987 and forecast were made up to the year 2065(1992)[2]. The model is widely known to be simple and has been used very successfully in U.S. and several countries[3]. The model is stochastic and combine demography model with statistical model time series[4] .

Colon cancer is a disease of large intestine which begins at a structure called the caecum, located in the right lower quadrant of the abdomen, and continues through all portions of the abdomen to its junction with the rectum, located in the deep pelvis. It arises from abnormal epithelial cells in the airway of colons[5]. In the Arab world show that (68.1%) of the Arab countries have colon cancer as one of the most frequent five types of cancer, they are gradually increasing in the region. Furthermore, there is great variation between different parts of the Arab world [6]. The (WHO) reported colon cancer deaths in Egypt reached 0.96% of total deaths[7].

Problem:

The cost of cancer care places a heavy burden on the health care system, and the long-term effects of cancer and its treatment on the quality of life of survivors take a toll at a population level and economics.

There are a few of studies applying statistical models to forecast the cancer mortality rate in Arab Countries, almost of the studies on incidence of cancer were conducted by the doctors or who in the field of health.

Objectives:

- We want to know about standard Lee-Carter method to model and forecast age-specific mortality rates of cancer (oral, lung and colon) for both sexes.
- How to use Singular Value Decomposition (SVD) and maximum likelihood estimation (MLE) to estimate the model's parameters.
- To determine the best method to use for estimation and forecasting the parameters' model for both sexes (male, female).
- To offers many important opportunities to improve public health from forecast of age-specific cancer mortality rate.

Importance:

To cover the lack of previous research on the age-specific cancer mortality rate, to apply new statistical models (Lee-Carter model) in our Arab countries. Cancer mortality rate can provide important information that influences practices, policies, and programs that directly affect the health sector.

Hypothesis:

- If the Lee-Carter model fit well for forecasting age-specific cancer mortality rate for colon cancer .
- If a Singular Value Decomposition (SVD) fit best than Maximum likelihood Estimation (MLE) to estimate the model's parameters for male.
- If a Singular Value Decomposition (SVD) fit best than Maximum likelihood Estimation (MLE) to estimate the model's parameters for female.

Methodology:

In this study we used articles, studies, books and previous studies. The statistical package R, Iterative lee carter package (ilc), demography and forecast packages has been used to execute Singular Value Decomposition (SVD) and Maximum Likelihood Estimation (MLE) to estimate the Lee-Carter model and forecasting age-specific cancer mortality rate.

Material and method :

Data source:

For this study the source of the data from the World Health Organization (WHO) Mortality database[8], , which contains number of deaths and population by country, year, sex, age-group and cause of death. For Egyptian colon cancer (male-female) from the period 2001-2014. The mortality rate of colon cancer was computed for age by dividing the number of death by the corresponding number of population for every age group and multiply by 100,000. The age groups are (5-9, 10-14,..., 70-74).

Lee-Carter Model:

The model is stochastic model to forecast the mortality rate[9]. Lee and Carter applied their model on U.S. mortality data from the time period 1933-1987 and forecast were made up to the year 2065. The model involves the main structure below:

$$lnm_{x,t} = a_x + b_x k_t + \epsilon_{x,t} \quad ....(1)$$

where $m_{x,t}$ is central death rate at age x in year t, a_x describes the overall mortality pattern across age, b_x represents the sensitivity of the log death rate to changes in the mortality index k_t, and $\epsilon_{x,t}$ is the error term with mean zero and variance σ_ϵ^2, they ignored it because it was too small. The a_x parameters are first estimation by averaging $lnm_{x,t}$ over time t. The b_x and k_t parameters were computed by applying singular value decomposition (SVD) to the matrix with components $lnm_{x,t} - \hat{a}_x$, with two constraints $\sum_x b_x$=1 and $\sum_t k_t$=0 . Finally, the k_t parameter was re-estimated in such a way that the fitted number of deaths and the actual number of deaths are to be equal for each year.

$$D_t = \sum_x N_{x,t}\, e^{\hat{a}_x + \hat{b}_x \hat{k}_t} \quad ....(2)$$

Where D_t is the total of deaths in year t. $N_{x,t}$ the population of age x in year t. A random walk with drift would be appropriate for modeling the series where d is the drift term and e’s are independent and identically distributed (iid) error terms with mean zero and variance.

In this study we applied two methods; Singular Value Decomposition (SVD) , and Maximum Likelihood Estimation (MLE) which was introduced by Wilmoth (1993) to estimate the model's parameters. Comparison of the two methods were based on mean error (ME) and mean square error (MSE), and mean percentage error MPE for forecasting[10].

$$\mathrm{ME} = \frac{1}{A}\sum_{x,t}\hat{\epsilon}_{x,t} \quad \ldots\ldots(3)$$

$$\mathrm{MSE} = \frac{1}{A}\sum_{x,t}\hat{\epsilon}_{x,t}{}^{2} \ldots\ldots(4)$$

$$\mathrm{MPE} = \frac{1}{A}\sum_{x,t}\frac{\hat{\epsilon}_{x,t}}{m_{x,t}} \ldots\ldots(5)$$

1 Singular Value Decomposition (SVD):

Singular Value Decomposition (SVD) was presented by Bozik and Bell (1987)and Bell and Monsell (1991) [11] . Is a method for transforming correlated variables into a set of uncorrelated ones that better expose the various relationships among the original data items. It is based on a theorem from linear algebra which says that a matrix Z can be broken down into the product of three matrices - an orthogonal matrix U, a diagonal matrix L, and the transpose of an orthogonal matrix V .

$A_{m,n}$ can be decomposed uniquely as.

First obtain the $a_x = \frac{1}{n}\sum_t \ln m_{x,t}$ ......(6)

Create matrix $Z_{x,t}$ for estimating b_x and k_t

Where $Z_{x,t} = \ln m_{x,t} - \widehat{a_x}$ ......(7)

Apply singular value decomposition to $Z_{x,t}$ to decompose the matrix $Z_{x,t}$ into the product three matrix

$$\mathrm{SVD}\,(Z_{x,t}) = ULV \quad \ldots\ldots(8)$$

U: represent the age component.

L: represent the singular values.

V: represent the time component.

