



UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

**Escola Tècnica Superior d'Enginyeria
de Telecomunicació de Barcelona**

**Approximation Model to Analyse Mortality Rate.
An Application for the case of Europe.**

A Master's Thesis

Submitted to the Faculty of the

Escola Tècnica d'Enginyeria de Telecomunicació de Barcelona

Universitat Politècnica de Catalunya

by

Hima Sahiti Seeram

In partial fulfilment

of the requirements for the degree of

MASTER IN TELECOMMUNICATIONS ENGINEERING

Advisor: Marcos Postigo Boix

Jose Luis Melus Moreno

Barcelona, October 2019

Acknowledgments

I would like to express my special thanks and gratitude to my advisors Marcos Postigo and Jose Luis who gave me the opportunity to work on this thesis. And for their constant support and guidance at every step. Through this thesis, I have realised how mortality plays an important role in many research areas as a base foundation.





Abstract

The study of mortality rates has a long history which has been used in the literature of fitting and forecasting. For Part-1, I have modelled mortality rates for 21 countries in Europe using three models: Lee-Carter (LC), Gompertz-Makeham (GM) and Polynomial Model (PM). Each model estimates parameters which contribute in calculating mortality. The accuracy of new Polynomial model directly depends on the number of parameters used. The dataset comprises available data from male and female aged 0-99 from the years 1985-2014. The comparison between the fitted mortality rates of the different proposed models and the actual mortality formula has been based on total mortality rate for each specific gender. This paper focuses on investigating the comparative evaluations of these models based on different errors. And then for Part-2, we examine the accuracy by predicting the missing mortality of one random year in the defined time period and measure the predicting error which resonates similarly to forecasting.

Keywords – Mortality modelling, Lee-Carter, Gompertz-Makeham, Polynomial Model, Fitting Errors, Prediction Errors

Table of Contents

Acknowledgments.....	i
Abstract.....	iii
List of Figures.....	vi
Lists of Tables.....	vii
1. Introduction.....	1
2. Fundamentals.....	3
2.1 Data Description.....	3
2.2 What is Mortality?.....	3
2.3 Age-group mortality.....	8
3. Fitting Mortality Models.....	10
3.1 Flowchart.....	10
3.2 Lee-Carter Model.....	11
3.3 Gompertz-Makeham Model.....	12
3.4 Polynomial Model.....	13
4. Mortality Modelling Results.....	14
4.1 Lee-Carter.....	14
4.1.1 Estimating Parameter.....	14
4.1.2 Mortality Rate.....	14
4.2 Gompertz-Makeham.....	17
4.2.1 Estimating Parameter.....	17
4.2.2 Mortality Rate.....	17
4.3 Polynomial.....	20
4.3.1 Estimating Parameter.....	20
4.3.2 Mortality Rate.....	21
5. Measuring Fitting Errors.....	24
5.1 Root Mean Square Error.....	24
5.2 Mean Relative Error.....	25
5.3 Mean Absolute Error.....	26
5.4 Mean Absolute Relative Error.....	26



6. Predicting Missing Mortality.....	28
6.1 Flowchart.....	28
6.2 Data Description.....	29
6.3 Mortality Rate.....	29
6.4 Measuring Predicting Errors.....	30
7. Conclusions.....	32
8. Future Work.....	33
Bibliography.....	35
Appendices.....	36

List of Figures

Fig. 2.2.1: Female’s Death – ES.....	5
Fig. 2.2.2: Male’s Death – ES.....	5
Fig. 2.2.3: Female’s Population – ES.....	6
Fig. 2.2.4: Male’s Population – ES.....	6
Fig. 2.2.5: Female’s Mortality – ES.....	7
Fig. 2.2.6: Male’s Mortality – ES.....	7
Fig. 2.3.1: Normalized Actual Mortality F/M – ES.....	9
Fig. 4.1.1: LC female mortality – ES.....	15
Fig. 4.1.2: LC male mortality – ES.....	15
Fig. 4.1.3: LC & Actual mortality (Norm.) F/M – ES.....	16
Fig. 4.2.1: GM female mortality – ES.....	18
Fig. 4.2.2: GM male mortality – ES.....	19
Fig. 4.2.3: GM & Actual mortality (Norm.) F/M – ES.....	19
Fig. 4.3.1: PM_5 female mortality – ES.....	22
Fig. 4.3.2: PM_5 male mortality – ES.....	22
Fig. 4.3.3: PM_10 female mortality – ES.....	22
Fig. 4.3.4: PM_10 male mortality – ES.....	22
Fig. 4.3.5: PM_15 female mortality – ES.....	22
Fig. 4.3.6: PM_15 male mortality – ES.....	22
Fig. 4.3.7: PM & Actual mortality (Norm.) F - ES.....	23
Fig. 4.3.8: PM & Actual mortality (Norm.) M - ES.....	23
Figure 6.4.1: Global Predicting Errors- Females.....	31
Figure 6.4.2: Global Predicting Errors- Males.....	31
Figure 8: Mobile Market Churning Process.....	33

List of Tables

Table 1: List of Countries for fitting.....	3
Table 2.3.1: Age-group & total mortality values.....	8
Table 4.1.1: Age-dependent parameter a_x	14
Table 4.1.2: LC mortality rate.....	14
Table 4.2.1: GM parameters.....	17
Table 4.2.2: GM mortality rate.....	17
Table 4.3.1: PM parameters.....	20
Table 4.3.2: Summary of Parameters.....	20
Table 4.3.3: PM mortality rate.....	21
Table 4.3.3: Summary of PM mortality – ES.....	21
Table 5.1: RMSE Fitting error.....	24
Table 5.2: MRE Fitting error.....	25
Table 5.3: MAE Fitting error.....	26
Table 5.4: MARE Fitting error.....	27
Table 5.5: Ranking of Fitting errors.....	27
Table 6.2: List of Countries for Predicting.....	29
Table 6.3: PM (5,10,15) Predicting vs. Actual Mortality.....	29
Table 6.4.1: RMSE Predicting error.....	30
Table 6.4.2: MRE Predicting error.....	30
Table 6.4.3: MAE Predicting error.....	30
Table 6.4.4: MARE Predicting error.....	31

1. Introduction

Social networks have been intensively studied by social scientists for several decades in order to understand both local phenomena, such as relationship formation and their dynamics, as well as network-wide processes, such as transmission of information. The recent substantial interest in the structural and functional properties of complex networks has been partially stimulated by attempts to understand the characteristics of social networks. In everyday social life or professional collaborations, people tend to form relationships, the existence of which is a prominent characteristic of social networks. Network formation has been studied in many research fields with their different focuses. Modelling social networks serves at least two purposes. Firstly, it helps us understand how social networks form and evolve. Secondly, in studying network-dependent social processes by simulation, such as diffusion or retrieval of information, successful network models can be used to specify the structure of interaction [1].

The development of stochastic models for the analysis of social networks is an important growth area within contemporary statistics. The last few decades have witnessed the rapid development of a variety of statistical models capable of representing the global structure for generating a social network. The formation of social network depends mainly on two factors: one, the dynamic ties between the individuals and two, on statistical model of mortality which provides information such as age, gender, population, country [2].

Mortality contributes significantly to population dynamics and is crucial in many fields such as economy, demography and social sciences. Early mortality tables were deterministic and static in nature assuming no further improvement in mortality rates over time and treating all lives as homogeneous with respect to mortality. In more recent years, and as populations age more rapidly, greater attention has been given to the modelling and forecasting of mortality and in particular to the uncertainty surrounding mortality rates.

Early attempts to model mortality did not take account of potential future improvements in mortality rates. Rather, they took current and past data and assumed that future mortality would behave in the same way. An important step towards the development of early age-continuous mortality models came from the early mortality laws originating from the fitting of a mathematical formulae to the mortality data. The first attempt to mathematically model mortality with a continuous formula was proposed in 1725 by Abraham De Moivre, who suggested

$$l_x = k \left(1 - \frac{x}{86}\right) \text{ for } 12 \leq x \leq 86$$

where l_x is the number of individuals still alive at age x and k is a normalizing constant, the assumption in this model being that all individuals will have died by age 86.

The models were more subjective than extrapolative, indicating that they relied heavily on the opinions of the modellers themselves. There are few ways, if at all possible, to obtain exact mortality rates. The most successful approach to modelling mortality in recent decades has been the extrapolative method which relies heavily on data which has become more and more available in recent years.

In 1825 Gompertz's law constitutes one of the most influential proposals in the early times of survival modelling. It is expressed in terms of what we now call a "force" of mortality. In the early 1990's researchers began to look at modelling mortality using time series to extrapolate the time trend based on historic mortality experience. These sorts of models make the implicit assumption that past trends identified in the data will continue into the future. The first and most recognized of these types of models is the Lee Carter mortality model which models the time trend using a one factor stochastic model [9].

The mortality models study various factors:

- **Historical trends:** characterize historical patterns of mortality and project future trends are based on age and time dependent changes in the rate of mortality
- **Contributing factors** responsible for variation in mortality that affect the risk of deaths

The number and complexity of mortality models have grown rapidly since the first law of mortality. A good model provides a simple but adequate mathematical description of mortality by age and/or time. The objective of this thesis is to study three different extrapolative mortality models and summarize their parameters. We design a new model called Polynomial model, which produces better mortality when compared with existing models such as Lee-Carter and Gompertz-Makeham. This new model gives a trade-off between the number of parameters and the accuracy of the mortality. After defining the models, we discuss model specification, estimation and assessment.

This thesis explores a progression of three models. The rest of the paper is organized as follows: Section-2 describes about what mortality is and how it is calculated. Section-3 defines the three models along with new Polynomial model. Section-4 assess the parameters and mortality produced by each model. Section-5 tests the goodness fit of three models by calculating four errors (RMSE, MAE, MRE and MARE). And finally, Section-6 calculates prediction errors for any missing mortality data for the new model.

2. Fundamentals

2.1 Data Description

Our study is based on data obtained from population and death tables of 21 European countries for 3 decades from 1985-2014. The tables are taken from the websites www.mortality.org and <https://ec.europa.eu/eurostat/data/database>. The original raw data collected from reliable sources (mostly as published or distributed by national statistical offices), from which all mortality estimates and life tables are derived. These data tables are separately available for female and male populations for the ages 0-99. To demonstrate the changes in our data, it has been classified into age-groups and total mortality rate for each country.

Table 1: List of 21 Countries for modelling mortality and their abbreviation as stated by Eurostat.

Countries - Abbreviation		
Austria – AT	Finland – FI	Norway – NO
Belarus – BY	France – FR	Poland – PL
Belgium – BE	Germany – DE	Portugal – PT
Bulgaria – BG	Greece – EL	Spain – ES
Czechia – CZ	Italy – IT	Slovakia – SK
Denmark – DK	Lithuania – LT	Sweden – SE
Estonia – EE	Netherlands – NL	Switzerland - CH

2.2 What is Mortality?

Mortality rates look at the number of people who die relative to the size and age of the population at a specified interval (1985-2014), and gives a general measure of that population. It is nothing more than the *incidence* of death in a certain time window.

$$m(x, t) = \frac{C(x,t)}{P(x,t)} = \frac{\text{deaths}(x,t)}{\text{average population}(x,t)}$$

where, Year = t; Age = x

The life table provides the most complete description of mortality in any population. The basic data input needed for its construction are the age-specific death rates calculated from information on both deaths and population by age and sex.

• Death Matrix

Processing of death counts are collected by sex, completed age, year of death. The matrix represents the number of deaths of a given sex at age completed i in year n (C_i^n). The matrix is arranged with the years (1985-2014) in column-wise and ages (0-99) in row-wise.

$$\mathbf{C} = \begin{pmatrix} C_0^{1985} & C_0^{1986} & \dots & C_0^{2014} \\ C_1^{1985} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ C_{99}^{1985} & \dots & \dots & C_{99}^{2014} \end{pmatrix} = \left[C_i^n \right]_{100 \times 30}$$

• Population Matrix

Estimating population size, census data and annual population estimates by sex and age are collected from statistical offices. The matrix represents population of a given sex at age completed i in year n (P_i^n). It is arranged with the years (1985-2015) in column-wise and ages (0-99) as row-wise.

$$\mathbf{P} = \begin{pmatrix} p_0^{1985} & p_0^{1986} & \dots & p_0^{2015} \\ p_1^{1985} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ p_{99}^{1985} & \dots & \dots & p_{99}^{2015} \end{pmatrix} = \left[p_i^n \right]_{100 \times 31}$$

Since we simulated for 21 countries, it is hard to display all the countries graphs, but in order to illustrate an example, we have decided to display Spain throughout the report. The below figures 2.2.1 and 2.2.2 represent death graphs while the figures 2.2.3 and 2.2.4 demonstrate population graphs for both female and male. The graphs are organized as follows: the number of deaths and population on y-axis, years on x-axis and ages on z-axis. And the estimated mortality rate for both genders are displayed in the figures 2.2.5 and 2.2.6 with the mortality rate on y-axis.

Figure 2.2.1: Female's Deaths

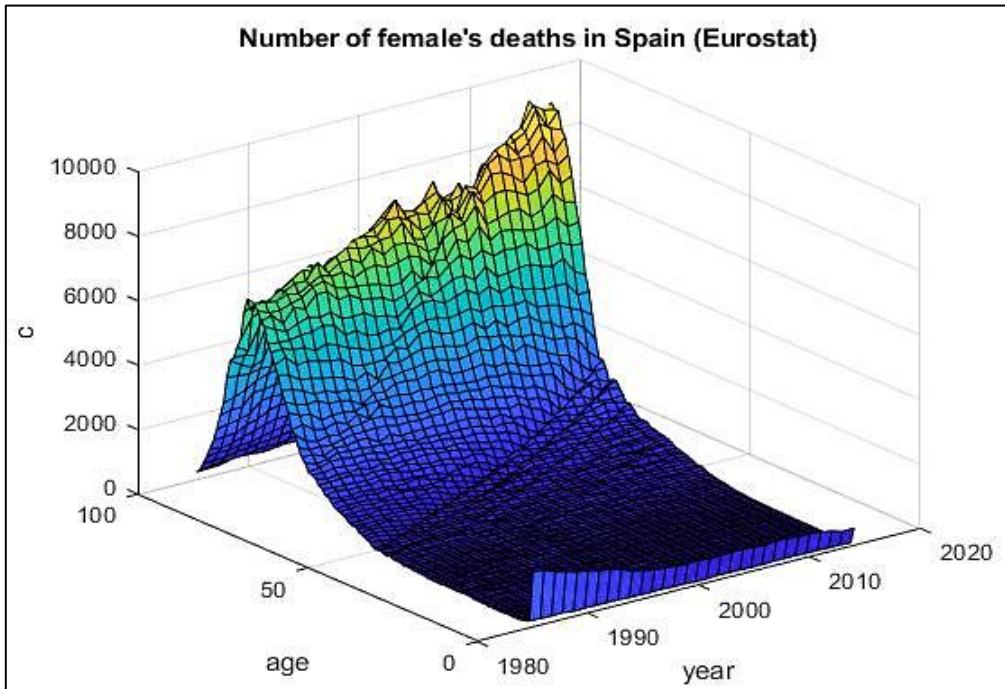
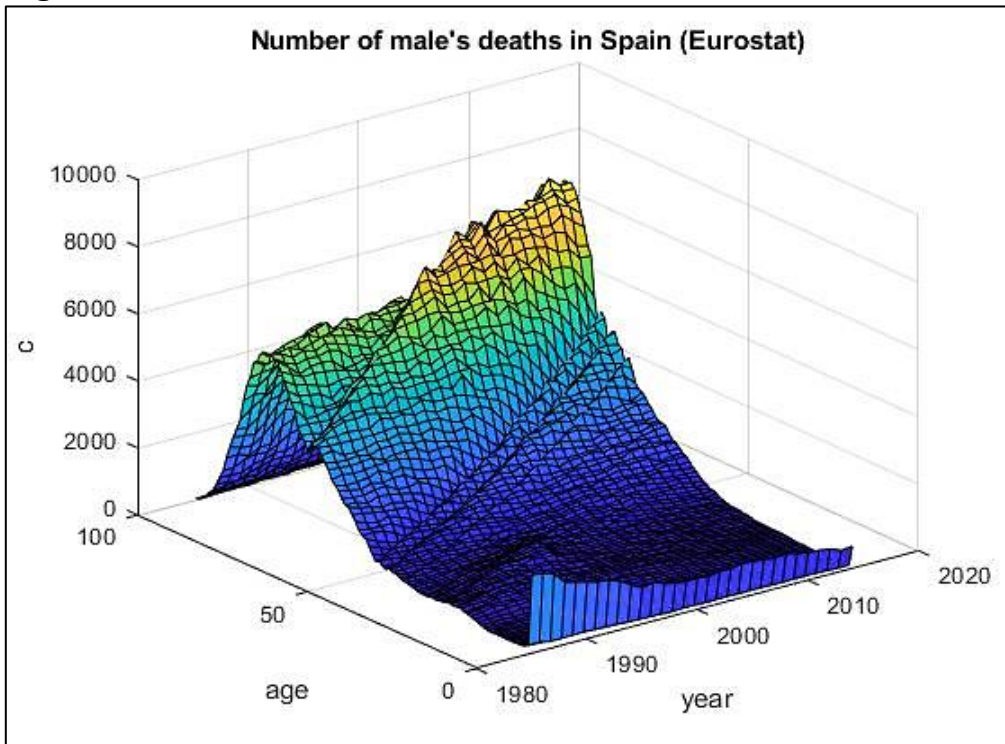


Figure 2.2.2: Male's Deaths



The deaths in males is higher than females by 9.92%. In terms of age-groups, the deaths are higher for age range 51-99 by 157.75% when compared to age range 0-50 for males and for females it is 179.55%. The occurrence of highest deaths in males happened in 2012 at age 83 and for females it is also in 2012 but at age 89.

