

A comparison of Poisson or Negative Binomial
Regression and Lee-Carter Models of
forecasting Norwegian male mortality

by

Coffie Emmanuel



THESIS

for the degree of

MASTER OF SCIENCE

(Modelling And Data Analysis)

Faculty of Mathematics and Natural Sciences

University of Oslo

May 26, 2015

A comparison of Poisson or Negative Binomial Regression and Lee-Carter Models of forecasting Norwegian male mortality

Coffie Emmanuel

May,2015

Abstract

Mortality rates are the ratio of death counts and estimates of the population exposed to risk of deaths in matched intervals of time and age. Mortality forecasting is the use of historical data of mortality to determine future pattern of trends. Mortality forecasting remains a major challenge for the planning of pension systems and management of annuity businesses. Having high mortality forecasting accuracy requires a model that provides good fit to the historical mortality data for consistent predictive performance. Poisson or Negative binomial regression and Lee-Carter (1992) demographic models are some of the predictive models being consistently used in mortality forecasting. Hence this thesis seeks to compare and evaluate forecasting accuracy of these two models in predicting Norwegian male mortality.

Keywords: mortality rates, poisson, negative binomial, orthogonal polynomials, lee-carter, random walk, mortality index, forecast, forecast errors

Acknowledgment

I am deeply grateful to the Almighty God for granting me strength and good health to successfully complete this thesis.

I am highly indebted to the Statistics and Mathematics Departments, most especially to all the intellectual lecturers of the Modeling and Data Analysis Master programme. I would also like to extend my sincere appreciation and gratitude to my supervisor Professor Anders Ryghs Swensen for his countless support and guidance.

Finally, I am thankful to my family and friends for their immense support and encouragement in my academic pursuance, most especially Prisca Mavhinga.

Contents

Abstract	1
Acknowledgment	2
Thesis Overview	5
1 The Data	6
1.1 Description of Data	6
1.1.1 Sources of the data	6
1.1.2 Lexis Diagram	6
1.1.3 Male mortality data	7
1.1.4 Risks of exposure data	9
1.1.5 Estimation of observed mortality Rates	11
2 Poisson and Negative Binomial Regression Models	13
2.1 Modeling of number of deaths	13
2.1.1 Basic concepts of Poisson Regression Model	13
2.1.2 Parameter estimation of Poisson Model	14
2.1.3 Basic concepts of Negative Binomial Regression	14
2.1.4 Parameter estimation of Negative Binomial Model	15
2.1.5 Smoothness of Poisson and Negative Binomial Models	15
2.1.6 Model Selection Criteria	15
2.1.7 Fitting the models	15
2.1.8 Checking for model assumptions	20
2.1.9 Goodness of fit	20
2.2 Modeling of risks of exposure	23
2.2.1 Basic Concepts	23
2.2.2 Fitting Negative binomial regression model	23
2.2.3 Goodness of fit	25
2.3 Estimation of mortality rates	27
2.3.1 Goodness of fit	27
3 The Lee-Carter Model	29
3.1 Basic concepts	29
3.2 Parameter Estimation	29
3.2.1 Singular Value Decomposition	30
3.3 Fitting Lee-Carter Model	31

3.4	Applying Lee-Carter Model	31
3.4.1	Fitted Lee-Carter model	31
3.4.2	Goodness of fit	35
4	Forecasting Techniques	37
4.1	Forecasting with Negative Binomial Model	37
4.1.1	Forecasting risks of exposure(2000 to 2009)	37
4.1.2	Forecasting number of deaths (2000 to 2009)	38
4.1.3	Forecasting rates of mortality (2000 to 2009)	40
4.2	Forecasting with Lee-Carter Model	41
4.2.1	Basic concepts of standard ARIMA (0,1,0) model	42
4.2.2	Forecasting rates of mortality (2000 to 2009)	44
5	Forecast errors	46
5.1	Basic concepts of forecast error	46
5.2	Distribution of forecast errors	46
5.3	Measurement of forecast errors	48
5.3.1	Basic concepts of Mean Squared Error (MSE)	48
5.3.2	Basic concepts of Mean Absolute Error (MAE)	48
5.3.3	Comparison of Sum of Squares of forecast errors	48
5.3.4	Comparison of MSE and MAE of forecast errors	48
5.4	Other options of Negative Binomial Regression Model	49
6	Conclusion	51
6.1	Summary Remarks	51
6.2	Known Weakness and Suggestion	52
	Appendix	53
	References	59