$\widehat{b_x}$ is derived from the first vector of U. And $\widehat{k_t}$ is derived from the first vector of V.

$$\widehat{b_x} = \frac{1}{\sum_x u_{x.i}^2}.\,(u_{1,1} \quad u_{2,1} \quad \ldots \quad u_{x,1}) \ldots\ldots(9)$$

$$\widehat{k_t} = \sum_t u_{x,1}^2\,.\,l_1.\,(v_{1,1} \quad v_{2,1} \quad \ldots \quad v_{t,1}) \ldots\ldots(10)$$

2. Maximum Likelihood Estimation (MLE):

The MLE referred as Poisson log bilinear model[12]. It gives optimal solution of the LC model under a Poisson model and avoid assumption of error with constant variance. This was introduced by (Wilmoth, 1993) [13].

$$D_{x,t} \sim P(m_{x,t}, E_{x,t}) \ldots\ldots(11)$$

Where

$$m_{x,t} = \exp(a_x + b_x k_t) \ldots\ldots(12)$$

Then the MLE is given by:

$$l(a,b,k) = \ln\left[\prod_x\prod_t \frac{\hat{D}_{x,t}^{D_{x,t}}\exp(-\hat{D}_{x,t})}{D_{x,t}!}\right] \ldots\ldots(13)$$

$$= \sum_x\sum_t (D_{x,t}\ln\hat{D}_{x,t} - \hat{D}_{x,t}\ln D_{x,t}!)$$

$$\sum_x\sum_t (D_{x,t}\left(a_x + b_x\hat{k}_t\right) - E_{x,t}\exp(a_x + b_x k_t)\mathrm{constant}) \ldots\ldots(14)$$

By differentiating both sides of equation (14), we can immediately see that the observed and fitted number deaths overtime are equal when the algorithm converges.

Random Walk with Drift:
In this study we used Auto Regressive Integrated Moving Average ARIMA models Random Walk with Drift (0,1,0) as standard Lee-Carter model to derived $k_t$values. Many techniques have been used and specified especially Random Walk with Drift because it is a simplest and important model. It presented as a current observation equal to previous observation with a random step up or down[14]. After k_t index is obtained it possible to forecast cancer mortality rate.

$$\hat{k}_t = \hat{k}_{t-1} + d + \epsilon_t \quad ....(15)$$

Where d: is drift parameter. ϵ_t is an error term with zero mean and constant variance.

$$\hat{d} = \frac{\hat{k}_t - \hat{k}_1}{T-1} \quad ....(16)$$

It depend on the first and last of k_t estimation.
The drift d estimated with uncertainty and standard error of it estimated and it used to form more complete measure of uncertainty in forecasting k_t.

$$\text{See} = \sqrt{\frac{1}{T-1}\sum_t(\hat{k}_t - \hat{k}_{t-1} - \hat{d})^2}................(17)$$

The age-specific cancer mortality rate was forecasted and it is computed from the following equations:

$$m_{x,2014+\Delta t} \approx e^{\hat{a}_x + \hat{b}_x \hat{k}_{2014+\Delta t} \cdot \theta}................................(18)$$

Results and Discussion:
The parameters of the Lee-Carter model were estimated by using the two methods (SVD, MLE), which were represented in the section above.

Singular Value Decomposition (SVD):

The parameter a_x is computed first from equation (6).

Table (1) Estimation of a_x.

Age	Male	Female
5-9	-3.53290945	-0.62931252
10-14	-3.10000329	-2.02133093
15-19	-1.81588120	-1.31984723
20-24	-0.98485968	-0.59637149
25-29	-0.61660929	-0.05982704
30-34	-0.28417757	0.14262063
35-39	0.01832391	0.52988338
40-44	0.33843088	0.87169901
45-49	0.81913661	1.14663662
50-54	1.34541791	1.40425981
55-59	1.75761235	1.56956913
60-64	1.98511767	1.75972606
65-69	2.16650715	1.69321958
70-74	2.53548953	1.68101701

Source: Author calculation by ilc and demography and R.

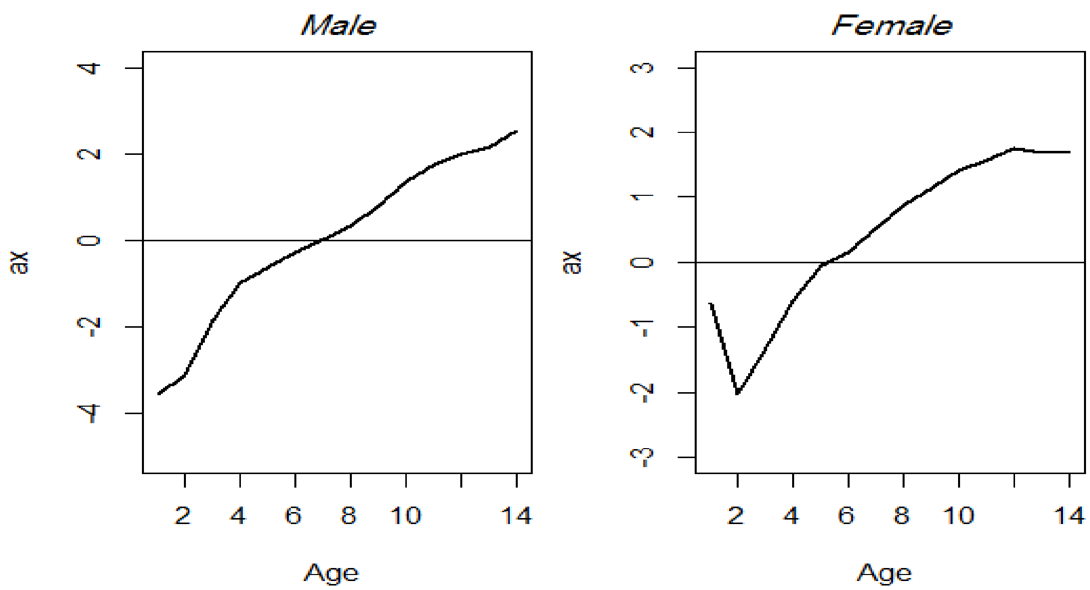


Figure (1) General pattern of mortality *ax*

Source: Author plotting by ilc and demography and R.

Table (1) shows the values of a_x, which it represents the general pattern (age shape) of mortality by age x for both sexes (male-female), and it increased overtime for both sexes (male, female). Figure (1) shows the pattern of a_x and it was up trend for both sexes (male, female), and this indicate that the younger ages had lower mortality than older ages. The negative trend in $\hat{a}_x$is in accord with improvement in cancer mortality rate.

The parameter b_x is computed from equation (9).

Table (2) Estimation of b_x.