Figure 2.2.3: Female's Population

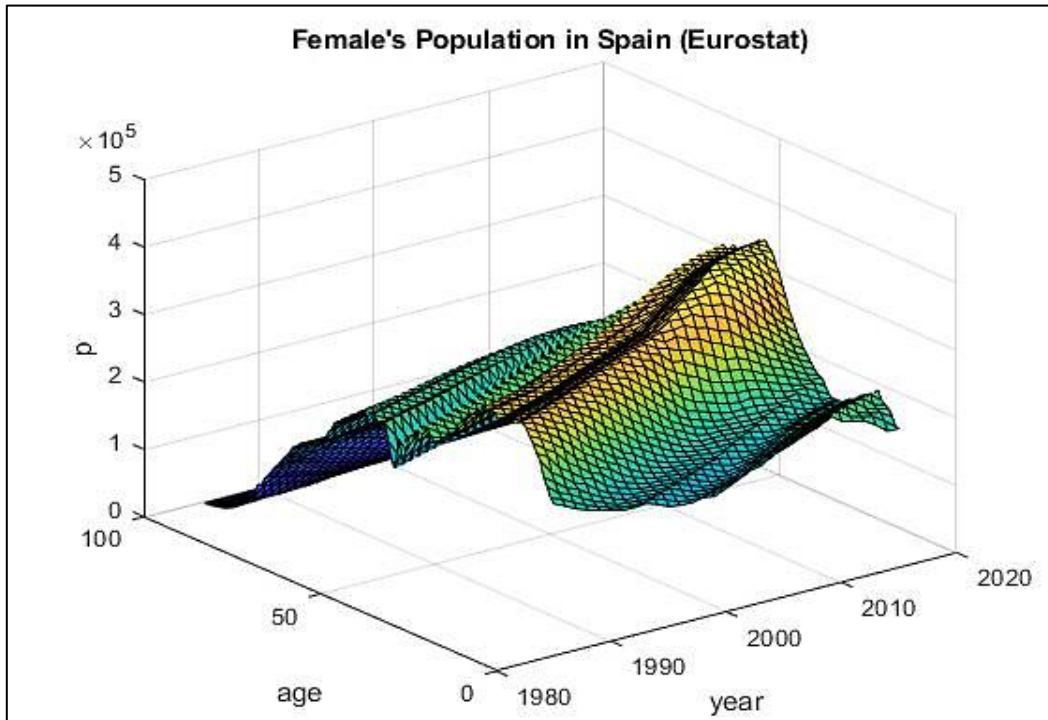
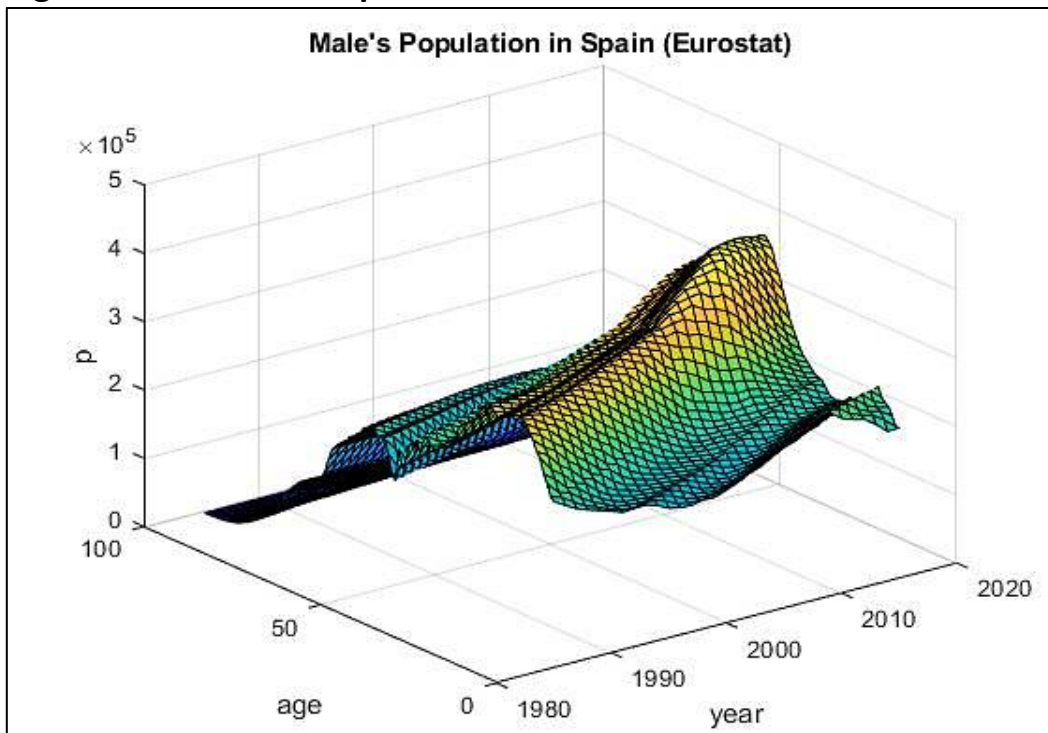


Figure 2.2.4: Male's Population



The population size is higher in females than males by 3.47%. In terms of age-groups, the population is higher for age range 0-50 by 85.16% when compared to 51-99 for males and for females it is 65.88%. The highest population is noticed in 2011 for females at age 34 while for males it is in 2009 at age 32.

Figure 2.2.5: Female's Mortality - Spain

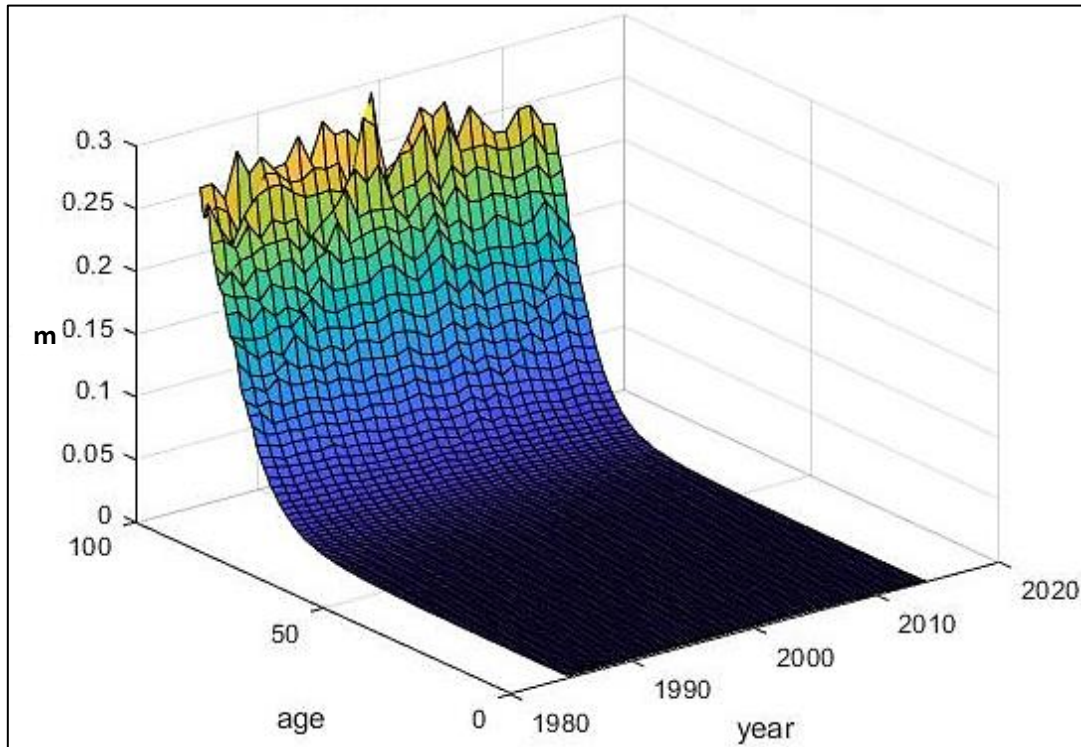
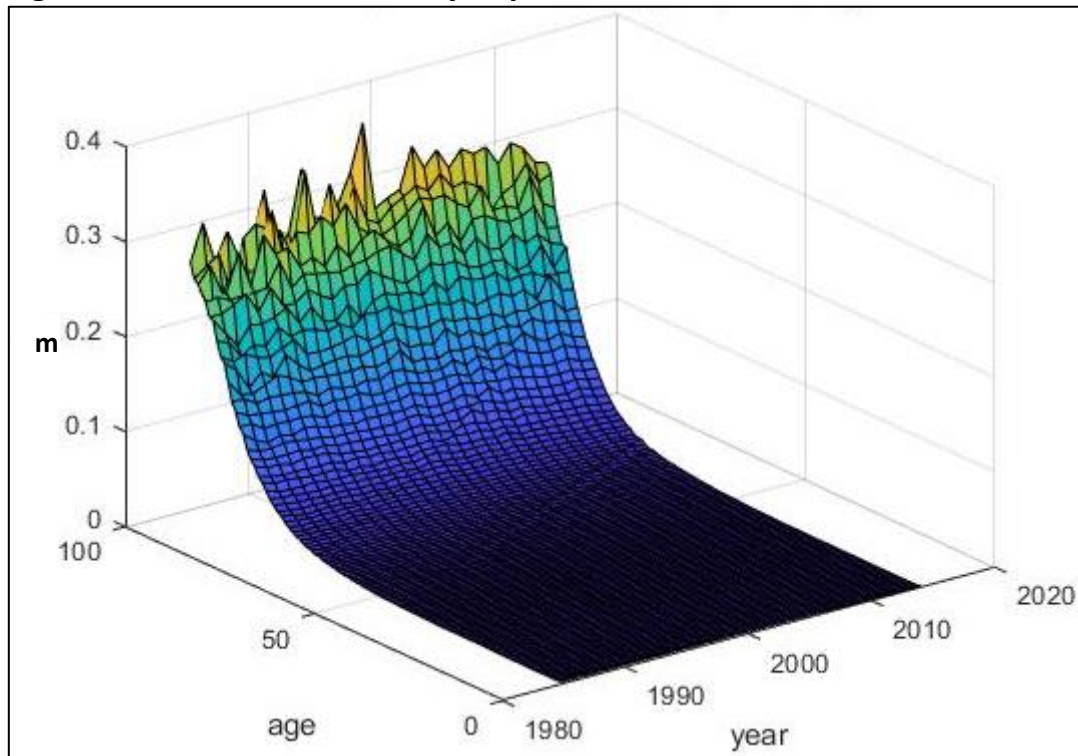


Figure 2.2.6: Male's Mortality - Spain



The males produce higher mortality than females by 26.61%. The highest mortality for both females and males are in the year 1999 at age 99. And the lowest mortality for females is in 2010 at age 8 while for males it is in 2014 at age 9.

2.3 Age-group Mortality

Comparing mortality between age-groups accounts for the difference in the structure of the populations. To observe the difference in the mortality patterns for both female and male populations, the data has been classified as age-group tables. The ages are divided into 5 categories: 0-19, 20-39, 40-59, 60-79, 80-99. Each age group summarizes mortality for the entire time period (1985-2014). And finally, at the end we add mortality of all ages and acquire total mortality for each country.

Tables 2.3.1 displays the mortalities for 21 countries with gender-specific. Mortality rates are very volatile and vary from age group to age group and also from year to year. Refer to Table 1 for abbreviations of countries.

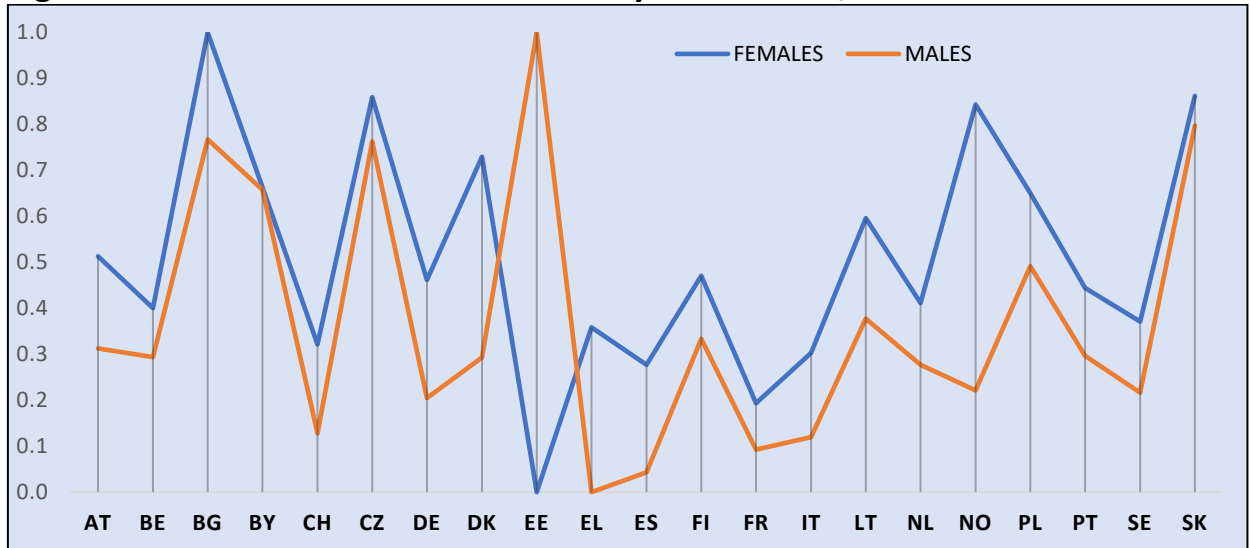
Table 2.3.1: Actual Mortality Rate – Age-group

	0-19		20-39		40-59		60-79		80-99		Total	
	F	M	F	M	F	M	F	M	F	M	F	M
AT	0.17	0.26	0.18	0.46	1.07	2.23	7.45	13.81	83.65	100.55	92.52	117.31
BE	0.17	0.26	0.21	0.48	1.17	2.13	7.33	14.28	77.18	99.19	86.06	116.34
BG	0.37	0.51	0.3	0.7	1.63	4.09	12.97	20.95	105.12	114.83	120.39	141.08
BY	0.29	0.45	0.4	1.51	2.05	6.3	12.78	25.29	85.61	101.83	101.13	135.38
CH	0.16	0.23	0.18	0.44	0.88	1.62	5.88	11.43	74.46	93.95	81.56	107.67
CZ	0.18	0.27	0.2	0.53	1.39	3.3	11.09	19.84	99.46	116.9	112.32	140.84
DE	0.16	0.23	0.19	0.42	1.13	2.27	7.84	14.45	80.23	94.29	89.55	111.66
DK	0.2	0.23	0.3	0.44	1.83	2.28	11.69	15.14	90.88	98.19	104.9	116.28
EE	0.68	0.81	0.4	1.37	1.77	4.77	11.45	23.96	48.84	122.36	63.14	153.27
EL	0.18	0.26	0.17	0.47	0.85	1.89	7.48	12.41	75	85.94	83.68	100.97
ES	0.16	0.23	0.18	0.49	0.83	1.98	5.98	12.17	71.84	88.36	78.99	103.23
FI	0.14	0.22	0.2	0.59	1.04	2.55	7.44	14.94	81.27	100.12	90.09	118.42
FR	0.16	0.24	0.22	0.56	1.04	2.45	5.66	12.14	67.15	90.42	74.23	105.81
IT	0.16	0.23	0.16	0.41	0.88	1.73	6.55	12.54	72.74	92.32	80.49	107.23
LT	0.3	0.46	0.37	1.44	1.89	5.7	10.2	20.63	84.5	92.46	97.26	120.69
NL	0.16	0.22	0.18	0.3	1.12	1.71	7.35	14.08	77.9	99.16	86.71	115.47
NO	0.33	0.23	0.21	0.42	1.24	1.72	10.36	13.05	99.25	97.13	111.39	112.55
PL	0.28	0.41	0.23	0.77	1.57	4.23	10.5	20.05	87.78	101.23	100.36	126.69
PT	0.23	0.35	0.24	0.71	1.08	2.49	7.86	14.39	79.16	98.53	88.57	116.47
SE	0.13	0.17	0.16	0.35	0.96	1.56	6.83	11.93	76.28	98.26	84.36	112.27
SK	0.26	0.36	0.22	0.64	1.54	4.20	11.53	21.25	98.93	116.2	112.48	142.65

It can be observed that males have higher mortality than females for all countries stated. The highest mortality for both males and females in the 80-99 age range and least mortality in the 0-39 age range which is the normal pattern. In comparison to all the countries mortality, it is observed that Bulgaria has the highest female mortality whereas Estonia has the lowest. Bulgaria has 62.39% higher female mortality than Estonia. And for males, Estonia has the highest and Greece has the lowest. Estonia shows 41.14% higher male mortality than Greece.