Thesis Overview

The datasets of observed Norwegian male mortality and their risks of exposure will be obtained from Human Mortality Database and used to fit both models. Poisson or negative binomial regression and Lee-Carter (1992) demographic models will be constructed from the datasets for the year period 1950 to 1999. The fitted models will be used to forecast mortality rates for the year period 2000 to 2009. The mortality set for the 50 age group in 1999 and 2009 will be used to validate both models. Descriptive summaries of both data sets will be covered in chapter one.

Poisson or Negative binomial regression model will be fitted to the Norwegian male mortality data by treating period and age groups as covariates. The number of deaths and their risks of exposure for the forecasting period will be predicted using the fitted models. Here some care may be necessary and a tradeoff between in sample fit and forecasting ability will be expected. To ensure the latter, relatively simple only functions will be allowed for the period covariate. The mortality rates for the forecasting period will then be estimated as the ratio of predicted number of deaths and their risks of exposure. These will be covered in chapters two and four.

Lee-Carter model is based on the original method for mortality forecasting suggested by Lee-Carter in "Lee, R.D. and L.R Carter: Modeling and forecasting U.S. Mortality, Journal of the American Statistical Association, 1992, vol. 87, pp. 659-67". The Lee-Carter model will be fitted to the the Norwegian male mortality data as the logarithmic transformation of age-specific central rate of mortality as a sum of an age-specific component that is independent of time, and the product of mortality index that describes the general level of mortality and an additional age-specific component that represents how rapidly or slowly mortality at each age varies when the mortality index changes. From this forecast of the mortality index, the actual age-specific mortality rates will be derived using the estimated age effects. These will be covered in chapters three and four.

Finally, both models will be compared using the mean squared error (MSE) and mean absolute error (MAE) with detailed considerations of the distribution of their respective forecast errors. Conclusion will then be drawn. These will be covered in chapters five and six.

Chapter 1

The Data

1.1 Description of Data

Both datasets of Norwegian male death counts and their risks of exposure are required to fit the models and perform forecasting analyses. The two sets of the data are from the calendar year period of 1950 to 2009. The data files of the male death counts are classified according to age, sex and time. The population size is given by one-year age groups from 0 up to 109 with an open interval for 110+, starting from 1950 through to 2009. The data files for risk exposure are organized accordingly with similar trend.

Data files for the Norwegian male risk exposures are estimates of the population exposed to risk of death during a 1x1 age-time interval based on annual (first January) population estimates, with a small correction that reflects the timing of deaths during the interval. Data files for the Norwegian male death counts are collected by 1x1 Lexis diagrams whilst their risk exposures are estimated from the diagram.

1.1.1 Sources of the data

Both demographic data files of Norwegian male death counts and their exposures to risk are available from the Human Mortality Database (maintained by the University of California, Berkeley (U.S.A), and the Max Planck Institute for Demographic Research (Germany)), www.mortality.org (data downloaded in January 2015).