Age	Male	Female
5-9	-0.659039790	-0.15717240
10-14	0.227754073	0.10033237
15-19	0.295034154	0.02465728
20-24	1.122244900	0.01449897
25-29	0.363905258	0.08518087
30-34	0.003015176	0.08635695
35-39	-0.088629174	0.02872251
40-44	0.068362719	0.09662582
45-49	-0.009866323	0.09911957
50-54	0.029212021	0.08127946
55-59	-0.064217863	0.13047972
60-64	-0.238798260	0.18146069
65-69	-0.156492757 0.107515	0.13262701
70-74	866	0.09583120

Source: Author calculation by ilc and demography and R.

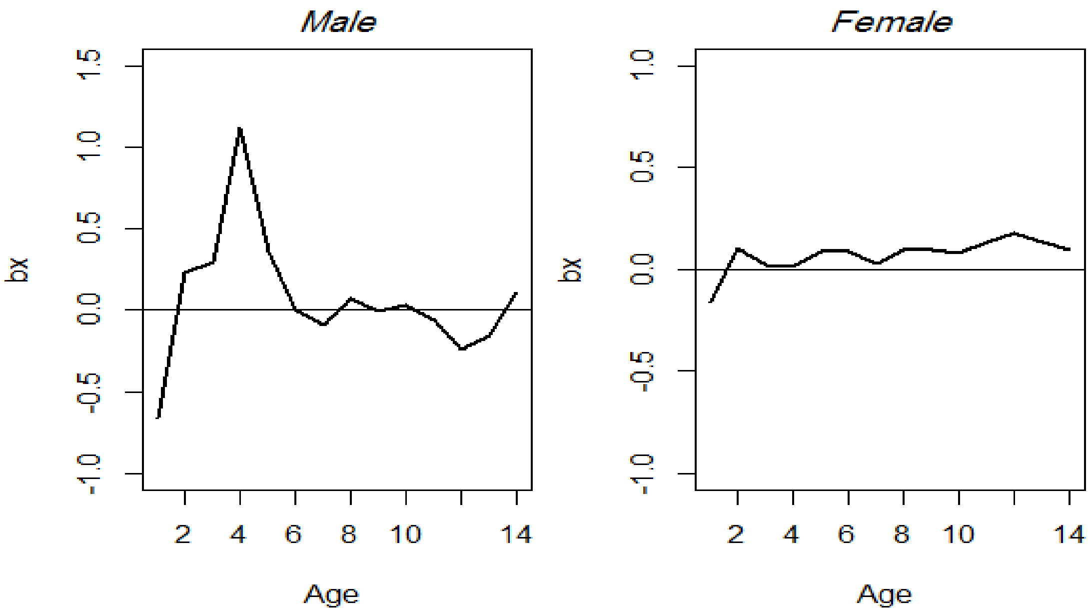


Figure (2) General pattern of b_x .

Source: Author plotting by ilc and demography and R.

Table (2) shows the values of $\hat{b}_x$ which represent the tendency of mortality at age x to change as the general level of mortality changes. The figure (2) shows the cancer mortality change for younger ages for male, and the cancer mortality among younger ages had higher values. For female the mortality the values of $\hat{b}_x$ was invariant for all ages. The high values of $\hat{b}_x$ indicate improvement in mortality at these ages, while the negative values at some ages indicate that mortality was increasing.

The First estimation of the parameter k_t, for 2001-2014 from the equation (10) and re-estimation of k_t from equation (2).

Table (3) Estimation of adjusted k_t .

Year	Male		Female	
	1st estimation	**2nd estimation**	**1st estimation**	**2nd estimation**
2001	1.146942967	0.74160675	-3.701077352	-3.51165155
2002	-0.406075192	-0.14210725	-2.263572299	-2.69803015
2003	-0.325588844	-0.14228783	-0.520527975	-1.66075716
2004	0.798716265	0.95131117	-1.644621178	-1.53820088
2005	0.870350407	1.74953844	-2.204503624	-0.85398142
2006	0.339109120	0.41057080	-0.814619651	-1.10909621
2007	0.004599658	1.06728620	-0.751976193	-0.08371683
2008	-0.134249018	-0.09566960	0.001437195	0.14081463
2009	-0.182098401	-0.09645981	-0.924889563	-0.31481464
2010	-0.032816015	0.04873011	1.267529793	0.75565107
2011	-0.411749723	1.46094517	3.068983558	2.53727672
2012	-0.580389202	1.40399144	2.551098646	3.06170198
2013	-0.462729122	1.26948739	2.340694394	3.07857939
2014	-0.624022900	1.91101848	3.596044251	3.83269770

Source: Author calculation by ilc and demography and R.

Figure (3) Estimation of *kt* , 2001–2014

Source: Author plotting by ilc and demography and R.

Table (3) shows the values of mortality index k_t for the period 2001–2014 for both sexes (male-female), which it captures the main time trend on the logarithmic scale in death rates at all ages. For the first and second estimation of k_t we solving equation (7) and (2). Figure (3) shows the mortality index $\hat{k}_t$ had non-linear trend overtime for male and linear trend for female. The decrease values of $\hat{k}_t$ indicate the mortality trend is decline.

Maximum Likelihood Estimation (MLE):

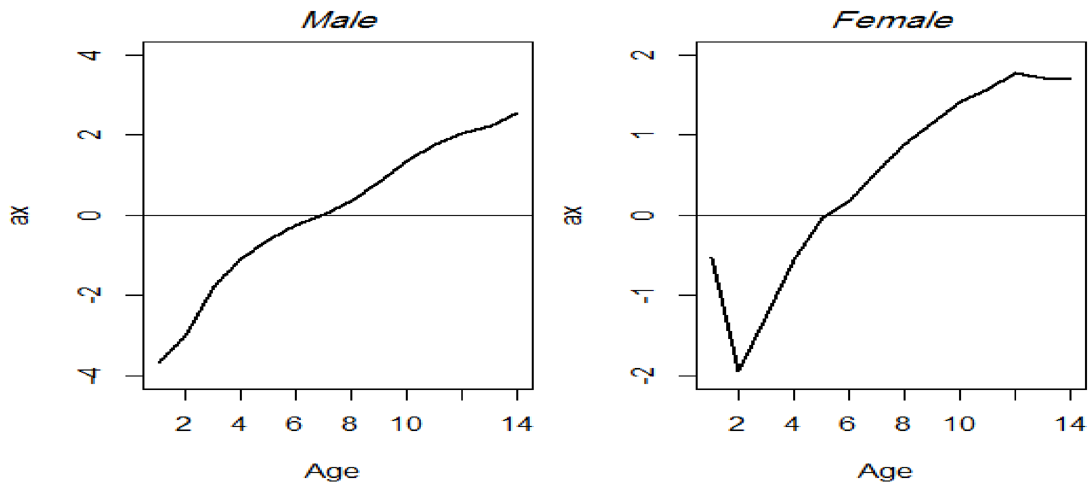
After fitting the technique of (MLE), we obtained at the following results.