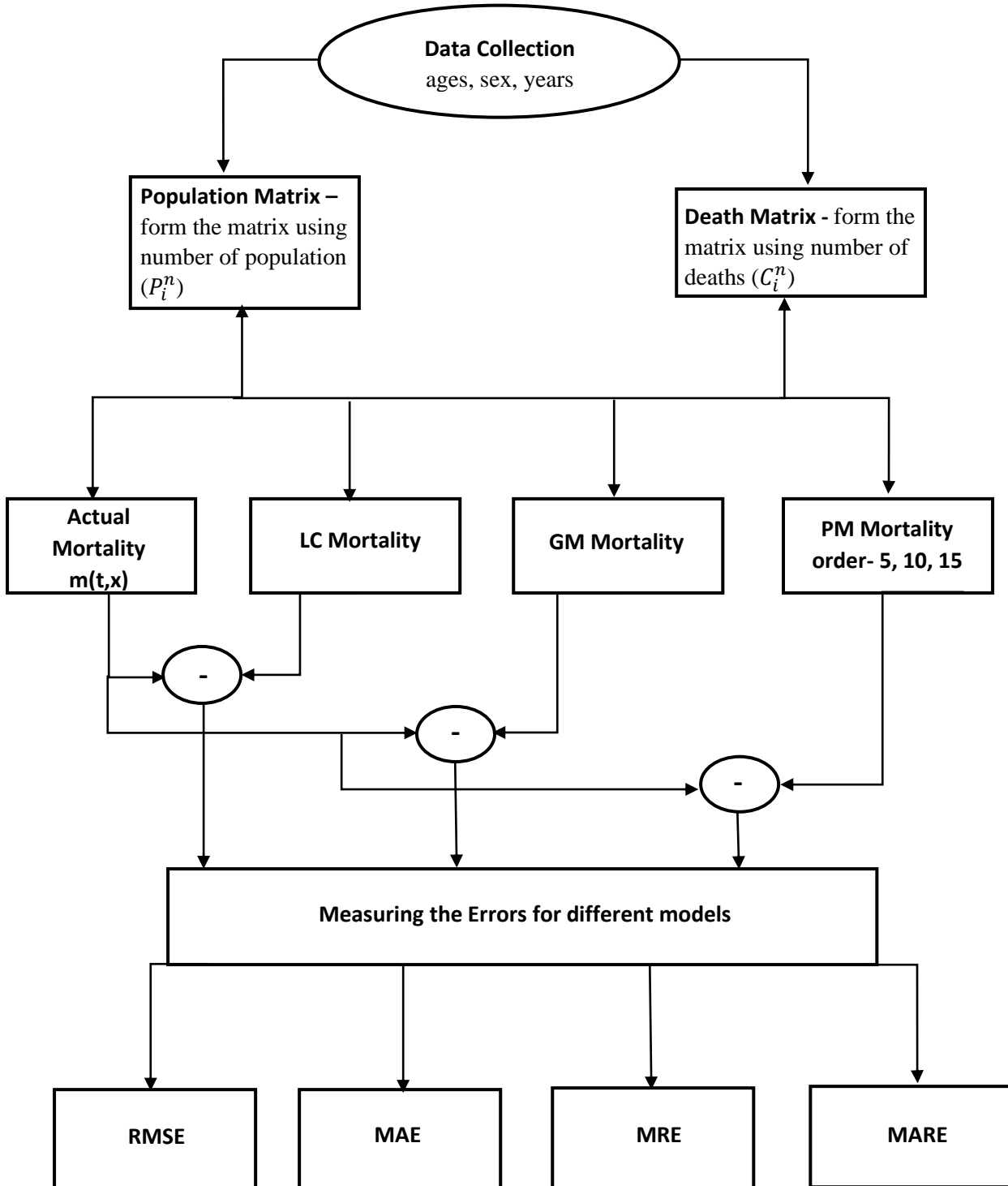
In order to observe the uniform differences in mortality data values mentioned in Table 2.3.1, we have normalized total mortality rate between 0 and 1 in the below figure 2.3.1.

Figure 2.3.1: Normalized Actual Mortality for Females/Males



3. Fitting Mortality Models

3.1 Flowchart



The above flowchart shows the steps involved in modelling mortality for 21 countries. At first, we obtain raw data with ages, deaths, population size for each gender. Then, we obtain death and population matrices and calculate mortality using the formula in Section 2.2. The next step, we fit mortality using three different models as stated in upcoming sections. And finally, calculate four errors (RMSE, MRE, MAE and MARE) to test the goodness of these models.

3.2 Lee-Carter Model

The Lee and Carter model, published in 1992, was the first attempt to model longevity data in a stochastic fashion by fitting the past mortality data and modelling the time trend as a stochastic process. The benefit of this model is it does not require subjective judgment or causes of death. Instead it models the data as a stochastic time series. It has become the baseline model against which all stochastic models of mortality have since been compared. Since it incorporates period (in years) and age mortality dynamics, the modelling mortality is based on past trends in age and time. The formulation of the Lee-Carter model is as follows [4]:

$$\log m_{xt} = a_x + b_x k_t + e_{xt}$$

Usually, x represents age completed i , and t year n in the death and population matrices. The Singular Value Decomposition (SVD is a factorization of a real or complex matrix) approach is used for estimation of the parameters \mathbf{b}_x and \mathbf{k}_t of the LC model. Lee and Carter seek to summarize and age-period surface of log-mortality rates $\log m(\mathbf{x}, \mathbf{t})$ in terms of vectors \mathbf{a} represents the general tendency in mortality for different ages and \mathbf{b} shows the rate of change in the mortality with respect to changes in \mathbf{k}_t . The mortality index, \mathbf{k}_t , demonstrates the period effect which is the relationship between time-dependent events and mortality rates, and the error term, e_{xt} takes into consideration the random historical fluctuations not captured by the model. The e_{xt} is assumed to be an independent and identically distributed Gaussian random variable with mean 0 and variance σ^2 [3].

Lee-Carter model estimate the parameters by imposing two constraints on \mathbf{b}_x and \mathbf{k}_t to avoid the identification issue that arises when there is more than one solution to a parameter estimate. Uniqueness of parameters is therefore ensured. \mathbf{b}_x is thus summed to unity and the summation of \mathbf{k}_t is brought to zero.

With the constraint imposed on the mortality index, it can be concluded that the parameter is being distributed equally with the instances eliminating each other so that the sum of all the time-dependent parameters gives a value of zero.


3.3 Gompertz-Makeham Model

The Gompertz–Makeham law states that the human death rate is the sum of an age-dependent parameter (the Gompertz parameter, named after Benjamin Gompertz), which increases exponentially with age and an age-independent parameter (the Makeham parameter, named after William Makeham). In a protected environment where external causes of death are rare (laboratory conditions, low mortality countries, etc.), the age-independent mortality component is often negligible. In 1825, Benjamin Gompertz proposed an exponential increase in death rates with age [8].


For many purposes, the Gompertz model provides a satisfactory fit to adult mortality rates. However, a close inspection of the difference between model estimates and observed death rates often reveals systematic underestimation of actual mortality at youngest adult ages (younger than 40) and overestimation at the oldest ages (over 80). The deviation at lower ages was addressed by Makeham (1860) with the addition of a constant to the Gompertz model. This new parameter is usually referred to as background mortality, which is the same for all the years. The Makeham model represents a clear improvement over the Gompertz model at younger ages, but it still overestimates mortality at the oldest ages [5].

The decline in the human mortality rate before the 1950s was mostly due to a decrease in the age-independent (Makeham) mortality parameter, while the age-dependent (Gompertz) mortality parameter was surprisingly stable.

$$m_{x,t} = h(x, t) \approx \beta_t e^{\gamma t^x} + \alpha_t$$



Gompertz parameter



Makeham parameter

The Makeham parameter, a constant mortality rate independent of aging added to the Gompertz law of human mortality, is proposed to be a measure of the impact on mortality rate by extrinsic causes of mortality, with the effect of aging removed.

3.4 Polynomial Model

This new model has been designed for this project whereas Lee-Carter and Gompertz-Makeham models have already existed. The new polynomial model has been introduced to show that the mortality fitting directly depends upon the number of parameters which varies on the order of the polynomial and also ought to produce better mortality in comparison with other models.

Step 1: We acquire logarithm of actual mortality and obtain polynomial fitting for each year over the ages (0-99)

$$\text{Poly_Fit} = \begin{pmatrix} \log(m_{0,1985}) & \cdots & \log(m_{0,2015}) \\ \vdots & \ddots & \vdots \\ \log(m_{99,1985}) & \cdots & \log(m_{99,2015}) \end{pmatrix}$$

Step 2: Then approximate coefficients for the above fitted polynomials (**Poly_Fit**). Number of coefficients is equal to the order of polynomial plus 1.

$$\text{Poly_Mod.p} = \begin{pmatrix} \mathbf{a}_{11} & \cdots & \mathbf{a}_{1, \text{coeff}} \\ \vdots & \ddots & \vdots \\ \mathbf{a}_{30,1} & \cdots & \mathbf{a}_{30, \text{coeff}} \end{pmatrix} \text{ years x coefficients}$$

Step 3: We approximate coefficient matrix with a polynomial over the years from the **Poly_Mod.p**.

$$\text{Poly_Mod.q} = \begin{pmatrix} \mathbf{q}_{1,1} & \cdots & \mathbf{q}_{1, \text{coeff}} \\ \vdots & \ddots & \vdots \\ \mathbf{q}_{\text{coeff}, 1} & \cdots & \mathbf{q}_{\text{coeff}, \text{coeff}} \end{pmatrix}$$

Step 4: We obtain mortality by applying the approximated coefficients (**Poly_Mod.q**) for both number of years and ages.

$$\mathbf{P}_{\text{coeff}} = \mathbf{q}_{1,1}\mathbf{t} + \mathbf{q}_{1,2}\mathbf{t}^2 + \dots + \mathbf{q}_{\text{coeff}, \text{coeff}}\mathbf{t}^{\text{coeff}}$$

$$\mathbf{m}_{x,t} = \mathbf{p}_1\mathbf{x} + \mathbf{p}_2\mathbf{x}^2 + \dots + \mathbf{p}_{\text{coeff}}\mathbf{x}^{\text{coeff}}$$

4. Mortality Modelling Results

4.1 Lee-Carter

4.1.1 Estimating Parameters

This section presents the results of estimation of parameters in LC model. Estimated values of age-dependent parameter a_x are reported in Table 4.1.1 which is a summation of all ages 0-99. The estimated parameter vector a_x is determined as the average over time of the logarithm of the mortality rates.

$$\sum_x b_x = 1 \text{ and } \sum_t k_t = 0$$

Table 4.1.1: Age-dependent parameter a_x

	a_x			a_x			a_x	
	F	M		F	M		F	M
AT	-609.77	-548.76	DK	-563.80	-550.81	LT	-562.43	-481.68
BE	-604.20	-547.19	EE	-554.87	-469.88	NL	-611.14	-565.65
BG	-558.38	-500.26	EL	-617.60	-555.68	NO	-581.27	-559.83
BY	-554.00	-473.10	ES	-622.16	-555.33	PL	-582.86	-508.65
CH	-622.34	-564.52	FI	-609.55	-540.07	PT	-595.60	-528.39
CZ	-589.05	-524.97	FR	-614.38	-546.28	SE	-621.70	-572.40
DE	-607.35	-551.65	IT	-623.07	-562.24	SK	-580.37	-510.25

4.1.2 Mortality Rate

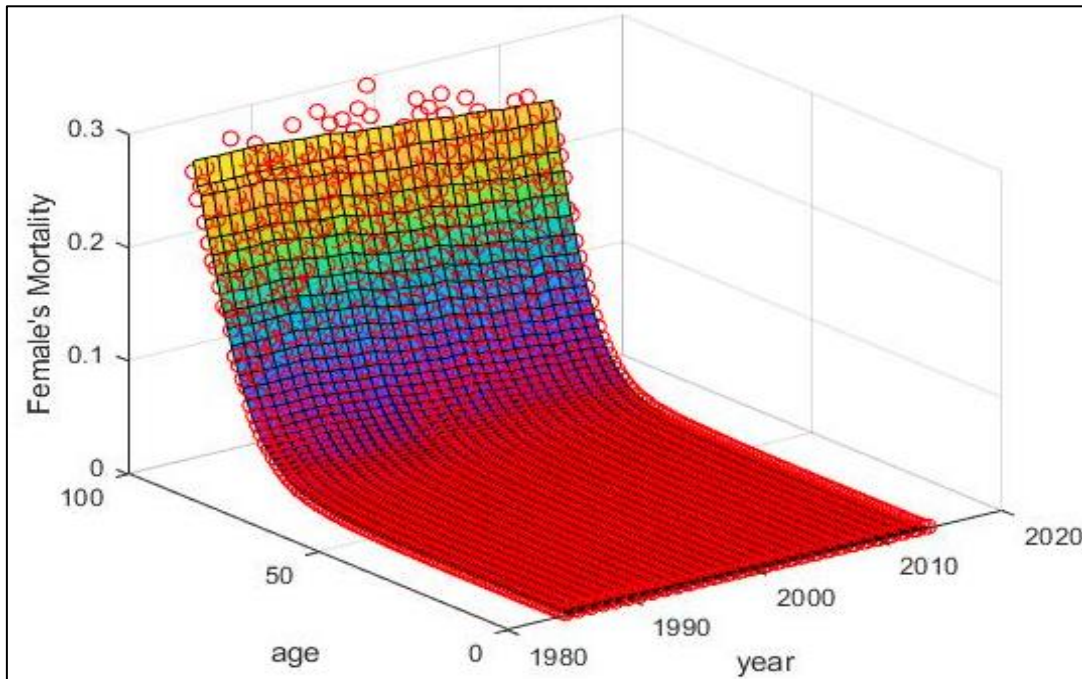
Table 4.1.2 represents mortality produced by LC model which is a total for all years (1985-2014) and ages (0-99). LC model produces similar mortality trends as the Actual mortality as stated in Table 2.3.1

Table 4.1.2: LC mortality rate

	F	M		F	M		F	M
AT	92.36	116.81	DK	104.00	115.68	LT	96.74	119.78
BE	85.93	115.90	EE	63.14	149.11	NL	86.62	115.26
BG	119.12	139.31	EL	83.39	100.72	NO	110.73	112.01
BY	100.45	134.58	ES	78.91	103.08	PL	100.20	126.39
CH	81.43	107.28	FI	89.77	117.42	PT	88.40	116.07
CZ	112.04	139.91	FR	74.15	105.69	SE	84.27	112.01
DE	89.38	110.96	IT	80.37	107.06	SK	111.92	140.43

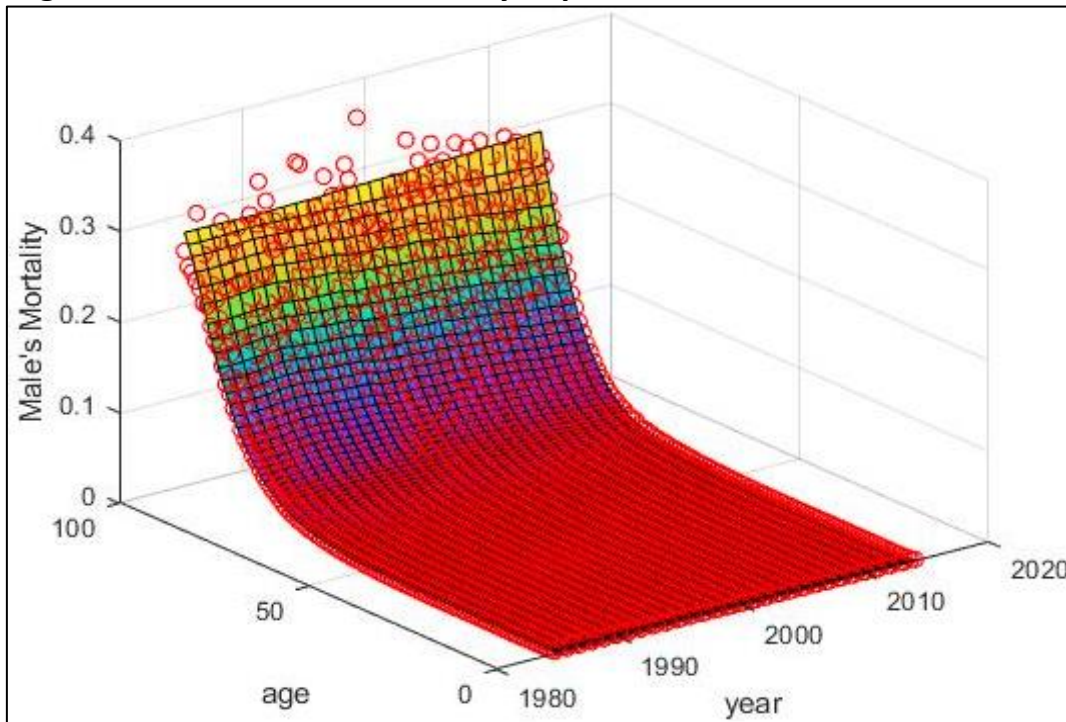
The red circles are scattered over the graph represents the Actual mortality to demonstrate the difference produced by LC model.

Figure 4.1.1: Female's LC Mortality - Spain



LC model produces 0.1% less total mortality than Actual for females. From figure 4.1.1, the highest mortality occurrence for LC happens in 1986 at age 99 whereas for Actual it was 1999 at same age. The lowest mortality appears in 2014 at age 8 for LC but for Actual, it was 2010 at same age.

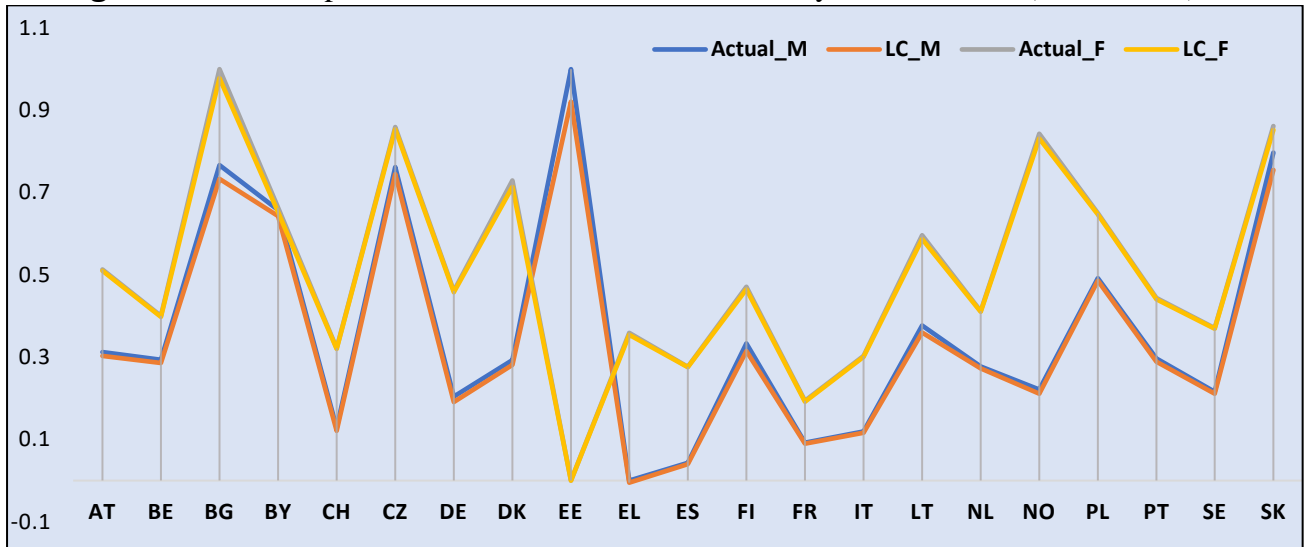
Figure 4.1.2: Male's LC Mortality - Spain



LC model shows 0.15% lower mortality than Actual for Males. From the above figure, the highest mortality occurrence for LC shows in 2014 at age 99 while for Actual mortality, it was in 1999 at the same age. And for the lowest mortality, it was in 2014 for both Actual and LC model but at different ages (Actual at age 9 and LC at 11).