1.1.2 Lexis Diagram

The Lexis Diagram displays the stock and flow of a population and the occurrence of demographic events over age and time. The diagram is a vital descriptive representation of demographic events. The risks of exposure from Human Mortality Database (2005) were computed using Lexis Diagram. The computational procedures are illustrated below:

Figure 1.1 depicts a Lexis diagram, which is a plot of a population's life experience in time against age. The graph is sectioned into one-year by one-year cells. Each 45°

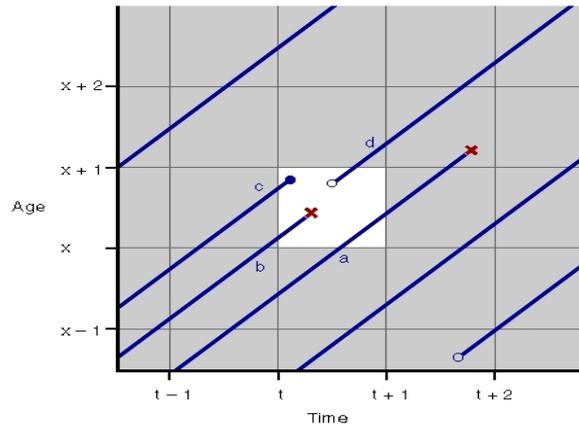


Figure 1.1: Example of Lexis Diagram

line represents an individual's life, which ends in death (red 'x') or out-migration (solid dot). An individual also may, at some time, migrate into the population (hollow dot).

We estimate risk of exposure for each 1x1 cell by using the highlighted 1x1 cell displayed in figure 1.1. As we can see, the cell starts at time t and age x . If we know the exact lines, then we can calculate the risk of exposure in person-years by adding up each of the line segment within the cell. We then divide the actual length of each segment by $\sqrt{2}$, since the life lines are 45° to the age and time axes.

However, exact life lines are rarely known in study of large national population. Instead, we often have the counts of individuals alive for each age at exact times t , $t + 1$, $t + 2$, etc. Considering the highlighted cell in figure 1.1 above again, for example, the count at time t and age x is 2 (lines b and c) and the count at time $t + 1$ and age x is 1 (line a). Given this information, our best estimate of the risk of exposure within the cell is merely the average of these two counts (thus, 1.5 person-years). Incidentally, line d does not contribute to this cell's risk of exposure estimate as it does not cross either of the boundaries at times t and $t + 1$.

1.1.3 Male mortality data

	Min.	Max.	Sum
Deaths	0.0	3,606	2,855,185
year	1950	2009	
Age	0	110+	

Table 1.1: Summary of Male mortality data

Table 1.1 gives a brief summary measures of the data file for number of deaths of Norwegian male from the calendar year 1950 to 2009. There are 2,855,185 total sum of death counts with 164 death counts for each calendar year. The counts range from a minimum value of 0.0 to a maximum value of 3606 with mean 156.8 .

The mortality data set has been classified into training set, used for fitting the models, test set, used for testing the fitted models and comparing forecasts and validation set, used for validating the fitted models and forecasts. The training set ranges from the year period of 1950 to 1999 and the test set is for the year period 2000 to 2009. The validation set consists of mortality rates for the 50 age group in the year 1999 and 2009. Diagrammatic presentations of these classifications are described below.

Observed number of male deaths (1950-1999)