We obtained the value of parameter a_x from equation (14).

Table (4) Estimation of a_x .

Age	Male	Female
5-9	-3.66637596	-0.52340236
10-14	-2.99544866	-1.93200492
15-19	-1.76819721	-1.23139904
20-24	-1.09702779	-0.55174437
25-29	-0.61624948	-0.03439554
30-34	-0.25596000	0.18268603
35-39	0.02772012	0.54980894
40-44	0.34389374	0.88246825
45-49	0.83107735	1.15147602
50-54	1.34510871	1.41186079
55-59	1.77697451	1.57879206
60-64	2.02234346	1.76559040
65-69	2.21900972	1.70106493
70-74	2.53552174	1.68967701

Source: Author calculation ilc and demography and R.



Figure(4) General pattern of ax.

Source: Author plotting by ilc and demography and R.

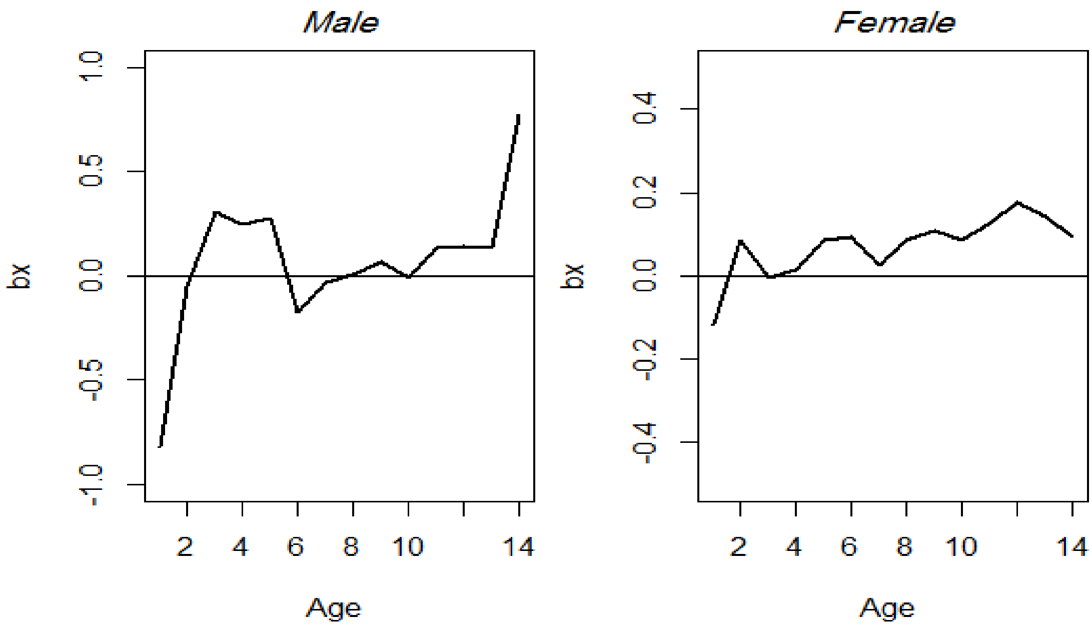
Table(4) shows the values of a_x ,which it represents the general pattern (age shape) of mortality by age for both sexes (male-female), and Figure(4) shows the pattern of a_x and shows the values of $\hat{a}_x$ increase over time for both sexes (male, female), and this indicate that they have up trend in mortality and the younger ages have lower mortality than older ages. The negative trend in $\hat{a}_x$is in accord with improvement in cancer mortality rate .

We obtained the value of parameter $b_x$from equation (14).

Table (5) Estimation of b_x .

Age	Male	Female
5-9	-0.821667657	-0.115667653 0.085
10-14	-0.051230827	441869
15-19	0.305852583	-0.005966264
20-24	0.247693920	0.015407662
25-29	0.278103135	0.083714243
30-34	-0.177637626	0.093985550
35-39	-0.031097756	0.024643654
40-44	0.007978901	0.085266973
45-49	0.068626016	0.109630276
50-54	-0.007935564	0.085783964
55-59	0.131536718	0.125462580
60-64	0.139910407	0.175256528
65-69	0.136732105	0.143468938
70-74	0.773135645	0.093571681

Source: Author calculation ilc and demography and R.



Figure(5) General pattern of (b_x)

Source: Author plotting by ilc and demography and R.

Table(5) shows the values of $\hat{b}_x$ which it represent the tendency of mortality at age x to change as the general level of mortality changes. the figure (5) shows the $\hat{b}_x$ has a negative value for younger ages for male, and positive value for older ages while the middle ages have invariant $\hat{b}_x$ for female the mortality the values of $\hat{b}_x$ was invariant for younger and middle ages and the older ages has higher values . The high values of $\hat{b}_x$ indicate improvement in mortality at all ages , while the negative values at some ages indicate that mortality was increasing.

We obtained the value of parameter kt from equation (14).

Table (6) Estimation of kt , 2001–2014 .

Year	Male	Female
2001	-0.928203921	-3.6784657
2002	-1.027160713	-2.6000211
2003	-0.594061609	-1.6982635
2004	0.654405599	-1.7673720
2005	1.161618504	-1.1815349
2006	-0.041611141	-1.2622395
2007	0.039622075	-0.2619323
2008	-0.218416341	0.2460449
2009	0.007840427	-0.6374291
2010	0.172756241	0.8148711
2011	0.134331804	2.4173220
2012	0.217326561	3.0160890
2013	0.012394591	2.9251521
2014	0.409157921	3.6677791

Source: Author calculation ilc and demography, forecast and R.

Figure(6) Estimation of adjusted k_t for (2001-2014)
Source: Author plotting by ilc and demography and R

Table (6) shows the values of mortality index k_t for the period 2001–2014 for both sexes (male-female) ,which it captures the main time trend on the logarithmic scale in death rates at all ages. Figure(6) shows the mortality index $\hat{k}_t$ has nonlinear trend for male and linear trend for female. The high values of $\hat{k}_t$ indicate there is no improvement of cancer mortality rate. for male 2005 has higher mortality rate, while for female have higher mortality on 2010-2014.