In figure 4.1.3, we represent normalized mortality differences between Lee-Carter and Actual mortality for both genders.

Figure 4.1.3: Comparison between LC & Actual mortality- Female/Male (Normalized)



4.2 Gompertz-Makeham

4.2.1 Estimating Parameters

GM model consists of three parameters that need to be fitted in order to calculate the mortality rate. α and β parameters vary between populations over years whereas γ is affected by age over time. The results of the estimated parameters are reported in Table 4.2.1 for all the countries for each gender.

Table 4.2.1: GM Parameters

	α		β		γ	
	F	M	F	M	F	M
AT	4.23E-08	0.00399	0.000319	0.001445	3.219	2.848
BE	4.17E-07	0.000927	0.000398	0.002151	3.113	2.700
BG	0.002773	0.00132	0.003288	0.007465	2.550	2.263
BY	1.04E-08	0.0017	0.003417	0.013736	2.407	1.980
CH	2.50E-05	0.002313	0.000127	0.000727	3.451	3.024
CZ	3.46E-08	0.002319	0.001008	0.003518	2.872	2.561
DE	0.000749	0.002831	0.000393	0.002574	3.119	2.699
DK	0.001473	0.004092	0.001138	0.001426	2.778	2.766
EE	7.52E-05	0.181499	0.009947	0.006361	1.851	3.430
EL	2.35E-08	0.000822	0.000947	0.001478	2.795	2.690
ES	6.68E-09	0.000427	0.000286	0.001269	3.216	2.771
FI	1.25E-05	0.00816	0.000316	0.00141	3.220	2.811
FR	1.02E-08	0.003215	0.000168	0.001001	3.349	2.874
IT	1.02E-07	5.65E-08	0.000237	0.001191	3.260	2.812
LT	0.007974	0.034933	0.001691	0.008061	2.852	2.254
NL	3.32E-07	2.40E-07	0.000239	0.001592	3.234	2.728
NO	0.003252	0.000579	0.000422	0.001128	3.113	2.811
PL	1.76E-08	0.001146	0.001411	0.006967	2.767	2.279
PT	1.11E-07	0.000125	0.000592	0.001879	2.957	2.654
SE	3.14E-06	6.11E-08	0.000197	0.000782	3.321	2.945
SK	0.00087	0.030258	0.000839	0.003549	2.912	2.593

4.2.2 Mortality Rate

Table 4.2.4 represents mortality produced by GM model which is a total for all years (1985-2014) and ages (0-99). GM model produces similar mortality trends as the Actual mortality as stated in Table 2.3.1

Table 4.2.4: GM mortality rate

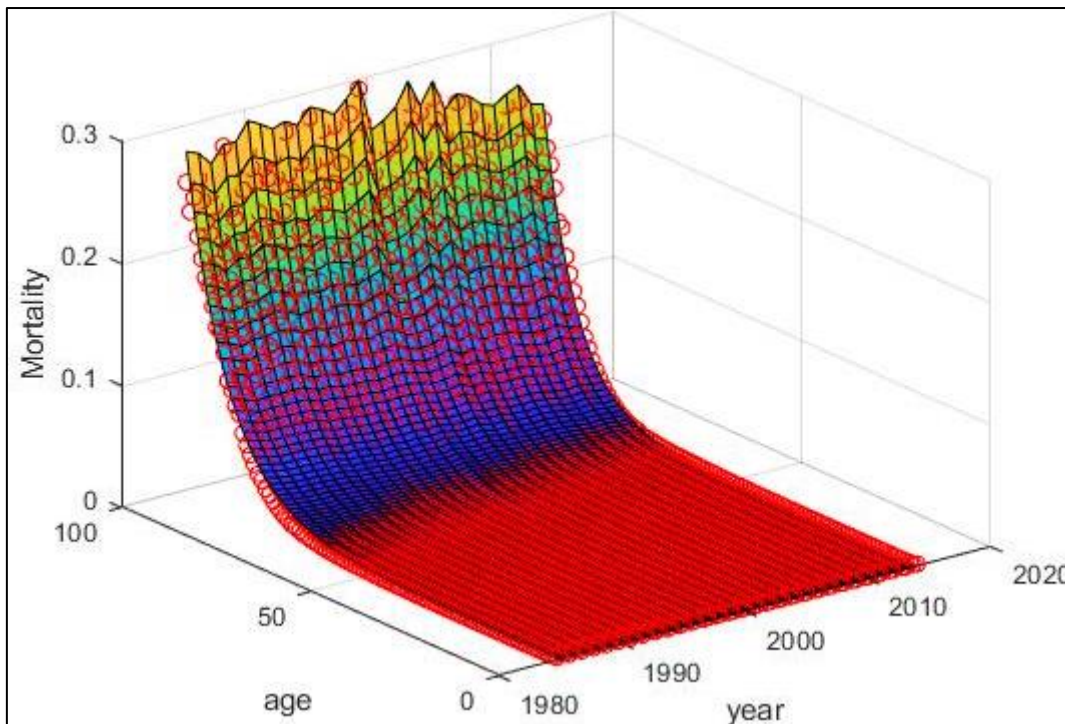
	F	M		F	M		F	M
AT	94.25	118.16	DK	106.08	116.48	LT	99.37	121.54
BE	87.75	118.26	EE	67.76	154.54	NL	87.89	117.81
BG	125.71	143.91	EL	87.16	102.32	NO	112.07	113.75
BY	104.59	137.13	ES	81.05	104.39	PL	103.35	128.56
CH	82.57	108.19	FI	91.39	118.78	PT	91.11	117.56

CZ	115.66	142.23	FR	75.21	106.27	SE	85.50	113.79
DE	91.05	112.47	IT	81.92	108.69	SK	114.20	143.65

The red circles are scattered over the graph represents the Actual mortality to demonstrate the difference produced by GM model.

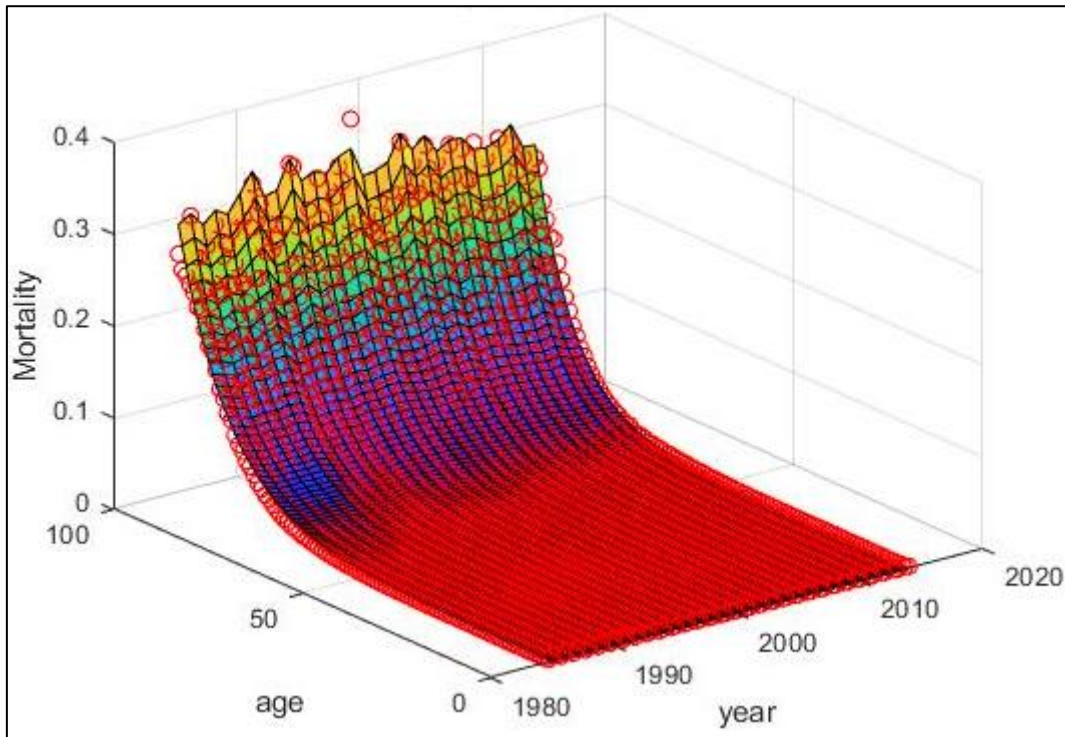
GM model produces 2.57% more total mortality than Actual mortality. From below figure 4.2.1, the highest occurrence of mortality for GM model is in 1999 at age 99 which is same for Actual mortality. While the lowest mortality for GM is in 2011 at age 0 but for Actual it was in 2010 at age 8.

Figure 4.2.1: Female’s GM Mortality - Spain



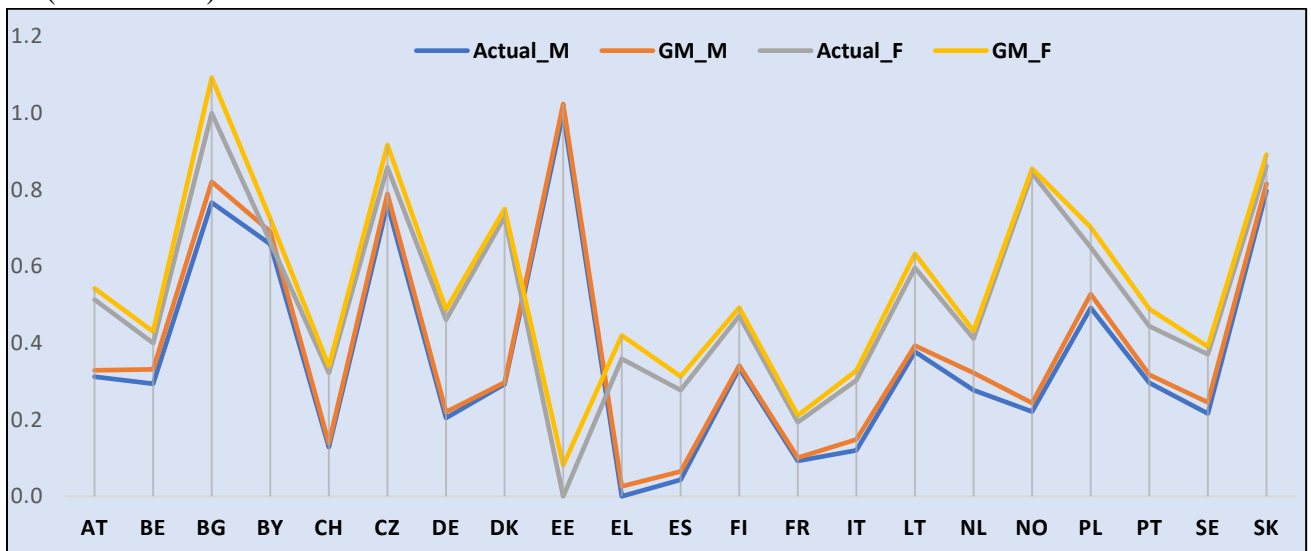
For the male mortality, GM produces 1.12% more than Actual mortality. From figure 4.2.2, the highest mortality occurrence for GM happens in 1994 at age 99 whereas for Actual it was 1999 at same age. The lowest mortality appears in 2014 at age 0 for GM and for Actual, also in the same year but at age 9.

Figure 4.2.2: Male's GM Mortality - Spain



In figure 4.2.3, we represent normalized mortality differences between Gompertz-Makeham and Actual mortality for both genders.

Figure 4.2.3: Comparison between GM & Actual mortality- Female/Male (Normalized)



4.3 Polynomial Model

4.3.1 Estimating Parameters

To demonstrate the difference in mortalities produced by different orders of polynomial, we have considered three orders (5, 10, 15). The below table 4.3.1 will display the summation of coefficients (**Poly_mod.q**) for each order with gender-specific. The table 4.3.2 compares the number of parameters used by each model and Polynomial model of order 15 uses the highest parameters followed by Lee-Carter.

Number of Coefficients = Order + 1

Table 4.3.1: PM parameters

	5		10		15	
	F	M	F	M	F	M
AT	-3.75	-3.25	-3.80	-3.27	-3.79	-3.27
BE	-3.74	-3.22	-3.81	-3.24	-3.77	-3.22
BG	-3.12	-2.71	-3.12	-2.78	-3.08	-2.78
BY	-3.15	-2.62	-3.18	-2.63	-3.11	-2.59
CH	-3.93	-3.41	-4.02	-3.45	-3.97	-3.45
CZ	-3.39	-2.92	-3.43	-2.99	-3.41	-3.00
DE	-3.71	-3.23	-3.75	-3.23	-3.72	-3.21
DK	-3.29	-3.15	-3.34	-3.15	-3.36	-3.13
EE	-3.28	-2.60	-3.33	-2.60	-3.21	-2.58
EL	-3.71	-3.30	-3.70	-3.29	-3.66	-3.25
ES	-3.97	-3.33	-3.96	-3.34	-3.95	-3.33
FI	-3.81	-3.23	-3.87	-3.25	-3.83	-3.24
FR	-4.05	-3.40	-4.08	-3.42	-4.06	-3.41
IT	-3.89	-3.33	-3.92	-3.34	-3.90	-3.33
LT	-3.44	-2.86	-3.45	-2.89	-3.39	-2.84
NL	-3.72	-3.21	-3.78	-3.22	-3.74	-3.20
NO	-3.38	-3.28	-3.38	-3.30	-3.35	-3.29
PL	-3.47	-2.93	-3.51	-2.94	-3.47	-2.93
PT	-3.74	-3.22	-3.71	-3.18	-3.70	-3.17
SE	-3.76	-3.32	-3.81	-3.36	-3.81	-3.35
SK	-3.29	-2.81	-3.27	-2.80	-3.27	-2.81

Table 4.3.2: Summary of Parameters

Models	Number of Parameters
LC	230
GM	90
PM_5	36
PM_10	121
PM_15	256

4.3.2 Mortality Rate

Table 4.3.4 displays the mortalities produced by orders (5, 10, 15). For all countries, it is observed that higher the order of polynomial the closer the mortality gets to the Actual mortality (Table 2.3.1). It can also be observed from the table 4.3.2 as the order of polynomial increases, the number of parameters required in estimating accurate mortality increases.

Table 4.3.3: PM mortality rate

	5		10		15	
	F	M	F	M	F	M
AT	92.26	117.04	92.36	116.89	92.46	117.02
BE	85.89	116.81	85.94	115.98	86.00	116.08
BG	120.13	139.86	119.35	139.66	119.59	139.91
BY	101.69	136.29	100.84	134.86	100.95	135.00
CH	81.10	107.65	81.45	107.33	81.53	107.42
CZ	112.09	140.02	112.05	139.92	112.18	140.17
DE	89.33	111.36	89.35	111.04	89.49	111.44
DK	104.13	116.16	104.07	115.75	104.51	115.89
EE	64.20	147.36	63.41	147.33	63.93	148.81
EL	84.16	101.52	183.40	100.70	83.50	100.83
ES	79.07	103.66	78.92	103.09	78.97	103.19
FI	89.88	117.98	89.84	117.57	89.94	117.86
FR	73.98	105.76	74.14	105.69	74.20	105.75
IT	80.26	107.46	80.41	107.09	80.47	107.22
LT	97.05	120.51	96.90	119.91	97.04	120.14
NL	86.38	115.74	86.64	115.27	86.71	115.36
NO	110.81	112.95	110.78	112.04	110.96	112.16
PL	100.27	126.95	100.19	126.34	100.30	126.49
PT	88.78	117.11	88.40	116.09	88.46	116.22
SE	83.94	112.47	84.26	112.01	84.32	112.14
SK	112.01	140.26	111.97	140.64	112.15	141.04

Figures 4.3.1- 4.3.6 displays mortality rates for Spain, both female and male of Polynomial model with orders – 5, 10 and 15 with Actual mortality scattered on the graphs to show the differences.

In females, polynomial order of 5 gives 0.1% higher mortality, order 10 produces 0.09% lower mortality and order 15 gives 0.03% lower mortality than Actual mortality.

And in males, polynomial order of 5 gives 0.42% higher mortality, order 10 gives 0.14% lower mortality and order 15 gives 0.04% lower mortality than Actual mortality.