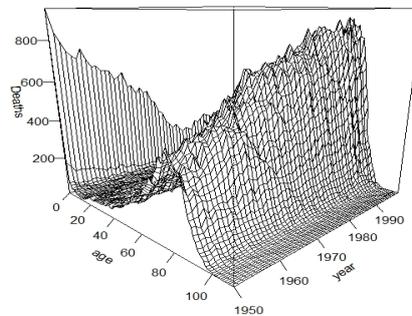


Figure 1.2: Observed Norwegian male death counts from 1950 to 1999

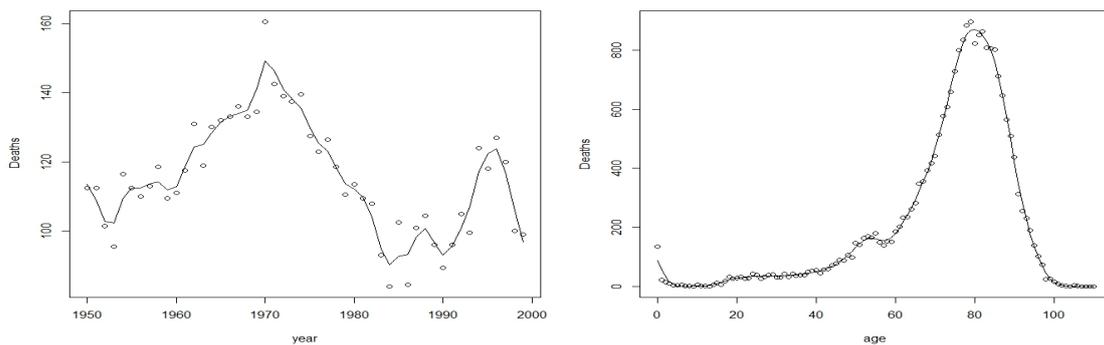


Figure 1.3: Observed male death counts (at age 50 and year 1999)

Figure 1.2 displayed above depicts a three-dimensional plot of observed number of deaths of Norwegian male from the year period 1950 to 1999 over the intervals of age and year. Figure 1.3 is a two-dimensional plot of figure 1.2 in the year 1999

for 50 age group, smoothed with splines of 25 degrees of freedom. The death counts of the Norwegian male against their ages are observed to be highest among the 70's and 80's age groups but decline steadily; recording minimal counts for 100's age groups. The death counts fluctuates with the years of counts. The maximum counts are recorded in the 1970's and the minimum counts in 1980's.

Observed number of male deaths (2000-2009)

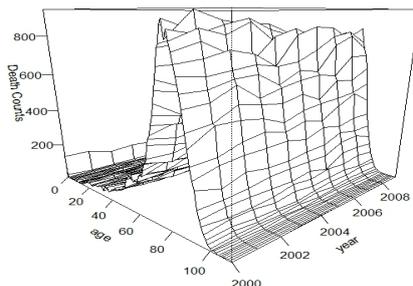


Figure 1.4: Observed Norwegian male death counts from 2000 to 2009

Displayed in figure 1.4 below is a three-dimensional plot of death counts against age and year for the year period 2000 to 2009. The plot follows the same trend of distribution of the training set in figure 1.2 and will be used in assessing the strength of the fitted models and performing forecast analyses of number of deaths.

1.1.4 Risks of exposure data

	Min.	Max.	Sum
Exposures	0.0	38,140	232,475,346
year	1950	2009	
Age	0	110+	

Table 1.2: Summary of risk exposure data

Summary measures of the data file of the Norwegian males exposed to risk for the year period 1950 to 1999 is tabulated above. About total sum of 232,475,346 observations are recorded with 10,530 for mean. Similar statistical summary recorded for age and year in the mortality data set are observed as well for the exposures.

Data set for Norwegian male exposed to risk are also categorised into training sets for the year period 1950 to 1999, test set from 2000 to 2009 and validation set consisting of 50 age group in the year 1999 and 2009. Models are as well fitted to the data set from 1950 to 1999 and forecasting strengths are examined and compared with the set from the year period 2000 to 2009.

Observed males exposed to risk (1950-1999)

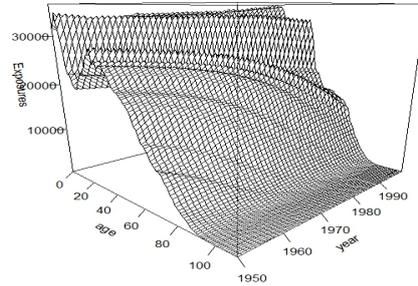


Figure 1.5: Observed Norwegian male exposure from 1950 to 1999

Figure 1.5 is a three-dimensional representation of the Norwegian males exposed to risk from the year period 1950 to 1999. This will be used in fitting the predictive model for forecasting.