Comparison between Singular Value Decomposition(SVD) and Maximum Likelihood Estimation(MLE):

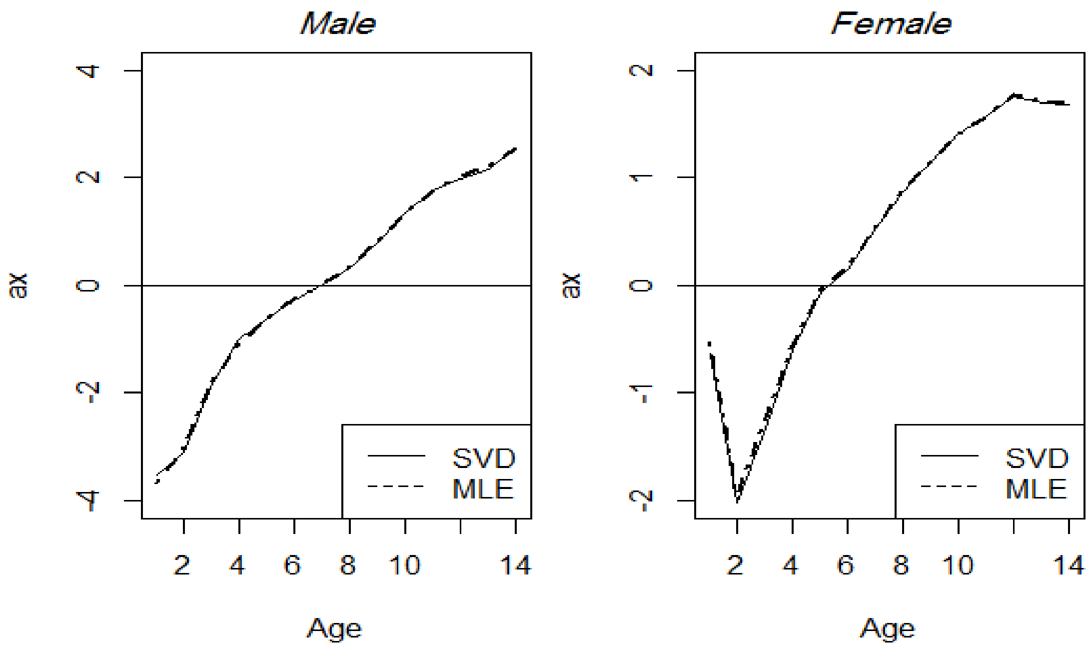
The comparison between two methods (SVD, MLE) to estimate the model's parameters based on ME and MSE, which were obtained from equations (3) and (4) respectively for both sexes (male, female).

Table (7) Comparison between SVD and MLE for estimation of a_x .

Age	Male		Female	
	SVD	**MLE**	**SVD**	**MLE**
5-9	-3.53290945	-3.66637596	-0.62931252	-0.52340236
10-14	-3.10000329	-2.99544866	-2.02133093	-1.93200492
15-19	-1.81588120	-1.76819721	-1.31984723	-1.23139904
20-24	-0.98485968	-1.09702779	-0.59637149	-0.55174437
25-29	-0.61660929	-0.61624948	-0.05982704	-0.03439554
30-34	-0.28417757	-0.25596000	0.14262063	0.18268603
35-39	0.01832391	0.02772012	0.52988338	0.54980894
40-44	0.33843088	0.34389374	0.87169901	0.88246825
45-49	0.81913661	0.83107735	1.14663662	1.15147602
50-54	1.34541791	1.34510871	1.40425981	1.41186079
55-59	1.75761235	1.77697451	1.56956913	1.57879206
60-64	1.98511767	2.02234346	1.75972606	1.76559040
65-69	2.16650715	2.21900972	1.69321958	1.70106493
70-74	2.53548953	2.53552174	1.68101701	1.68967701

Source: Author calculation ilc and demography, forecast and R.

SUST Journal of Natural and Medical Sciences (JNMS) vol. 18 (2) 2017
ISSN (Print): 1858-6805 e-ISSN (Online): 1858-6813



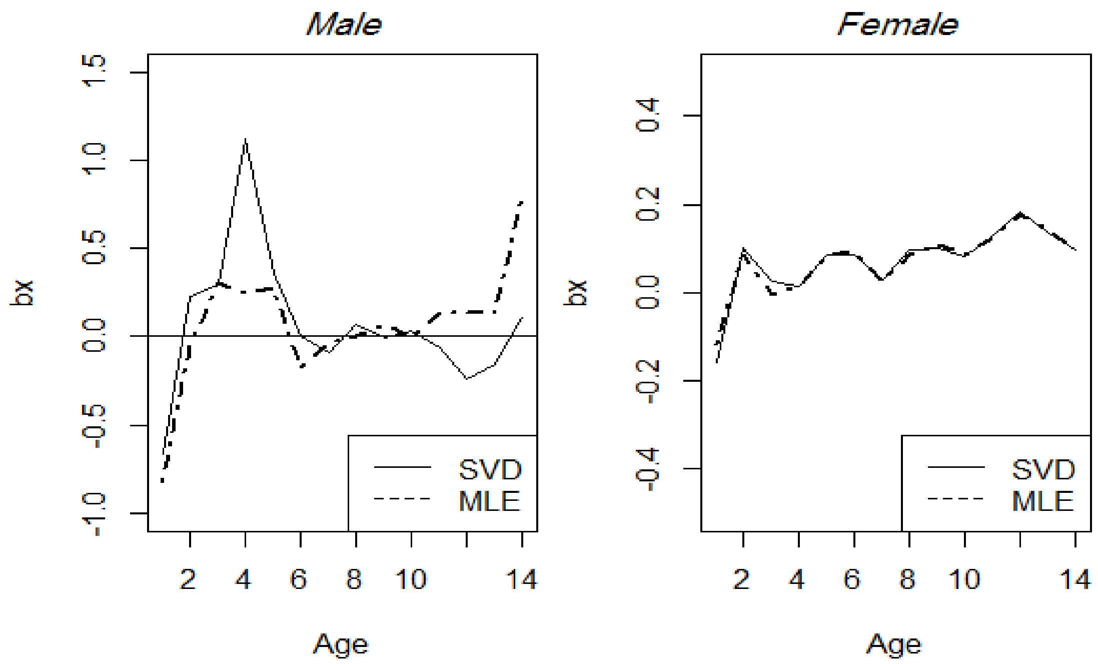
Figure(7) Comparison between SVD and MLE for estimation a_x.
Source: Author plotting by ilc and demography and R.