Table 4.3.3: Summary of low and high mortality with year and age

	PM_5		PM_10		PM_15		Actual	
	F	M	F	M	F	M	F	M
Low	2014(12)	2014(11)	2014(6)	2014(7)	2014(9)	2014(10)	2010(8)	2014(9)
High	2014(99)	2010(99)	1995(99)	1995(99)	1997(98)	1996(98)	1999(99)	1999(99)

Figure 4.3.1: PM-5

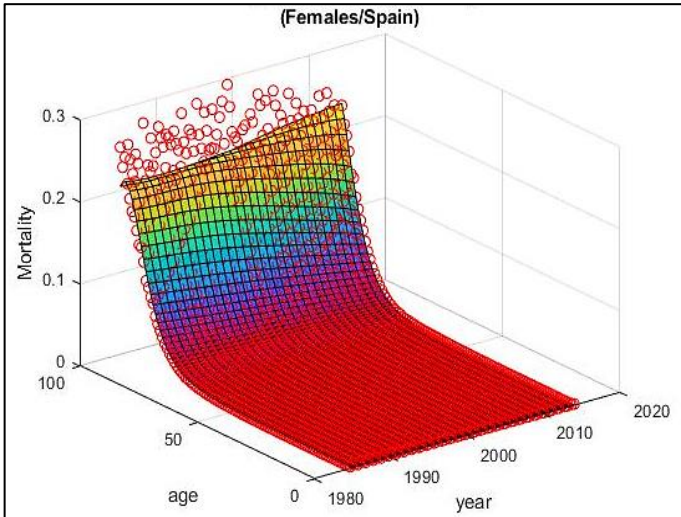


Figure 4.3.2: PM-5

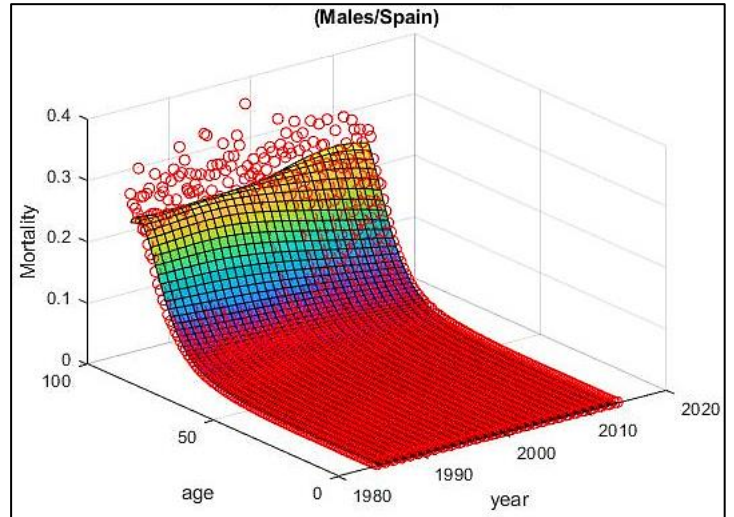


Figure 4.3.3: PM-10

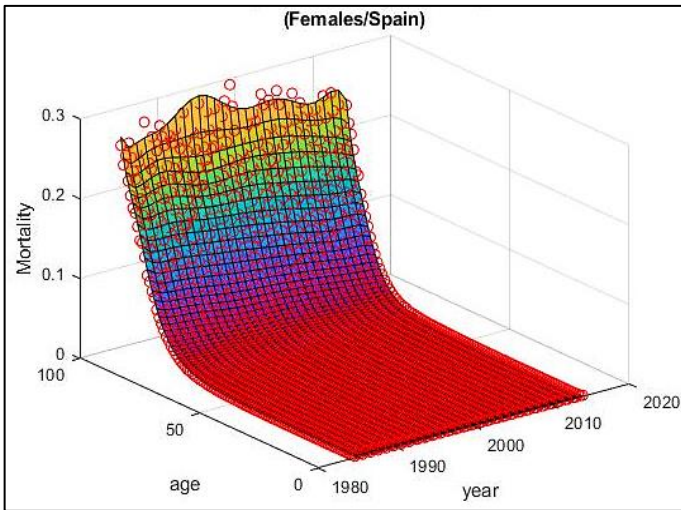


Figure 4.3.4: PM-10

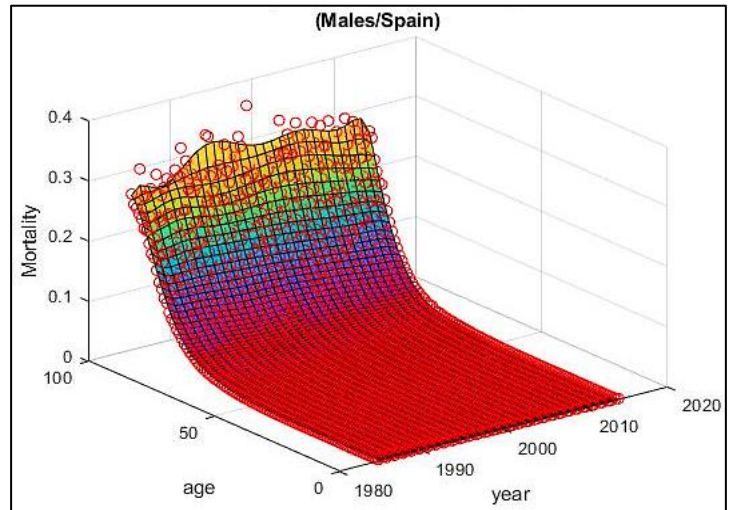


Figure 4.3.5: PM-15

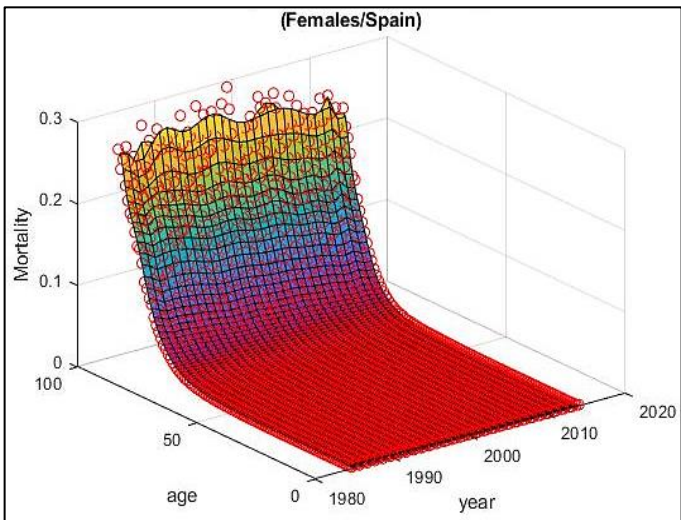
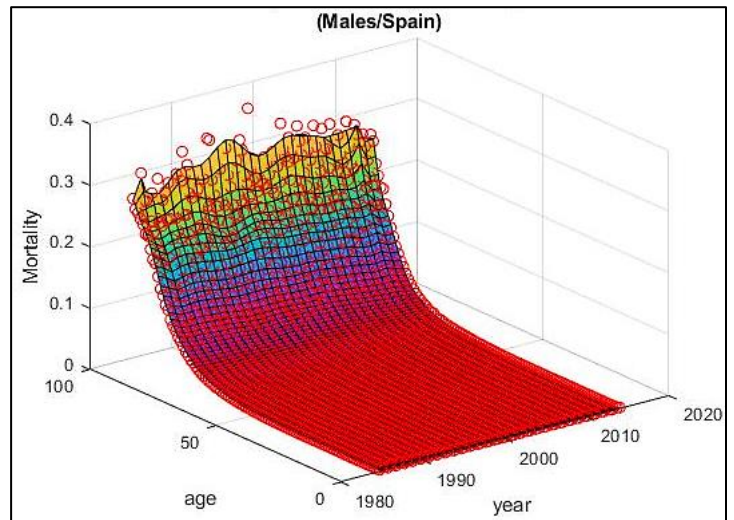


Figure 4.3.6: PM-15



In the below figures 4.3.7 and 4.3.8, we demonstrate the differences of normalized mortality produced by different orders (5, 10 and 15) of Polynomial model and Actual mortality for both genders.

Figure 4.3.7: Comparison between PM & Actual mortality- Female (Normalized)

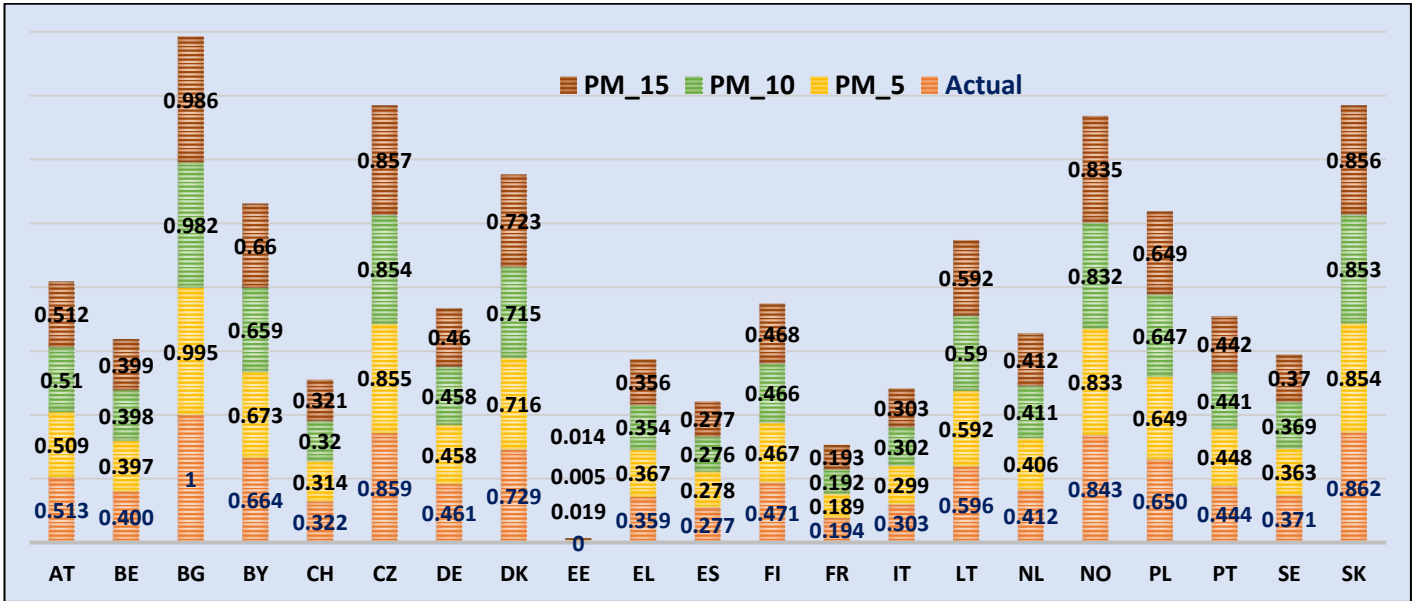
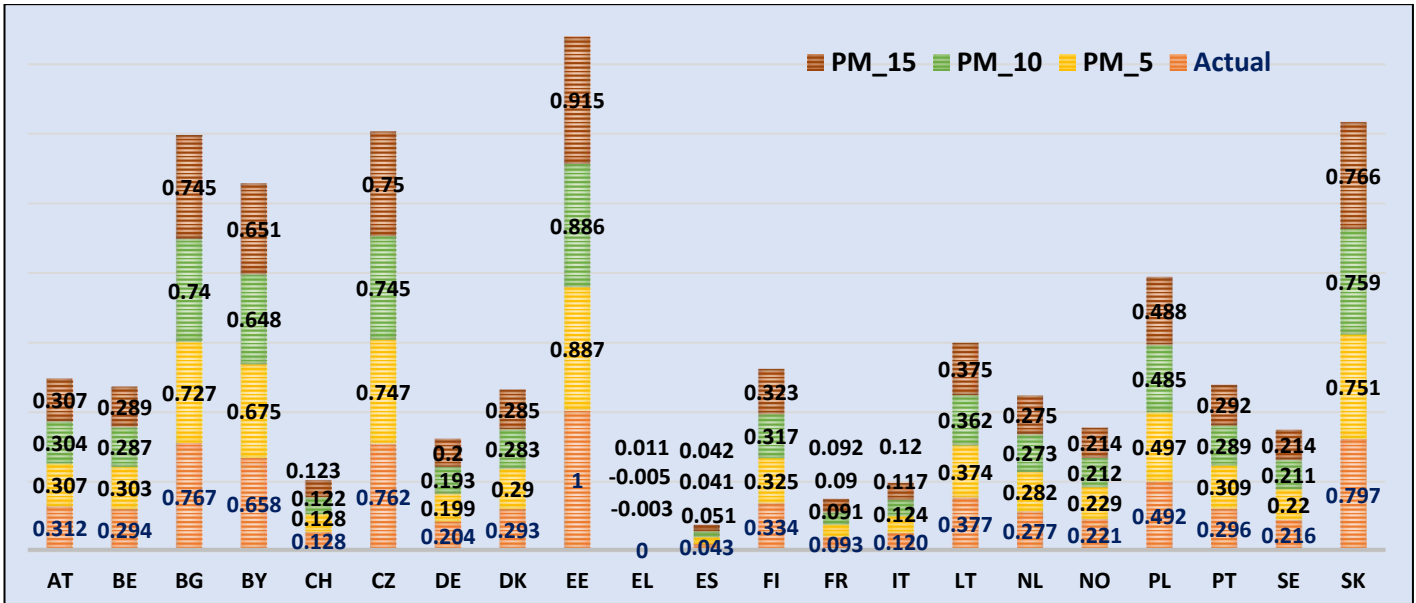


Figure 4.3.8: Comparison between PM & Actual mortality- Male (Normalized)



5. Measuring Fitting Errors

The error calculation is a useful tool in determining the precision of the models. To determine which models produces the best results for four different types of errors. The first step to estimate error depends on calculating the difference in mortality produced by the models and the Actual mortality, that is, the difference between estimated values and actual values.

$$\text{Error} = \text{Models Mortality } (\widehat{m}_i) - \text{Actual Mortality } (m_i)$$

5.1 Root Mean Square Error (RMSE)

It is the standard deviation of the residuals (prediction errors). Residuals are a measure of how far from the regression line data points are. Table 5.1 displays RMSE error for each model with gender specific and for every country.

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^N (\widehat{m}_i - m_i)^2}{N}}$$

where, N = number of data points, 3000(100 ages x 30 years)

Table 5.1: RMSE Fitting error

	LC		GM		PM_5		PM_10		PM_15	
	F	M	F	M	F	M	F	M	F	M
AT	0.148	0.245	0.191	0.242	0.249	0.368	0.142	0.237	0.153	0.203
BE	0.134	0.218	0.182	0.214	0.216	0.378	0.122	0.198	0.122	0.179
BG	0.389	0.385	0.382	0.378	0.415	0.374	0.341	0.334	0.306	0.314
BY	0.259	0.199	0.241	0.2	0.315	0.326	0.173	0.17	0.151	0.154
CH	0.148	0.226	0.182	0.212	0.252	0.421	0.143	0.203	0.127	0.187
CZ	0.173	0.293	0.216	0.287	0.223	0.352	0.162	0.287	0.148	0.264
DE	0.178	0.275	0.217	0.282	0.303	0.4	0.175	0.239	0.128	0.166
DK	0.336	0.26	0.348	0.237	0.405	0.395	0.327	0.242	0.287	0.218
EE	0.442	0.576	0.479	0.846	0.498	1.099	0.373	0.733	0.314	0.578
EL	0.2	0.164	0.273	0.172	0.351	0.391	0.183	0.159	0.186	0.162
ES	0.109	0.13	0.171	0.145	0.254	0.315	0.109	0.123	0.114	0.132
FI	0.208	0.327	0.222	0.307	0.285	0.469	0.191	0.308	0.186	0.26
FR	0.098	0.107	0.149	0.132	0.203	0.28	0.096	0.104	0.102	0.112
IT	0.114	0.119	0.136	0.123	0.225	0.326	0.106	0.104	0.127	0.13
LT	0.251	0.259	0.247	0.25	0.283	0.366	0.237	0.238	0.207	0.203
NL	0.109	0.148	0.136	0.15	0.173	0.32	0.107	0.137	0.105	0.139
NO	0.31	0.244	0.279	0.249	0.418	0.435	0.29	0.247	0.279	0.234
PL	0.141	0.164	0.227	0.214	0.171	0.243	0.132	0.15	0.138	0.15
PT	0.16	0.204	0.194	0.217	0.297	0.364	0.154	0.184	0.151	0.183
SE	0.109	0.179	0.145	0.191	0.204	0.386	0.113	0.17	0.126	0.152
SK	0.254	0.444	0.255	0.412	0.3	0.52	0.228	0.421	0.211	0.387

From the above table, it is noticed that all models give high RMSE error for both female and male for the country Estonia (EE). The models LC, PM_10 and PM_15

shows low error both female and male for the country France (FR). PM_5 gives low error for country Poland (PL) for both genders. While, GM model gives Netherlands (NL) low female error and Italy (IT) low male error.

5.2 Mean Relative Error (MRE)

This type of error is relative to the size of the item being measured. It is also be used to describe accuracy. Table 5.2 displays MRE errors for all models and countries, both female and male.

$$\text{MRE} = \frac{1}{N} \sum_{i=1}^N \frac{(\hat{m}_i - m_i)}{m_i}$$

Table 5.2: MRE Fitting error

	LC		GM		PM_5		PM_10		PM_15	
	F	M	F	M	F	M	F	M	F	M
AT	0.025	0.015	-0.101	0.198	0.075	0.096	0.035	0.032	0.026	0.016
BE	0.017	0.012	-0.109	0.168	0.056	0.081	0.025	0.025	0.018	0.013
BG	0.015	0.011	0.844	0.458	0.042	0.046	0.021	0.018	0.015	0.01
BY	0.016	0.016	0.465	0.530	0.039	0.048	0.017	0.018	0.011	0.008
CH	0.036	0.02	-0.262	-0.034	0.087	0.097	0.049	0.032	0.038	0.02
CZ	0.017	0.011	0.287	0.312	0.053	0.067	0.028	0.026	0.017	0.013
DE	0.003	0.004	-0.023	0.305	0.036	0.058	0.011	0.016	0.004	0.004
DK	0.022	0.029	0.136	0.332	0.049	0.093	0.03	0.042	0.021	0.03
EE	0.053	0.028	1.148	3.257	0.087	0.076	0.06	0.039	0.049	0.028
EL	0.025	0.014	0.492	0.087	0.066	0.089	0.037	0.034	0.025	0.015
ES	0.006	0.006	-0.083	-0.059	0.033	0.051	0.011	0.014	0.006	0.004
FI	0.039	0.031	-0.108	0.44	0.08	0.098	0.053	0.049	0.036	0.031
FR	0.003	0.003	-0.295	0.024	0.034	0.058	0.011	0.013	0.004	0.003
IT	0.005	0.005	-0.146	-0.015	0.037	0.065	0.014	0.017	0.006	0.005
LT	0.031	0.021	0.635	1.57	0.06	0.069	0.04	0.032	0.03	0.018
NL	0.011	0.007	-0.187	0.23	0.044	0.057	0.017	0.017	0.012	0.008
NO	0.031	0.031	0.154	0.067	0.061	0.116	0.038	0.048	0.031	0.031
PL	0.004	0.003	0.253	0.395	0.038	0.052	0.013	0.017	0.005	0.004
PT	0.015	0.01	0.029	-0.068	0.039	0.052	0.021	0.017	0.014	0.009
SE	0.029	0.021	-0.189	-0.027	0.074	0.093	0.038	0.33	0.028	0.02
SK	0.026	0.018	0.234	1.272	0.061	0.068	0.038	0.031	0.028	0.018

From the table 5.2, it is demonstrated that the models LC, GM and PM_15 give low MRE error in the country France (FR) and high error for Estonia (EE) for females. While PM_5 and PM_10 shows low error in the country Spain (ES) and high for Estonia. In terms for males, models LC, PM_10 and PM_15 give low error in France while GM model shows low for Portugal (PT) and PM_5 for Bulgaria (BG). For high male errors, LC and PM_15 shows in countries Finland (FI) and Norway (NO), PM_5 in Norway, PM_10 in Sweden (SE) and GM shows in Estonia.