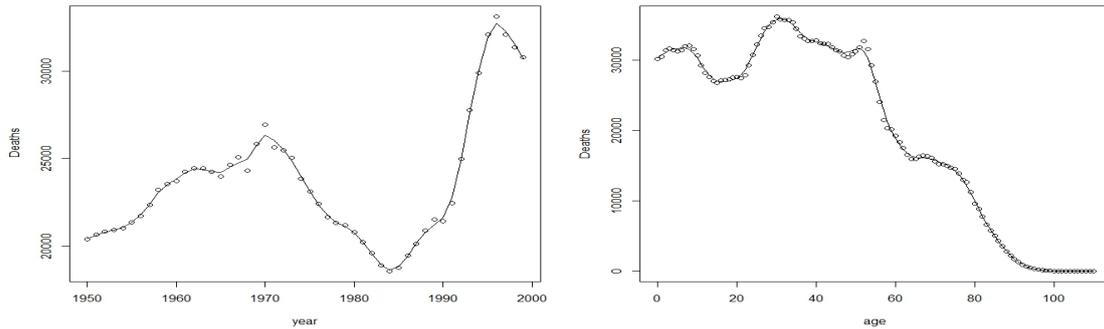


Figure 1.6: Observed male exposure (at age 50 and year 1999)

Displayed in figure 1.6 are two dimensional plots of figure 1.5 for observed males exposed to risk in the year 1999 for the 50 age group, smoothed with splines of 25 degrees of freedom. Risks of exposure are observed to fluctuate from 0 age group to 50 age group, declined roughly steadily thereafter to 100 age group in the year period 1999. The year period 1980's records the least risks of exposure with the highest occurring in the 1990's for the 50 age group.

Observed males exposed to risk (2000-2009)

Displayed in figure 1.7 is a three dimensional plot of observed Norwegian male risks of exposure from the year period of 2000 to 2009. The plot shows the same distribution trend of the training set in figure 1.5 and will be used in forecast analyses.

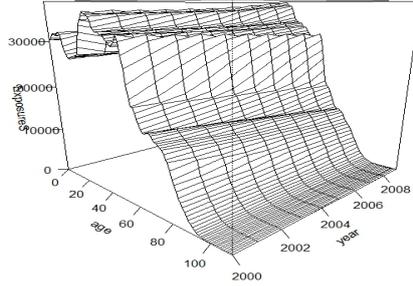


Figure 1.7: Observed Norwegian male exposure from 2000 to 2009

1.1.5 Estimation of observed mortality Rates

Mortality rate is the probability of an individual aged exactly x at time t will die before time $t + 1$. Mortality rate is estimated as the ratio of observed number of Norwegian males and their risk of exposure in 1×1 age-period interval. This implies mortality rates are highly influenced by risks of exposure.

Observed rates of mortality (1950-1999)

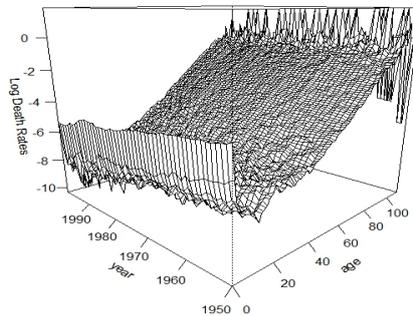


Figure 1.8: Observed Norwegian male log mortality from 1950 to 1999

Figure 1.8 is a three dimensional plot of observed log mortality for Norwegian males from the year period 1950 to 1999. The plot will be used in comparing and assessing fitted and forecast rates of mortality .

Figure 1.9 shown below are two dimensional plots of mortality rates on logarithmic scale in the year 1999 for the 50 age group, smoothed with splines of 25 degrees of freedom. The mortality rate attain minimum value for the 10 years age group and maximum for 110+ age group.