If we take a look to table(7) and figure(7), we will notice that the estimation of parame ters a_x from SVD and MLE slight differenc e , and this is very clear in the figure (7) for both sexes (male, female). The maximum di fference value of estimation a_x for male wa s 0.1334665, while for female was 0.10591 02.

Table (8) Comparison between SVD and MLE for estimation of b_x.

Age	Male		Female	
	SVD	**MLE**	**SVD**	**MLE**
5-9	-0.659039790	-0.821667657	-0.15717240	-0.115667653
10-14	0.227754073	-0.051230827	0.10033237	0.085441869
15-19	0.295034154	0.305852583	0.02465728	-0.005966264
20-24	1.122244900	0.247693920	0.01449897	0.015407662
25-29	0.363905258	0.278103135	0.08518087	0.083714243
30-34	0.003015176	-0.177637626	0.08635695	0.093985550
35-39	-0.088629174	-0.031097756	0.02872251	0.024643654
40-44	0.068362719	0.007978901	0.09662582	0.085266973
45-49	-0.009866323	0.068626016	0.09911957	0.109630276
50-54	0.029212021	-0.007935564	0.08127946	0.085783964
55-59	-0.064217863	0.131536718	0.13047972	0.125462580
60-64	-0.238798260	0.139910407	0.18146069	0.175256528
65-69	-0.156492757	0.136732105	0.13262701	0.143468938
70-74	0.107515866	0.773135645	0.09583120	0.093571681

Source: Author calculation ilc and demography, forecast and R.



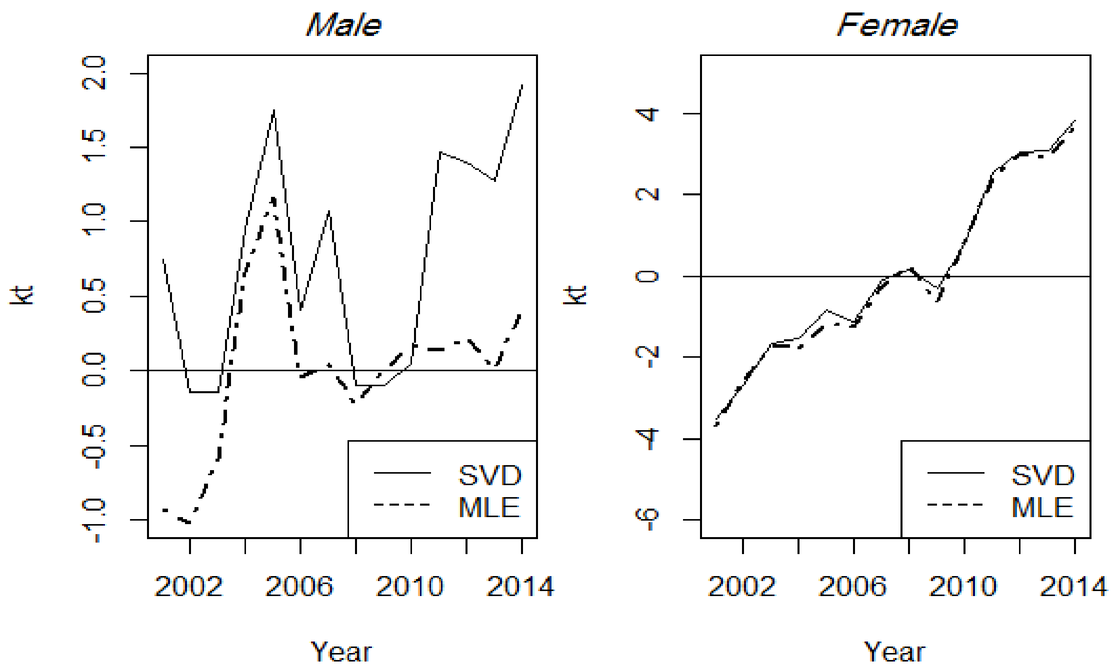
Figure(8) Comparison between SVD and MLE for estimation b_x.
Source: Author plotting by ilc and demography and R.

If we take a look to table(8) and figure(8), we will notice that the estimation of parame ters b_x from SVD and MLE was difference for male, while for female was slight differe nce, and this was very clear in the figure (8) for both sexes (male, female). The maximu m difference value of estimation b_x for male was 0.874551, while for female was 0.041 50475.

Table(9) Comparison between SVD and MLE for estimation of k_t.

Year	Male		Female	
	SVD	**MLE**	**SVD**	**MLE**
2001	0.74160675	-0.928203921	-3.51165155	-3.6784657
2002	-0.14210725	-1.027160713	-2.69803015	-2.6000211
2003	-0.14228783	-0.594061609	-1.66075716	-1.6982635
2004	0.95131117	0.654405599	-1.53820088	-1.7673720
2005	1.74953844	1.161618504	-0.85398142	-1.1815349
2006	0.41057080	-0.041611141	-1.10909621	-1.2622395
2007	1.06728620	0.039622075	-0.08371683	-0.2619323
2008	-0.09566960	-0.218416341	0.14081463	0.2460449
2009	-0.09645981	0.007840427	-0.31481464	-0.6374291
2010	0.04873011	0.172756241	0.75565107	0.8148711
2011	1.46094517	0.134331804	2.53727672	2.4173220
2012	1.40399144	0.217326561	3.06170198	3.0160890
2013	1.26948739	0.012394591	3.07857939	2.9251521
2014	1.91101848	0.409157921	3.83269770	3.6677791

Source: Author calculation ilc and demography and R.



Figure(9) Comparison between SVD and MLE for estimation k_t.
Source: Author plotting by ilc and demography and R.

If we take a look to table(9) and figure(9), we will notice that the estimation of parame ters k_t. from SVD and MLE was difference for male, while for female was slight differe nce, and this was very clear in the figure (9) for both sexes (male, female). The maximu m difference value of estimation k_t for mal e was 1.669811, while for female was 0.327 5535.

Table(10) Comparison between SVD and MLE for Errors .

Sex	Method	ME	MSE
Male	**SVD**	0.05377	0.28382
	MLE	**0.00506**	**0.11065**
Female	**SVD**	**0.00835**	**0.07022**
	MLE	0.03347	0.07085

Source: Author calculation by ilc and demography and R
Note: The numbers in bold are best .