5.3 Mean Absolute Error

MAE measures the average magnitude of the errors in a set of predictions, where all individual differences have equal weight. Table 5.3 displays MAE errors for all models and countries, both female and male.

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |\hat{m}_i - m_i|$$

Table 5.3: MAE Fitting error

	LC		GM		PM_5		PM_10		PM_15	
	F	M	F	M	F	M	F	M	F	M
AT	0.001	0.002	0.003	0.003	0.002	0.005	0.001	0.003	0.001	0.002
BE	0.001	0.002	0.003	0.004	0.002	0.005	0.001	0.002	0.001	0.002
BG	0.004	0.005	0.007	0.007	0.005	0.005	0.003	0.004	0.003	0.004
BY	0.003	0.004	0.004	0.004	0.003	0.006	0.002	0.003	0.002	0.002
CH	0.001	0.002	0.002	0.003	0.002	0.005	0.001	0.002	0.001	0.002
CZ	0.002	0.003	0.004	0.005	0.003	0.005	0.002	0.003	0.002	0.003
DE	0.001	0.003	0.003	0.004	0.002	0.005	0.001	0.002	0.001	0.002
DK	0.003	0.003	0.004	0.003	0.004	0.005	0.003	0.003	0.003	0.002
EE	0.004	0.008	0.005	0.016	0.004	0.012	0.003	0.009	0.003	0.008
EL	0.002	0.002	0.004	0.003	0.003	0.004	0.002	0.002	0.002	0.002
ES	0.001	0.001	0.002	0.002	0.002	0.004	0.001	0.001	0.001	0.001
FI	0.002	0.003	0.003	0.004	0.002	0.006	0.002	0.003	0.002	0.003
FR	0.001	0.001	0.002	0.002	0.002	0.004	0.001	0.001	0.001	0.001
IT	0.001	0.001	0.002	0.002	0.002	0.004	0.001	0.001	0.001	0.002
LT	0.003	0.004	0.004	0.005	0.003	0.005	0.002	0.003	0.002	0.003
NL	0.001	0.002	0.002	0.003	0.002	0.004	0.001	0.002	0.001	0.002
NO	0.003	0.003	0.004	0.004	0.004	0.006	0.003	0.003	0.003	0.003
PL	0.001	0.002	0.004	0.004	0.002	0.004	0.001	0.002	0.001	0.002
PT	0.001	0.002	0.003	0.003	0.003	0.005	0.001	0.002	0.001	0.002
SE	0.001	0.002	0.002	0.003	0.002	0.005	0.001	0.002	0.001	0.002
SK	0.003	0.005	0.004	0.006	0.003	0.007	0.002	0.005	0.002	0.005

From the above table, it is concluded that all models show high MAE error for males in the country Estonia. While models LC, PM_10 and PM_15 gives high female error in both Estonia and Bulgaria but GM and PM_5 gives only in Bulgaria.

5.4 Mean Absolute Relative Error (MARE)

It is a way to measure the performance of an estimation method in statistics, for example in trend estimation, also used as a loss function for regression. Table 5.4 displays MARE errors for each model with gender specific and for every country.



$$\text{MARE} = \frac{1}{N} \sum_{i=1}^N \left| \frac{\hat{m}_i - m_i}{m_i} \right|$$

Table 5.4: MARE Fitting error

	LC		GM		PM_5		PM_10		PM_15	
	F	M	F	M	F	M	F	M	F	M
AT	0.136	0.104	0.356	0.474	0.232	0.293	0.157	0.152	0.139	0.110
BE	0.115	0.097	0.355	0.395	0.199	0.277	0.132	0.129	0.114	0.095
BG	0.115	0.098	1.001	0.558	0.176	0.195	0.127	0.113	0.111	0.089
BY	0.128	0.137	0.585	0.583	0.177	0.217	0.116	0.111	0.097	0.078
CH	0.158	0.118	0.403	0.362	0.256	0.303	0.181	0.147	0.160	0.118
CZ	0.113	0.090	0.472	0.434	0.188	0.235	0.136	0.130	0.113	0.094
DE	0.057	0.062	0.348	0.476	0.16	0.227	0.085	0.107	0.061	0.06
DK	0.136	0.139	0.386	0.506	0.19	0.280	0.152	0.164	0.131	0.136
EE	0.237	0.163	1.273	3.425	0.292	0.267	0.249	0.186	0.221	0.158
EL	0.135	0.103	0.674	0.341	0.219	0.288	0.16	0.142	0.136	0.103
ES	0.074	0.072	0.427	0.262	0.154	0.218	0.085	0.091	0.068	0.062
FI	0.175	0.144	0.386	0.667	0.244	0.289	0.198	0.174	0.167	0.139
FR	0.057	0.053	0.437	0.379	0.156	0.233	0.081	0.09	0.057	0.051
IT	0.069	0.068	0.355	0.287	0.159	0.253	0.088	0.095	0.071	0.066
LT	0.159	0.139	0.852	1.675	0.211	0.247	0.175	0.149	0.154	0.114
NL	0.092	0.76	0.314	0.399	0.171	0.223	0.107	0.105	0.096	0.082
NO	0.161	0.148	0.457	0.347	0.224	0.329	0.177	0.179	0.161	0.148
PL	0.061	0.05	0.445	0.475	0.157	0.206	0.09	0.099	0.068	0.057
PT	0.106	0.095	0.391	0.277	0.166	0.216	0.123	0.105	0.104	0.083
SE	0.139	0.119	0.332	0.287	0.23	0.289	0.16	0.143	0.141	0.115
SK	0.144	0.114	0.448	1.369	0.207	0.231	0.17	0.146	0.151	0.112

From the table 5.4, it is demonstrated that models LC, PM_10 and PM_15 give low MARE error for France and high error for Estonia, in females. While GM and PM_5 models give high error for Estonia as well but low errors for Netherlands (GM model) and Spain (PM_5 model). For males, models PM_10 and PM_15 give low error for France and high for Estonia, while GM also display high error for Estonia but low for Spain. And finally, for PM_5 model displays low for Bulgaria and high for Norway.

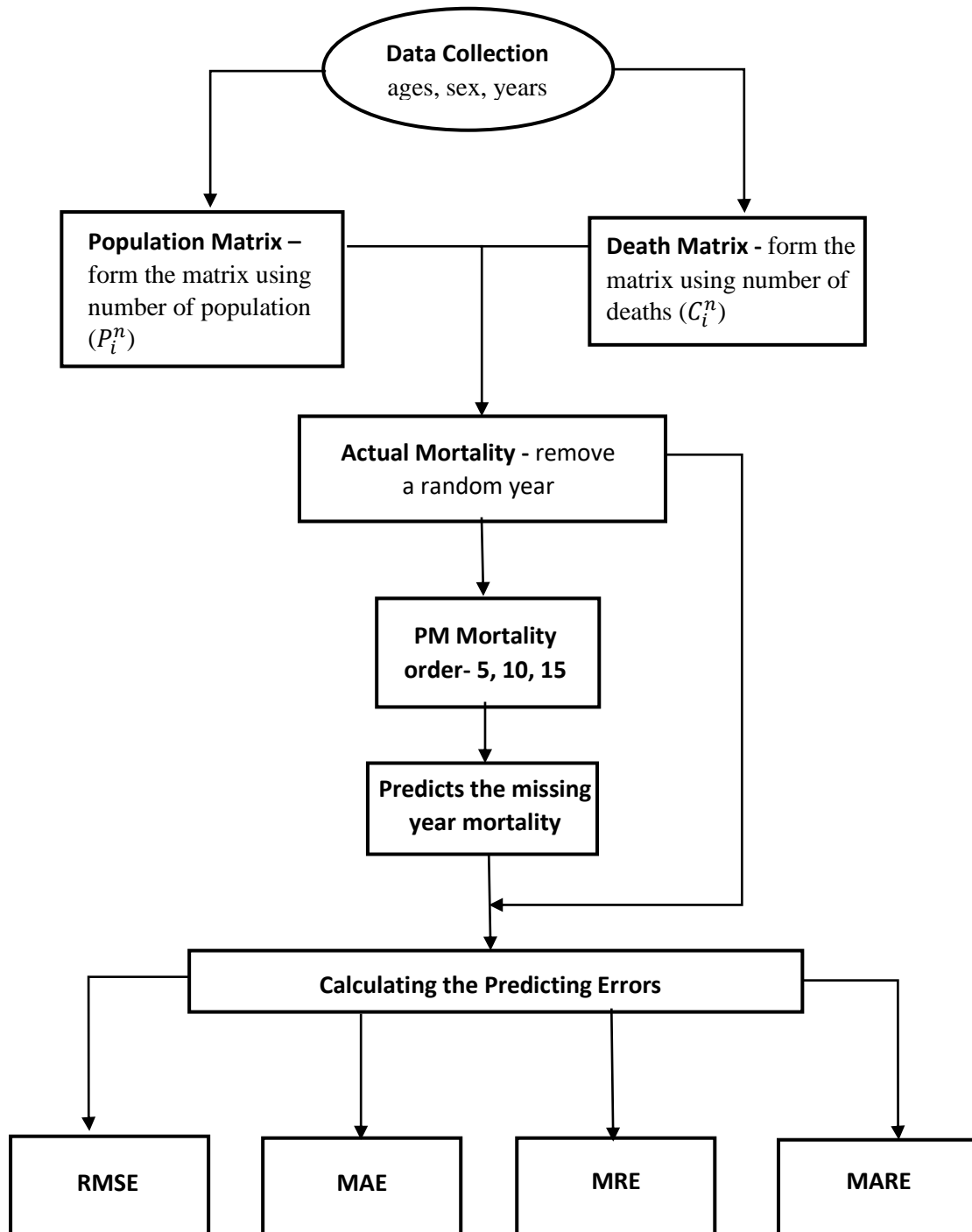
Table 5.5: Ranking of Fitting errors

	RMSE		MRE		MAE		MARE	
	F	M	F	M	F	M	F	M
LEAST   HIGH	PM_15		PM_15		PM_15		PM_15	
	PM_10		LC		PM_10		LC	PM_10
	LC		PM_10		LC		PM_10	LC
	GM		PM_5		PM_5	GM	PM_5	
	PM_5		GM		GM	PM_5	GM	

From all the above tables (5.1-5.4), it is observed that males show higher errors than females. MAE produces the least error compared to the other errors for all models. It is noticed from the table 5.5 that based on four errors, Polynomial model of order 15 gives least error while order-5 and GM model gives highest error.

6. Predicting Missing Mortality

6.1 Flowchart



The above flowchart demonstrates the steps for predicting missing year mortality for 8 countries. The first three steps follow the same procedure as mentioned for fitting mortality such as collection of data and organising it into death and population matrices. Then the actual mortality is calculated but with exception of removing one random year. The polynomial model estimates this missing mortality with different orders (5, 10, 15). And finally, errors (RMSE, MRE, MAE, MARE) are calculated in testing the accuracy of prediction.

6.2 Data Description

For the second part of the thesis, we are trying to evaluate the predictability of the new Polynomial model for the missing mortality values in the defined time period (1985-2014). In this case, we can randomly delete any one year. And then calculate four different errors as in the first part to test the accuracy in predicting. This part closely resonates to forecasting as the model is trying to calculate the incomplete mortality data.

Table 6.2 lists the countries used for prediction for each gender specific and missing years.

	F	M	F	M
BE	2012	2011	FR	2003
BG	2011	2003	IT	2011
DE	1989	1988	PL	2011
ES	2011	2008	PT	2011

6.3 Mortality Rate

In this section, we compare the predicting year mortality of polynomial model with the actual mortality. For this, we summarize the ages (0-99) to obtain mortality for the predicting year.

Table 6.3 illustrates the difference in mortality predicted by different orders of Polynomial model

	F				M			
	Actual	PM_5	PM_10	PM_15	Actual	PM_5	PM_10	PM_15
BE	2.575	2.464	2.493	2.438	3.248	3.477	3.5	3.426
BG	4.108	3.657	3.774	3.8	4.672	4.55	4.637	4.674
DE	3.522	3.332	3.302	3.398	4.273	4.296	4.295	4.422
ES	2.257	2.285	2.277	2.223	3.217	3.205	3.192	3.193
FR	2.568	2.332	2.333	2.365	2.976	3.064	3.099	3.05
IT	2.322	2.342	2.372	2.299	3.846	3.953	3.935	3.911
PL	2.811	2.925	2.915	2.92	3.901	3.827	3.859	3.891
PT	2.412	2.544	2.577	2.556	4.028	4.22	4.333	4.473

6.4 Measuring Predicting Errors

For this section, we look at four error types as mentioned previously for fitting to give insight about the precision of different orders of Polynomial model.

From below table 6.4.1, it is shown that PM_5 model gives high RMSE female error for Germany (DE) while PM_10 and PM_15 gives high error for Bulgaria. For low female error, PM_5 (Poland), PM_10 (Spain) and PM_15 (Italy). For low RMSE male errors, all three models show for Spain and for high error, PM_5 (Portugal), PM_10 (Belgium) and PM_15 (Germany).

Table 6.4.1: RMSE

	5		10		15	
	F	M	F	M	F	M
BE	0.184	0.226	0.125	0.275	0.167	0.264
BG	0.711	0.296	0.652	0.232	0.636	0.228
DE	0.817	0.346	0.649	0.218	0.613	0.4
ES	0.11	0.192	0.063	0.077	0.123	0.104
FR	0.335	0.243	0.264	0.124	0.241	0.12
IT	0.118	0.331	0.125	0.111	0.087	0.193
PL	0.104	0.243	0.116	0.108	0.144	0.138
PT	0.176	0.384	0.164	0.268	0.143	0.385

Table 6.4.2: MRE

	5		10		15	
	F	M	F	M	F	M
BE	0.008	0.048	-0.006	0.012	-0.014	-0.002
BG	0.012	0.039	-0.003	-0.011	0.022	-0.005
DE	0.064	0.088	0.04	0.004	0.061	-0.033
ES	0.019	0.033	0.001	0.007	-0.027	0.009
FR	-0.006	0.051	-0.026	0.032	-0.021	0.02
IT	0.026	0.121	0.018	0.048	-0.005	0.031
PL	0.048	0.019	0.007	0.018	-0.002	0.006
PT	0.054	0.119	0.013	0.082	-0.005	0.06

For low female MRE errors, both PM_5, PM_10 shows France and PM_15 (Spain). For high female errors, all models show Germany. For low male errors, PM_5 (Poland), PM_10 (Bulgaria) and PM_15 (Germany). And finally, for high male errors, PM_5 (Italy), both PM_10 and PM_15 shows Portugal.

Table 6.4.3: MAE

	5		10		15	
	F	M	F	M	F	M
BE	0.001	0.004	0.001	0.003	0.002	0.003
BG	0.007	0.005	0.007	0.004	0.007	0.003
DE	0.004	0.005	0.003	0.002	0.003	0.003
ES	0.001	0.002	0.001	0.001	0.001	0.001
FR	0.003	0.003	0.002	0.002	0.002	0.002
IT	0.001	0.005	0.001	0.001	0.001	0.002
PL	0.001	0.003	0.001	0.001	0.001	0.002
PT	0.002	0.006	0.002	0.004	0.002	0.005

For high female MAE errors, all three models show Bulgaria. And for high male errors, PM_5, PM_15 show Portugal while PM_10 shows Bulgaria and Portugal.

Table 6.4.4: MARE

	5		10		15	
	F	M	F	M	F	M
BE	0.182	0.218	0.136	0.111	0.132	0.096
BG	0.159	0.205	0.133	0.113	0.135	0.101
DE	0.221	0.306	0.145	0.165	0.116	0.093
ES	0.12	0.154	0.066	0.081	0.066	0.06
FR	0.146	0.186	0.093	0.071	0.069	0.041
IT	0.121	0.416	0.078	0.141	0.06	0.109
PL	0.145	0.167	0.079	0.087	0.072	0.059
PT	0.165	0.392	0.126	0.173	0.098	0.143

For low female MARE errors, both PM_5, PM_10 shows Spain and PM_15 (Italy). For high female errors, both PM_5, PM_10 show Germany and PM_15 (Bulgaria). For low male errors, PM_5 (Spain), both PM_10, PM_15 (France). And finally, for high male errors, PM_5 (Italy), both PM_10 and PM_15 shows Portugal.

In order to observe the difference in four errors, we have calculated Global values which summarizes values of all the countries for each error type with gender specific.