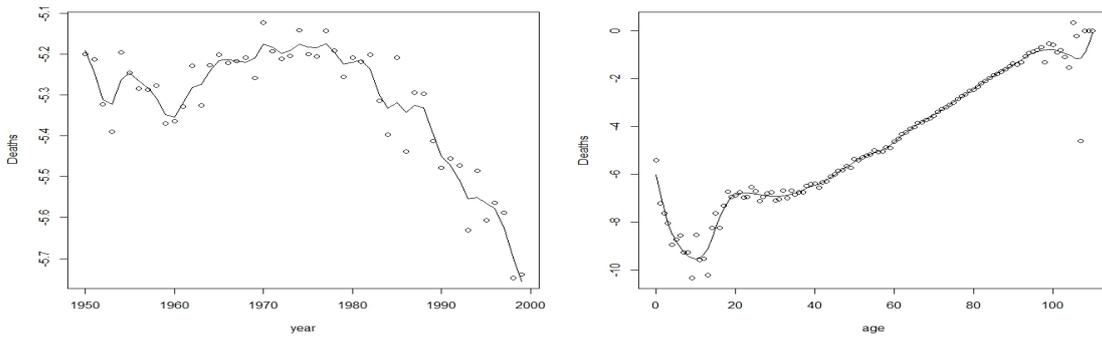


Figure 1.9: Observed male log mortality (at age 50 and year 1999)

Nevertheless, the mortality increased along with observed calendar years. A significant decline in the rates is observed from 0 to 20 age groups. The mortality rates increased with age thereafter. Further decline is observed after the 100 age group. This indicates young age groups are less exposed to risk of deaths and thus have less rates of mortality relative to the older groups.

Observed rates of mortality (2000-2009)

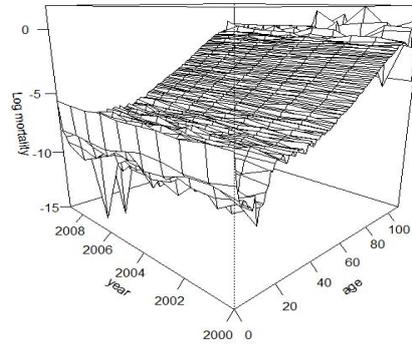


Figure 1.10: Observed death rates from 2000 to 2009, for comparing forecast

Figure 1.10 displays a three dimensional plot of observed log mortality of Norwegian males from the year period 2000 to 2009. The plot follows the same trend of distribution of the training set in figure 1.8. The plot will be used for comparing forecast of mortality rates of the same year period.

Poisson and Negative Binomial Regression Models

2.1 Modeling of number of deaths

We model the training set of the mortality data described in subsection 1.1.3 with Poisson and negative binomial regression. Described in subsections below are the basic concepts of the two distributions and how they are fitted to the training set.

2.1.1 Basic concepts of Poisson Regression Model

The poisson regression models the mortality data by treating both age and year covariates x_{ij} as categorical variables with the response variable y_{ij} as the number of deaths. Thus,

$$y_{ij} \sim \text{poisson}(u_{ij}) \text{ and } u_{ij} = E(y_{ij})$$

then,

$$\begin{aligned} \log(u_{ij}) &= \log(n_{ij}) + (x_{ij})^T \beta \\ \implies u_{ij} &= \exp\{\log(n_{ij}) + (x_{ij})^T \beta\} \end{aligned} \quad (2.1)$$

Where $\log(u_{ij})$ is the natural log link, $i = 0, 1, \dots, 110$ and $j = 1950, \dots, 1999$. n_{ij} is the number of risk exposure. The vector x_{ij} contains the values of the explanatory variables for each age i , year j combination. The explanatory variables coded as polynomials with the usual intercept 1, indicator variables corresponding to age, and indicators corresponding to calendar year.

The probability mass function (pmf) of Poisson distribution is given by:

$$P(Y_{ij} = y_{ij}) = \frac{\mu_{ij}^{y_{ij}} \exp(-\mu_{ij})}{y_{ij}!} \quad (2.2)$$

Hence from equation 2.1, we obtain the Poisson regression model to be:

$$P(Y_{ij} = y_{ij}) = \frac{(\exp\{\log(n_{ij}) + (x_{ij})^T \beta\})^{y_{ij}} \exp(-\exp\{\log(n_{ij}) + (x_{ij})^T \beta\})}{y_{ij}!}$$

The mean and variance of Poisson regression are $E(y_{ij}) = u_{ij}$ and $Var(y_{ij}) = u_{ij}$.