The table(10) shows the errors from the tow methods (SVD,MLE), and they were satisfactory well for estimating the parameters, but MLE was better than SVD for male with errors (ME=0.00506, MSE=0.11065). While for female The SVD better Than SVD with errors (ME=0.00835, MSE=0.07022) .

Forecast k_t and Age-specific Cancer Mortality Rate:

The ARIMA (0, 1, 0) Walk Random with drift was used to forecast the mortality index for both sexes (male, female) from equations (15), For male we used k_t from The MLE, while for female we used k_t from SVD. The drift is calculated from equation(16) and error from equation (17).

Table (11) Estimation of drift and Error of ARIMA model.

Sex	**Male**	**Female**
Method	**MLE**	**SVD**
$\hat{d}$	0.102874	0.5649499
$\hat{\sigma}^2$	0.3069034	0.3842088

Source: Author calculation by ilc, demography, forecast and R.

Table (12) Forecast k_t for period 2015 – 2020 for male.

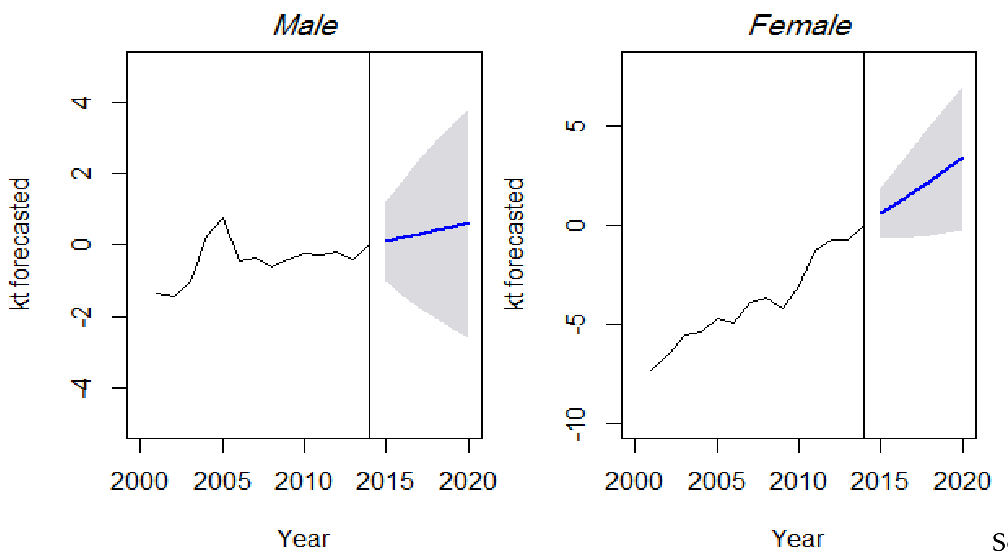
Year	**$k_t$forecast**	**lower**	**Upper**
2015	0.1028740	-1.023911	1.229659
2016	0.2057480	-1.443697	1.855193
2017	0.3086220	-1.777779	2.395023
2018	0.4114960	-2.071818	2.894810
2019	0.5143699	-2.342552	3.371292
2020	0.6172439	-2.598116	3.832604

Source: Author calculation by ilc and demography and R.

Table (13) Forecast k_t for period 2015 – 2020 for female.

Year	**$k_t$forecast**	**lower**	**Upper**
2015 2016 2017 2018 2019 2020	0.5649499 1.1298 999 1.6948498 2. 2597998 2.82474 97 3.3896997	-0.6957835 -0.7 156249 -0.6395 749 -0.5187214 -0.3717930 -0.2 078905	1.825683 2.975 425 4.029275 5. 038321 6.02129 2 6.987290

Source: Author calculation by ilc and demography and R.



Source: Author

Figure (10) Forecast mortality index for (2000-2020).

plotting by ilc and demography and R.

Table (11) shows the estimated drift and err or form Random walk with drift.Tables (12) and (13) and figure (10) show that the k_t in creased overtime for both sexes (male, fema le).

After obtaining the cancer mortality index, it was used to forecast age-specific cancer mortality rate for the period 2015-2020 for both sexes. It was computed from the equation:

Table (14) Error from forecasting age-specific mortality rate.

Sex	**MPE**
Male	0.30682
Female	0.04833

Source: Author calculation by ilc and demography, forecast and R.

Table (14) shows the errors from forecasting age-specific cancer mortality rate for both sexes (male, female), and were small especially for female.

Table (15) Forecast age-specific cancer mortality rate for the period 2015–2020 for male.

Age	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**
5-9	0.02	0.02	0.01	0.01	0.01	0.01
10-14	0.05	0.05	0.05	0.05	0.05	0.05
15-19	0.20	0.21 0.	0.21	0.22	0.23	0.23
20-24	0.38	39	0.40	0.41	0.42	0.43
25-29	0.62	0.64	0.66	0.68	0.70	0.72
30-34	0.71	0.69	0.68	0.67	0.66	0.65
35-39	1.01	1.01	1.01	1.00	1.00 1	1.00
40-44	1.42	1.42	1.42	1.42	.42	1.422
45-49	2.38	2.39	2.41	2.43	2.45	2.46
50-54	3.82	3.82	3.82	3.81	3.81	3.81
55-59	6.32	6.41	6.50	6.59	6.68	6.77
60-64	8.12	8.23	8.35	8.47	8.60	8.72 10.5
65-69	9.87	10.01	10.15	10.29	10.44	8
70-74	18.75	20.31	21.99	23.81	25.78	27.91

Source: Author calculation by ilc and demography, forecasting and R.

Table (16) Forecasting age-specific cancer mortality rate for period 2015–2020 for female.

Age	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**
5-9	0.27	0.24	0.22	0.20	0.19	0.17
10-14	0.21	0.22	0.23	0.24	0.26	0.27
15-19	0.30	0.30	0.31	0.31	0.31	0.32
20-24	0.59	0.59	0.60	0.60	0.61	0.61
25-29	1.37	1.44	1.51	1.58	1.66	1.74
30-34	1.69	1.77	1.86	1.95	2.05	2.15
35-39	1.93	1.96	1.99	2.02	2.06	2.09
40-44	3.66	3.86	4.08	4.31	4.55	4.80
45-49	4.87	5.15	5.44	5.76	6.09	6.44
50-54	5.82	6.10	6.38	6.68	7.00	7.32
55-59	8.53	9.18	9.88	10.64	11.45	12.33
60-64	12.91	14.30	15.84	17.55	19.45	21.55
65-69	9.74	10.50	11.32	12.20	13.15	14.17
70-74	8.19	8.64	9.12	9.63	10.17	10.73

Source: Author calculation by ilc and demography, forecasting and R.