Figure 6.4.1: Global Predicting Errors- Females

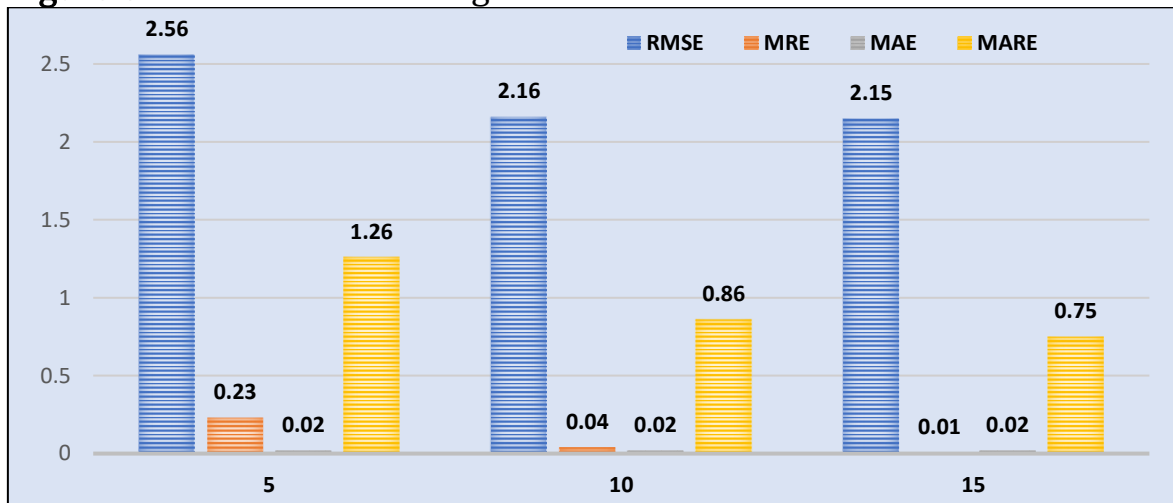
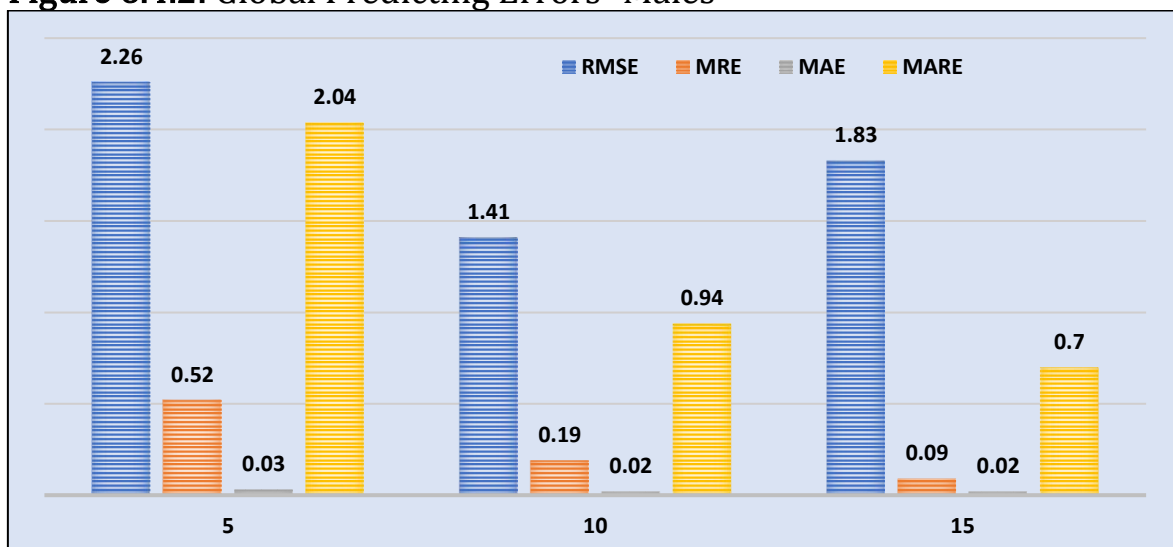


Figure 6.4.2: Global Predicting Errors- Males



It can be noticed that RMSE produces higher predicting error followed by MARE for both female and male. While MAE and MRE gives least errors. Polynomial model of order 15 gives least predicting errors for all types which is similar to fitting errors.

7. Conclusions

A considerable effort has been devoted in recent years to the understanding of complex systems such as Social network. These networks possess a complex community structure in which individuals typically belong to groups or communities, with a high density of internal connections and loosely connected among them, giving rise to a hierarchy of nested social ties. One of the defining features of statistical models for the analysis of social network is their ability to represent directly the underlying mechanisms generating dependence between network ties. Mortality statistics are a useful tool for assessment in Social relationships, Age composition which helps to build Social networks (family network, work network and friend network).

Mortality contributes significantly to population dynamics and is crucial in many fields such as economy, demography and social sciences. In this thesis, we look deeply into different mortality models that contribute towards fitting values to actual mortality.

One of the stochastic mortality models used is the Lee and Carter (1992) model is the most frequently used, due to some advantages like the small number of parameters compared to other models and the robustness. The second model used is Gompertz-Makeham model which is the pioneer of a new approach to survival modelling in 1825. The third, is a new model designed during this thesis called Polynomial model which is further divided into three polynomial orders of 5, 10 and 15. All models uses different parameters in order to produce mortality. Lee-Carter uses an age-dependent parameter, Gompertz-Makeham uses three parameters: one age-independent and two age-dependent. And finally, Polynomial model uses coefficients as parameters. The new model provides a trade-off between the number of parameters require to produce accurate mortality which means as the order of the polynomial increases, the parameters involved increases and produces closely to Actual mortality.

In this paper, 21 countries in Europe have been fitted mortality using these models. All the models have shown same mortality trends: Bulgaria shows the highest female mortality with Estonia the least, and Greece shows the least male mortality with Estonia the highest. It can also be noticed from all the tables shown in the report that males demonstrate higher mortality than females.

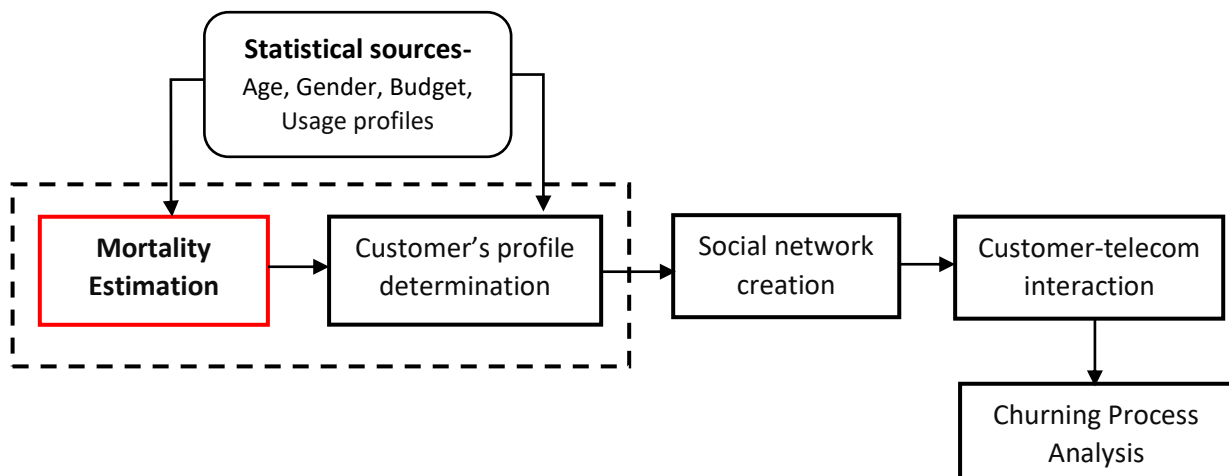
In order to test the precision of these fitting models, we have calculated four different errors (RMSE, MAE, MRE and MARE) for each model. It is observed that MAE is the gives the least error for all models. And also, the Polynomial model of order 15 shows the best results for all the error types. With Gompertz-Makeham the worst model followed by Polynomial of order 5.

For the second part of the thesis, we dive deeper to study the new Polynomial model by predicting the mortality of a random missing year from ages 0-99. We calculated the same error types to study the difference in mortality produced by different orders of polynomial. It is seen that as the order of the polynomial increases the prediction of mortality gets better. Hence, PM_15 is best model for fitting and predicting.

8. Future Work

- 8.1 Based on the second part of thesis, in terms of predicting the missing mortality. We have demonstrated the goodness of new Polynomial model by predicting the missing mortality of one random year for all the ages from 0-99. In the future study case, we can actually remove more than one year which could be either consecutive or randomly picked years. And can also be applied in terms of age which could predict for any randomly missing specific age or an age-group. This method helps to cope for any incomplete statistical data tables obtained from Eurostat or from any source.
- 8.2 There were 6.8 billion estimates for mobile subscriptions worldwide by end of 2013. As the mobile market gets saturated, it becomes harder for telecom providers to acquire new customers, and makes it essential for them to retain their own. Due to the high competition between different telecom providers and the ability of customers to move from one provider to another, all telecom service providers suffer from customer churn. As a result, churn prediction has become one of the main telecom challenges. The primary goal of churn prediction is to predict a list of potential churners, so that telecom providers can start targeting them by retention campaigns. This work describes work in progress in which modelling for churn as a dynamic social behaviour, where customer churn propagates in the telecom network over strong social ties.

Figure 8: Mobile Market Churning Process



Statistical Sources: Collect life tables with data such as age, population size, gender.

Mortality Modelling: Use stochastic models to estimate mortality.

Customer's Profile: define customers by their personal profile or characteristics and usage profiles, budget, number of devices.

Social Network Creation: takes into account nodal attributes such as strong and loose ties, homophily (the tendency of individuals to associate with others of same kind).

Customer-telecom interaction: The customer's behaviour determines the interaction between customers and telecom providers in a mobile market.

Churning Process: the analysis of churning process considers the user profiles of customers, satisfaction with the service and price.

The creation of social network is based on customers' profile. We accommodate customer's profile information based on statistical sources and establish a mechanism to form ties between them based on affinity. Thus, ties are more probable between similar people. Customers will decide which telecom provider to choose based on the information shared by their friends and the features of the mobile service. The churning process is based on the satisfaction with the provider.

Bibliography

1. Amati, Viviana and Lomi, Alessandro and Mira, Antonietta, Social Network Modeling (March 2018). Annual Review of Statistics and Its Application, Vol. 5, Issue 1, pp. 343-369, 2018
2. Tom A.B. Snijders, Statistical Models for Social Networks. Annual Review of Sociology August 2011; 37:1, 131-153
3. https://en.wikipedia.org/wiki/Lee%E2%80%93Carter_model
4. Chavhan, R.N. and Shinde, R.L., 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), pp.51–68
5. Missov TI, Lenart A. Gompertz-Makeham life expectancies: expressions and applications. Theor Popul Biol. 2013 Dec; 90:29-35
6. Marie Angèle Cathleen Alijean & Jason Narsoo, 2018. "Evaluation of the Kou-Modified Lee-Carter Model in Mortality Forecasting: Evidence from French Male Mortality Data," Risks, MDPI, Open Access Journal, vol. 6(4), pages 1-26, October.
7. https://www.wikiwand.com/en/Gompertz%E2%80%93Makeham_law_of_mortality
8. John Bongaarts, 2005. "Long-range trends in adult mortality: Models and projection methods," Demography, Springer; Population Association of America (PAA), vol. 42(1), pages 23-49, February.
9. Booth, H. & Tickle, L., 2008. "Mortality Modelling and Forecasting: A Review of Methods," Annals of Actuarial Science, Cambridge University Press, vol. 3(1-2), pages 3-43, September.
10. Postigo-Boix, Marcos & Melús-Moreno, José L., 2018. "A social model based on customers' profiles for analyzing the churning process in the mobile market of data plans," Physica A: Statistical Mechanics and its Applications, Elsevier, vol. 496(C), pages 571-592.

Appendices

We present Matlab Code used for our thesis. In part-1 we have fitting mortality models (LC; GM and PM) and also errors calculation (RMSE, MRE, MAE and MARE).

Part – 1

The datasets

We use Eurostat or HMD to obtain data for the models.

```

debug_level = 0; % 0, 1 or 2
Eurostat_dir = "Eurostat Data";
Error_type = 'mortality'; % Types: 'mortality' or 'deaths'
Error_ref = 'Eurostat'; % Types: 'Eurostat' or 'Statistical'
Poly_order_max = 15;
results_dir = "C:\Users\cash\Documents\MATLAB\Results\ES";
results_dir = sprintf("%s_%s_%s", results_dir, Error_type, Error_ref);
if ~exist(results_dir, 'dir')
    mkdir(results_dir);
end
Country_table = readtable("geo.txt");
Country_map = containers.Map(Country_table.EUR, Country_table.Europe);
%Countries =
{'BE', 'BG', 'BY', 'CH', 'CZ', 'DK', 'DE', 'EE', 'EL', 'ES', 'FR', 'IT', 'LT', 'NL',
'AT', 'SK', 'FI', 'SE', 'NO', 'PL', 'PT'}
Countries = {'ES'};
Sex_code = {'M'}; % M: male or F: female
%Sex_code = {'F'}; % M: male or F: female

View_graphs = true;
Save_files = true;

RMSE_global_LC = zeros(1, Poly_order_max);
MRE_global_LC = zeros(1, Poly_order_max);
MARE_global_LC = zeros(1, Poly_order_max);
MAE_global_LC = zeros(1, Poly_order_max);
RMSE_global_LC2 = zeros(1, Poly_order_max);
MRE_global_LC2 = zeros(1, Poly_order_max);
MARE_global_LC2 = zeros(1, Poly_order_max);
MAE_global_LC2 = zeros(1, Poly_order_max);
RMSE_global_GM = zeros(1, Poly_order_max);
MRE_global_GM = zeros(1, Poly_order_max);
MARE_global_GM = zeros(1, Poly_order_max);
MAE_global_GM = zeros(1, Poly_order_max);
RMSE_global_Poly = zeros(1, Poly_order_max);
MRE_global_Poly = zeros(1, Poly_order_max);
MARE_global_Poly = zeros(1, Poly_order_max);
MAE_global_Poly = zeros(1, Poly_order_max);

yearStart = 1985;
yearEnd = 2014;

```

```
years = (yearStart:yearEnd);

for country_i = 1:length(Countries)
    for sex_i = 1:length(Sex_code)

        MinimumAge = 0;
        MaximumAge = 99;
        % ages = (MinimumAge:MaximumAge);
        sex_EuStat = Sex_code{sex_i}; % M: male or F: female
        if (sex_EuStat=='M')
            sex = 'male';
        else
            sex = 'female';
        end
        country_EuStat = Countries{country_i}; % Eurostat Country Code
        country = Country_map(country_EuStat); % Corresponding Country Name (for
figures)
```

The datasets are the following:

```
file_mortality = dir(sprintf("%s\\%s_mortality_%s*", Eurostat_dir, country_EuStat,
sex_EuStat));
file_population = dir(sprintf("%s\\%s_population_%s*", Eurostat_dir,
country_EuStat, sex_EuStat));
segments_m = split (file_mortality.name,{'_', '.'});
yearStart_m = str2double(cell2mat(segments_m(4)));
yearEnd_m = str2double(cell2mat(segments_m(5)));
segments_p = split (file_population.name,{'_', '.'});
yearStart_p = str2double(cell2mat(segments_p(4)));
yearEnd_p = str2double(cell2mat(segments_p(5)));
MinimumAge_m = MinimumAge;
MaximumAge_m = MaximumAge;
MinimumAge_p = MinimumAge;
MaximumAge_p = MaximumAge;
```

Death Matrix

```
C = dlmread(sprintf("%s\\%s",Eurostat_dir,file_mortality.name)); % Number of
deaths -- Dimensions: ...
% {(MaximumAge - MinimumAge + 1), (yearEnd - yearStart +1)}
figure;
surf(yearStart_m:yearEnd_m,MinimumAge:MaximumAge,C);
title(strcat("Number of ", sex, "'s deaths in ",country, " (Eurostat)"))
xlabel('year')
ylabel('age')
zlabel('c')

% if necessary trim the matrix
i=1;
while any(isnan(C(:,i))) % Trim left side (if missing values)
    C(:,i) = [];
    i = i + 1;
    yearStart_m = yearStart_m + 1;
end
```

```

i=size(C,2);
while any(isnan(C(:,i))) % Trim right side (if missing values)
    C(:,i) = [];
    i = i - 1;
    yearEnd_m = yearEnd_m - 1;
end
i=1;
while any(isnan(C(i,:))) % Trim top side (if missing values)
    C(i,:) = [];
    i = i + 1;
    MinimumAge_m = MinimumAge_m + 1;
end
i=size(C,1);
while any(isnan(C(i,:))) % Trim bottom side (if missing values)
    C(i,:) = [];
    i = i - 1;
    MaximumAge_m = MaximumAge_m - 1;
end
if any(isnan(C))
    fprintf("Matrix C not valid (NaN values)\n");
    exit;
end

```

Population Matrix

```

P = dlmread(sprintf("%s\\%s",Eurostat_dir,file_population.name)); % Population --
Dimensions: ...
% { (MaximumAge - MinimumAge + 1), (yearEnd - yearStart +2) }

figure;
surf(yearStart_p:yearEnd_p,MinimumAge:MaximumAge,P);
title(strcat(regexprep(sex, '\<[a-z]', '${upper($1)}'), "'s Population in
",country," (Eurostat)"))
xlabel('year')
ylabel('age')
zlabel('p')

% if necessary trim the matrix
i=1;
while any(isnan(P(:,i))) % Trim left side (if missing values)
    P(:,i) = [];
    yearStart_p = yearStart_p + 1;
end
i=size(P,2);
while any(isnan(P(:,i))) % Trim right side (if missing values)
    i=size(P,2);
    P(:,i) = [];
    yearStart_p = yearStart_p - 1;
end
i=1;
while any(isnan(P(i,:))) % Trim top side (if missing values)
    P(i,:) = [];
    MinimumAge_p = MinimumAge_p + 1;
end

```

```

i=size(P,1);
while any(isnan(P(i,:))) % Trim bottom side (if missing values)
    i=size(P,1);
    P(i,:) = [];
    i = i - 1;
    MaximumAge_p = MaximumAge_p - 1;
end
if any(isnan(P))
    fprintf("Matrix C not valid (NaN values)\n");
    exit;
end

```

Number of deaths and Population matrices for the models

```

yearStart = max([yearStart_m,yearStart_p]);
yearEnd = min([yearEnd_m,yearEnd_p]);
if (yearEnd == yearEnd_p)
    yearEnd = yearEnd - 1; % Since we need one more year to calculate the mortality
end
MinimumAge = max([MinimumAge_m,MinimumAge_p]);
MaximumAge = min([MaximumAge_m,MaximumAge_p]);

MAE_LC_years_val = zeros(Poly_order_max,yearEnd-yearStart+1);
MAE_GM_years_val = zeros(Poly_order_max,yearEnd-yearStart+1);
MAE_Poly_years_val = zeros(Poly_order_max,yearEnd-yearStart+1);
MAE_LC2_ind_years_val = zeros(Poly_order_max,yearEnd-yearStart+1);
MAE_LC_age_val = zeros(Poly_order_max,MaximumAge-MinimumAge+1);
MAE_GM_age_val = zeros(Poly_order_max,MaximumAge-MinimumAge+1);
MAE_Poly_age_val = zeros(Poly_order_max,MaximumAge-MinimumAge+1);
MAE_LC2_ind_age_val = zeros(Poly_order_max,MaximumAge-MinimumAge+1);
MARE_LC_years_val = zeros(Poly_order_max,yearEnd-yearStart+1);
MARE_GM_years_val = zeros(Poly_order_max,yearEnd-yearStart+1);
MARE_Poly_years_val = zeros(Poly_order_max,yearEnd-yearStart+1);
MARE_LC2_ind_years_val = zeros(Poly_order_max,yearEnd-yearStart+1);
MARE_LC_age_val = zeros(Poly_order_max,MaximumAge-MinimumAge+1);
MARE_GM_age_val = zeros(Poly_order_max,MaximumAge-MinimumAge+1);
MARE_Poly_age_val = zeros(Poly_order_max,MaximumAge-MinimumAge+1);
MARE_LC2_ind_age_val = zeros(Poly_order_max,MaximumAge-MinimumAge+1);
ACR_LC_years = zeros(Poly_order_max,yearEnd-yearStart+1);
ACR_GM_years = zeros(Poly_order_max,yearEnd-yearStart+1);
ACR_LC2_years = zeros(Poly_order_max,yearEnd-yearStart+1);
ACR_Poly_years = zeros(Poly_order_max,yearEnd-yearStart+1);
ACR_LC_ages = zeros(Poly_order_max,MaximumAge-MinimumAge+1);
ACR_GM_ages = zeros(Poly_order_max,MaximumAge-MinimumAge+1);
ACR_LC2_ages = zeros(Poly_order_max,MaximumAge-MinimumAge+1);
ACR_Poly_ages = zeros(Poly_order_max,MaximumAge-MinimumAge+1);

```

Mortality Matrix

```
T = C ./ (P(:,1:(size(P,2)-1)) + P(:,2:size(P,2))/2); % Mortality rate
h = figure;
if ~View_graphs
    h.Visible = 'off';
end
surf(yearStart:yearEnd,MinimumAge:MaximumAge,T);
title(strcat("Mortality rate for ", sex, "s in ",country," (Eurostat)"))
xlabel('year')
ylabel('age')
zlabel('t')

Full_T = T; % This is the full T matrix
column2remove = randi([2,size(T,2)-1]);
T(:,column2remove) = []; % Now, we remove one column (year) randomly.
Full_years = years;
years(:,column2remove) = [];
```

Lee-Carter model

```
[LC_Mod.axhat,LC_Mod.bxhat,LC_Mod.kthat,LC_Mod.That] = LC1(T);

h = figure;
if ~View_graphs
    h.Visible = 'off';
end
surf (years,ages,LC_Mod.That)
hold on;
[Years, Ages, T_scat] = prepareSurfaceData(years, ages', T);
scatter3 (Years,Ages,T_scat,'red')
title(["Lee-Carter Model";strcat("(" ,country,")")]
xlabel('year')
ylabel('age')
zlabel(strcat(regexprep(sex,'(\[a-z])','${upper($1)}'), "'s Mortality"))

% Plot estimated number of deaths
LC_Mod_Deaths = LC_Mod.That.*(P(:,1:(size(P,2)-1)) + P(:,2:size(P,2))/2);
h = figure;
if ~View_graphs
    h.Visible = 'off';
end
surf (years,ages,LC_Mod_Deaths)
hold on;
[Years, Ages, C_scat] = prepareSurfaceData(years, ages', C);
scatter3 (Years,Ages,C_scat,'red')
title(["Lee-Carter Model";strcat("(" ,country,")")]
xlabel('year')
ylabel('age')
zlabel(strcat(regexprep(sex,'(\[a-z])','${upper($1)}'), "'s Deaths"))
```


Gompertz-Makeham model

```
GM_Mod = GM_model_exp_poly (ages,years,T,nan, nan, nan);

h = figure;
if ~View_graphs
    h.Visible = 'off';
end
surf (years,ages,GM_Mod.That)
hold on;
scatter3 (Years,Ages,T_scat,'red')
title(["Gompertz-Makeham Model";strcat("(",regexprep(sex,'(\<[a-
z])','${upper($1)}'),'s/',country,"")"])
xlabel('year')
ylabel('age')
zlabel("Mortality")

% Plot number of deaths
GM_Mod_Deaths = GM_Mod.That.*(P(:,1:(size(P,2)-1)) + P(:,2:size(P,2)))/2);
h = figure;
if ~View_graphs
    h.Visible = 'off';
end
surf (years,ages,GM_Mod_Deaths)
hold on;
scatter3 (Years,Ages,C_scat,'red')
title(["Gompertz-Makeham Model";strcat("(",regexprep(sex,'(\<[a-
z])','${upper($1)}'),'s/',country,"")"])
xlabel('year')
ylabel('age')
zlabel("Deaths")
```

The polynomial model

```
StartPoint_val = 0.01;
Lower_val = 0;
Upper_val = 1;
for poly_order = 1:Poly_order_max
    clearvars Pol_Mod;
% Approximate mortality for each year with a polynomial
for i=1:size(years,2)
    method = sprintf('poly%d',poly_order);
    method = char(method);
    Polfit = Poly_fit(ages',log(T(:,i)),poly_order);
    Pol_Mod.p(i,:) = Polfit(:)';
end
% Approximate the coefficients with a polynomial
for i=1:poly_order+1
    p_a = sprintf('p%d',i);
    p_a = char(p_a);
    Polfit = Poly_fit(years',Pol_Mod.p(:,i),poly_order);
    Pol_Mod.q(i,:) = Polfit(:)';
end
```

```

for y=1:size(years,2)
    for a=1:size(ages,2)
        vect_a=zeros(1,poly_order+1);
        vect_y=zeros(1,poly_order+1);
        vect_a(1)=1;
        vect_y(1)=1;
        norm_ages = normalize(ages);
        norm_years = normalize(years);
        for i=2:poly_order+1
            vect_a(i)=vect_a(i-1)*(norm_ages(a));
            vect_y(i)=vect_y(i-1)*(norm_years(y));
        end
        Pol_Mod.That(a,y) = exp(vect_a*Pol_Mod.q*vect_y');
    end
end
end
% Plot mortality
h = figure;
if ~View_graphs
    h.Visible = 'off';
end
surf (years,ages,Pol_Mod.That)
hold on;
scatter3 (Years,Ages,T_scat,'red')
title_str = sprintf("Polynomial Model (order = %d)",poly_order);
title([title_str;strcat("(",regexp(sex,'\<[a-
z])','${upper($1)}'),"s/",country,"")])
xlabel('year')
ylabel('age')
zlabel("Mortality")

% Plot Number of deaths
Poly_Mod_Deaths = Pol_Mod.That.*(P(:,1:(size(P,2)-1)) + P(:,2:size(P,2)))/2);
h = figure;
if ~View_graphs
    h.Visible = 'off';
end
surf (years,ages,Poly_Mod_Deaths)
hold on;
scatter3 (Years,Ages,C_scat,'red')
title_str = sprintf("Polynomial Model (order = %d)",poly_order);
title([title_str;strcat("(",regexp(sex,'\<[a-
z])','${upper($1)}'),"s/",country,"")])
xlabel('year')
ylabel('age')
zlabel("Deaths")

```

Measuring the Error

The Error

```
switch (Error_type)
  case 'mortality' % Error respect to the mortality (statistical data)
    switch (Error_ref)
      case 'Statistical'
        V_LC=LC_Mod.That;
        V_GM=GM_Mod.That;
      case 'Eurostat'
        V_LC=T;
        V_GM=T;
    end
    Error_LC = (LC_Mod.That-V_LC);
    Error_GM = (GM_Mod.That-V_GM);
    Error_Poly = (Pol_Mod.That-V_GM);

  case 'deaths' % Error respect to the deaths (statistical data)
    switch (Error_ref)
      case 'Statistical'
        V_LC=LC_Mod.That.*(P(:,1:(size(P,2)-1))+ P(:,2:size(P,2))));
        V_GM=GM_Mod.That.*(P(:,1:(size(P,2)-1))+ P(:,2:size(P,2))));
      case 'Eurostat'
        V_LC=C;
        V_GM=C;
    end
    Error_LC = (LC_Mod.That.*(P(:,1:(size(P,2)-1)) + P(:,2:size(P,2)))-
V_LC);
    Error_GM = (GM_Mod.That.*(P(:,1:(size(P,2)-1)) + P(:,2:size(P,2)))-
V_GM);
    Error_Poly = (Pol_Mod.That.*(P(:,1:(size(P,2)-1)) + P(:,2:size(P,2)))-
V_GM);
end
```

Squared Errors (SE)

```
S_error_LC = Error_LC.^2;
S_error_GM = Error_GM.^2;
S_error_Poly = Error_Poly.^2;
```

Relative Errors

```
Rel_error_LC = Error_LC./V_LC;
Rel_error_GM = Error_GM./V_GM;
Rel_error_Poly = Error_Poly./V_GM;
```

```
[X_coord,Y_coord] = meshgrid(years,ages);
```

```
% Plot Relative Errors
```

```
h = figure;
if ~View_graphs
  h.Visible = 'off';
end
levels=[0.1 0.2 1];
contourf(X_coord,Y_coord, Rel_error_LC,levels, 'ShowText', 'on')
```

```

title(["Relative Error (LC)";strcat("(" ,regexprep(sex, '\<[a-
z]'),' ${upper($1)}'),'s/','country,")")]
xlabel('age')
ylabel("years")

h = figure;
if ~View_graphs
    h.Visible = 'off';
end
contourf(X_coord,Y_coord, Rel_error_GM,levels, 'ShowText', 'on')
title(["Relative Error (GM)";strcat("(" ,regexprep(sex, '\<[a-
z]'),' ${upper($1)}'),'s/','country,")")]
xlabel('age')
ylabel("years")
h = figure;
if ~View_graphs
    h.Visible = 'off';
end
contourf(X_coord,Y_coord, Rel_error_Poly,levels, 'ShowText', 'on')
title_str = sprintf("Relative Error (Poly) (order = %d)",poly_order);
title([title_str;strcat("(" ,regexprep(sex, '\<[a-
z]'),' ${upper($1)}'),'s/','country,")")]
xlabel('age')
ylabel("years")

```

Absolute Errors

```

Abs_error_LC = abs(Error_LC);
Abs_error_GM = abs(Error_GM);
Abs_error_Poly = abs(Error_Poly);

% Plot Absolute Errors
h = figure;
if ~View_graphs
    h.Visible = 'off';
end
levels=[0.003 0.03 0.1 0.2 0.3];
contourf(X_coord,Y_coord, Abs_error_LC,levels, 'ShowText', 'on')
title(["Absolute Error (LC)";strcat("(" ,regexprep(sex, '\<[a-
z]'),' ${upper($1)}'),'s/','country,")")]
xlabel('age')
ylabel("years")
h = figure;
if ~View_graphs
    h.Visible = 'off';
end
contourf(X_coord,Y_coord, Abs_error_GM,levels, 'ShowText', 'on')
title(["Absolute Error (GM)";strcat("(" ,regexprep(sex, '\<[a-
z]'),' ${upper($1)}'),'s/','country,")")]
xlabel('age')
ylabel("years")

h = figure;
if ~View_graphs

```

```

    h.Visible = 'off';
end
contourf(X_coord,Y_coord, Abs_error_Poly,levels, 'ShowText', 'on')
title_str = sprintf("Absolute Error (Poly) (order = %d)",poly_order);
title([title_str;strcat("(",regexp(sex, '\<[a-
z]'),'${upper($1)}'),"s/",country,")")]
xlabel('age')
ylabel("years")

```

Summary of Results

Global RMSE (Root Mean Square Error)

```

RMSE_LC = sqrt(sum(sum(S_error_LC))/numel(S_error_LC))/mean(V_LC(:));
RMSE_GM = sqrt(sum(sum(S_error_GM))/numel(S_error_GM))/mean(V_GM(:));
RMSE_Poly = sqrt(sum(sum(S_error_Poly))/numel(S_error_Poly))/mean(V_GM(:));

```

Global MAE (Mean Absolute Error)

```

MAE_LC = sum(sum(sqrt(S_error_LC)))/numel(S_error_LC);
MAE_GM = sum(sum(sqrt(S_error_GM)))/numel(S_error_GM);
MAE_Poly = sum(sum(sqrt(S_error_Poly)))/numel(S_error_Poly);

```

Global MRE (Mean Relative Error)

```

MRE_LC = sum(sum(Rel_error_LC))/numel(Rel_error_LC);
MRE_GM = sum(sum(Rel_error_GM))/numel(Rel_error_GM);
MRE_Poly = sum(sum(Rel_error_Poly))/numel(Rel_error_Poly);

```

Global MARE (Mean Absolute Relative Error)

```

MARE_LC = sum(sum(abs(Rel_error_LC)))/numel(Rel_error_LC);
MARE_GM = sum(sum(abs(Rel_error_GM)))/numel(Rel_error_GM);
MARE_Poly = sum(sum(abs(Rel_error_Poly)))/numel(Rel_error_Poly);

```

RMSE_Global vs Poly order

```

h = figure;
if ~View_graphs
    h.Visible = 'off';
end
plot(1:Poly_order_max,log10(RMSE_global_LC));
hold on
plot(1:Poly_order_max,log10(RMSE_global_GM));
plot(1:Poly_order_max,log10(RMSE_global_Poly));
legend('LC','GM','Poly','location','best');
title(["RMSE Global vs. Poly order";strcat("(",regexp(sex, '\<[a-
z]'),'${upper($1)}'),"s/",country,")")]
xlabel('order')
ylabel('log(RMSE)')

```

MAE_Global vs Poly order

```

h = figure;
if ~View_graphs
    h.Visible = 'off';
end
plot(1:Poly_order_max,log10(MAE_global_LC));
hold on

```

```
plot(1:Poly_order_max,log10(MAE_global_GM));
plot(1:Poly_order_max,log10(MAE_global_Poly));
legend ('LC','GM','Poly','location','best');
title(["Global MAE vs. Polynomial degree";strcat("(" ,regexprep(sex, '\\<[a-
z]'),' ${upper($1)}'),"s/",country,"")])
xlabel('degree')
ylabel('log(MAE)')
```

MRE_Global vs Poly order

```
h = figure;
if ~View_graphs
    h.Visible = 'off';
end
plot(1:Poly_order_max,log10(MRE_global_LC));
hold on
plot(1:Poly_order_max,log10(MRE_global_GM));
plot(1:Poly_order_max,log10(MRE_global_Poly));
legend ('LC','GM','Poly','location','best');
title(["MRE Global vs. Poly order";strcat("(" ,regexprep(sex, '\\<[a-
z]'),' ${upper($1)}'),"s/",country,"")])
xlabel('order')
ylabel('log(MRE)')
```

MARE_Global vs Poly order

```
h = figure;
if ~View_graphs
    h.Visible = 'off';
end
plot(1:Poly_order_max,log10(MARE_global_LC));
hold on
plot(1:Poly_order_max,log10(MARE_global_GM));
plot(1:Poly_order_max,log10(MARE_global_Poly));
legend ('LC','GM','Poly','location','best');
title(["MARE Global vs. Poly order";strcat("(" ,regexprep(sex, '\\<[a-
z]'),' ${upper($1)}'),"s/",country,"")])
xlabel('order')
ylabel('log(MARE)')
end
end
```

Part – 2: we have code of Polynomial model for predicting missing years mortality and errors calculation (RMSE, MAE, MRE and MARE)

Error Estimation for missing years

```
Error_fitting = Pol_Mod.That(:,column2remove) - Full_T(:,column2remove);

% Plot Error
h = figure;
if ~View_graphs
    h.Visible = 'off';
end
plot (ages,Error_fitting)
title_str = sprintf("Polynomial Model Fitting Error (order = %d, year =
%d)",poly_order, Full_years(column2remove));
```

```
title([title_str;strcat("(",regexp(sex, '\<[a-
z]'),'${upper($1)}'),"s/",country,"")])
xlabel('age')
ylabel("Error")
```

Squared Errors (SE)

The square errors:

```
S_error_Poly = Error_fitting.^2;

% Plot Squared Error
h = figure;
if ~View_graphs
    h.Visible = 'off';
end
plot (ages,Error_fitting)
title_str = sprintf("Polynomial Model Fitting Squared Error (order = %d, year =
%d)",poly_order, Full_years(column2remove));
title([title_str;strcat("(",regexp(sex, '\<[a-
z]'),'${upper($1)}'),"s/",country,"")])
xlabel('age')
ylabel("Error")
```

Relative Errors

```
Rel_error_Poly = Error_fitting./Full_T;
```

Absolute Errors

```
Abs_error_Poly = abs(Error_fitting);

% Plot Absolute Error
h = figure;
if ~View_graphs
    h.Visible = 'off';
end
plot (ages,Error_fitting)
title_str = sprintf("Polynomial Model Fitting Absolute Error (order = %d, year =
%d)",poly_order, Full_years(column2remove));
title([title_str;strcat("(",regexp(sex, '\<[a-
z]'),'${upper($1)}'),"s/",country,"")])
xlabel('age')
ylabel("Error")
```

Summary of Results

Global RMSE (Root Mean Square Error)

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^N (\hat{m}_i - m_i)^2}{N}}$$

```
RMSE_Poly = sqrt(sum(sum(S_error_Poly))/numel(S_error_Poly))/mean(Full_T(:));
```

Global MAE (Mean Absolute Error)

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |\hat{m}_i - m_i|$$

```
MAE_Poly = sum(sum(sqrt(S_error_Poly)))/numel(S_error_Poly);
```

Global MRE (Mean Relative Error)

$$\text{MRE} = \frac{1}{N} \sum_{i=1}^N \frac{(\hat{m}_i - m_i)}{m_i}$$

```
MRE_Poly = sum(sum(Rel_error_Poly))/numel(Rel_error_Poly);
```

Global MARE (Mean Absolute Relative Error)

$$\text{MARE} = \frac{1}{N} \sum_{i=1}^N \left| \frac{(\hat{m}_i - m_i)}{m_i} \right|$$

```
MARE_Poly = sum(sum(abs(Rel_error_Poly)))/numel(Rel_error_Poly);
```