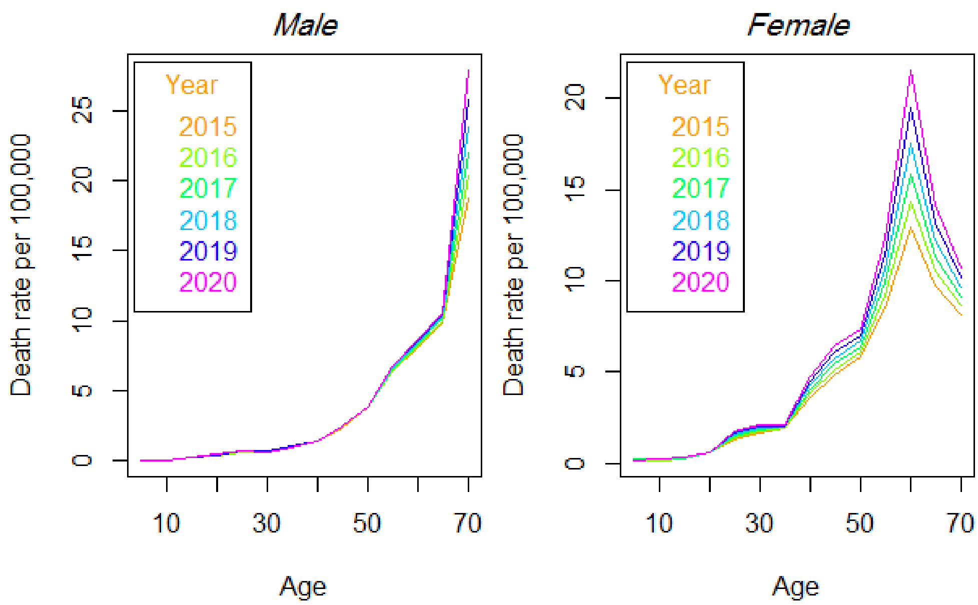


Figure (11) Forecast age-specific mortality rate (2015-2020).
Source: Author plotting by ilc and demography, forecast and R.

Tables(15) and (16) and figure(11) show the age-specific cancer mortality rates were increasing for aged (5-24), (40-59) and (65-74) year-old, while for female age-specific cancer mortality rates were increasing for aged 15-64 to except age (65-74) year-old. When comparing both sexes the male had lower cancer mortality rate than female overtime.

Conclusions:

- The two methods (SVD,MLE) were satisfactory with (ME =0.05377,0.00506) and (MSE=0.28382,0.11065) respectively for male. While for female (ME=0.00835,0.03347) and (MSE=0.07022,0.07085) respectively.
- MLE is better than SVD to estimate the Lee-Carter model for male with error (ME=0.00506, MSE=0.11065), while for female SVD is better than MLE with error (ME=0.00835, MSE=0.07022). The errors of forecasting age-specific cancer mortality rate MPE=0.30682 for male, while for female MPE= 0.04833**.**
- The higher age-specific cancer mortality rate for male found in age-group (70-74) for all years (2015,2016,2017,2018,2019 and 2020) was (18.75,20.31,21.99,23.81,25.78 and 27.91).
- The higher age-specific cancer mortality rate for female found in age-group (60-64) for all years (2015, 2016, 2017, 2018, 2019 and 2020) was (12.91,14.30, 5.84, 17.55, 19.45 and 21.55)

Recommendations:

- Recommend to use Lee-Carter model to forecast mortality rate, because its simplicity and provides a description of mortality change that is easy to understand.
- The model has a few variables and combine demographic and statistical models other than the mortality models.

- Recommended to record data accurately because incomplete data affect the performance of the model.

References:

1. Benjamin Gompertz. (2010). On the Nature of the Function Expressive of the Law of Human Mortality, and on a New Mode of Determining the Value of Life Contingencies. *Philosophical Transactions of the Royal Society of London*. **115** : 513-583 **.**

2. Lee, Ronald D. and Carter, Lawrence R. (1992). Modeling and Forecasting U.S. Mortality. *Journal of the American Statistical Association.* **87**:659-671.

3. Jenny Zheng Wang. (2007). Fitting and Forecasting Mortality for Sweden: Applying the Lee-Carter Model. Mathematical Statistics Stockholm University. (http://www.math.su.se/mathstat).

4. Tabeau, E., Van Den Berg Jeths, A. and Heathcote, C. (2001). *Forecasting Mortality in Developed Countries: Insights from a Statistical Demographic and Epidemiological Perspective*, Kluwer Academic Publishers.

5. National Comprehensive Cancer Network Staff (NCCN). (2014). *Colon Cancer*.**1.2014**. National Comprehensive Cancer Network (NCCN).

6. Amal S. Ibrahim, Hussein M. Khaled, Nabiel NH Mikhail, Hoda Baraka, and Hossam Kamel. (2014). *Cancer Incidence in Egypt: Results of the National Population-Based Cancer Registry Program* Journal of Cancer Epidemiology **2014** .

7. American Cancer Society. (2010) .The global economic cost of cancer, No.005444.

8. World Health Organization Mortality database. http://www.who.int/healthinfo/statistics/mortality_rawdata/en. (downloaded 11/5/2017)

9. Rajendra N. Chavhan and Ramkrishna L. Shinde. (2016). Modeling and Forecasting Mortality Using the Lee-Carter Model for Indian Population Based on Decade-wise Data *.Sri Lankan Journal of Applied Statistics.* 17-1:51-68.

10. Wouter van Wel . (2015). Mortality Modelling and Forecasting using Cross-Validation Techniques, *Marble* **1 (2015)**: 89-113.

11. Federico Girosi and Gary King. (2008) *.Demographic Forecasting*. Princeton University Press.

12. Wilmoth, J.(1993).Computational Methods for Fitting and Extrapolation the Lee – Carter Model of Mortality Change. Technical report, Department of Demography, University of California, Berkeley.

13. Peter J. Brockwell and Richard A. Davis.(2002). *Introduction to time series and forecasting . 2nd ed* .Springer-verlag New York.