2.1.2 Parameter estimation of Poisson Model

Fitting Poisson distribution 2.2 to the Norwegian male mortality requires estimation of β parameters with maximum likelihood method. We express the likelihood function of equation 2.2 as;

$$L(\beta) = \prod_{ij} P(y_{ij}) = \prod_{i,j} \frac{\mu_{ij}^{y_{ij}} \exp(-\mu_{ij})}{y_{ij}!} \quad (2.3)$$

with μ_{ij} as defined in 2.1. The log-likelihood function of equation 2.3 becomes:

$$\ln L(\beta) = \sum_{i,j} \{y_{ij} \ln \mu_{ij} + \mu_{ij} - \ln y_{ij}!\} \quad (2.4)$$

The maximum likelihood estimates are the values of β that maximize $L(\beta)$ of equation 2.3. The variance-covariance matrix of the estimators is $\Sigma = -H^{-1}$, where H^{-1} is the Hessian matrix of second derivatives of equation 2.4. H^{-1} is used to find the p-values of β estimates .

2.1.3 Basic concepts of Negative Binomial Regression

Negative binomial regression is a type of Poisson regression in which the dependent variable y_{ij} is a count of the number of times death occurs. It has one extra parameter more than Poisson regression that adjusts the variance independently from the mean. The appropriate parameterisation of Negative binomial distribution given by Hilbe (2011) follows:

$$f(y_{ij}) = \frac{\Gamma(y_{ij} + \frac{1}{\kappa})}{y_{ij}! \Gamma(\frac{1}{\kappa})} \left(\frac{1}{1 + \kappa \mu_{ij}} \right)^{\frac{1}{\kappa}} \left(\frac{\kappa \mu_{ij}}{1 + \kappa \mu_{ij}} \right)^{y_{ij}} \quad (2.5)$$

where $\mu_{ij} > 0$ is the mean of y_{ij} and $\kappa = \frac{1}{\theta} > 0$ is the dispersion parameter. The mean and variance of negative binomial regression model are:

$$E(y_{ij}) = \mu_{ij} \text{ and } Var(y_{ij}) = u_{ij}(1 + \kappa u_{ij})$$

Thus if ,

$$y_{ij} \sim negb(u_{ij})$$

then from 2.1,

$$\implies u_{ij} = \exp\{\log(n_{ij}) + (x_{ij})^T \beta\}$$

Hence, the pmf of Negative binomial regression model becomes:

$$f(y_{ij}) = \frac{\Gamma(y_{ij} + \frac{1}{\kappa})}{y_{ij}! \Gamma(\frac{1}{\kappa})} \left(\frac{1}{1 + \kappa \exp\{\log(n_{ij}) + (x_{ij})^T \beta\}} \right)^{\frac{1}{\kappa}} \left(\frac{\kappa \exp\{\log(n_{ij}) + (x_{ij})^T \beta\}}{1 + \kappa \exp\{\log(n_{ij}) + (x_{ij})^T \beta\}} \right)^{y_{ij}} \quad (2.6)$$

With $E(y_{ij}) = \mu_{ij}$ and $Var(y_{ij}) = u_{ij}(1 + \kappa u_{ij})$. κ measures overdispersion by controlling extra variation compared to the Poisson regression model.

2.1.4 Parameter estimation of Negative Binomial Model

In other to fit the distribution 2.5 to the Norwegian male mortality, we also need to estimate κ and β with maximum likelihood method. The likelihood function is:

$$L(\kappa, \beta) = \prod_{i,j} f(y_{ij}) = \prod_{i,j} \frac{\Gamma(y_{ij} + \frac{1}{\kappa})}{y_{ij}! \Gamma(\frac{1}{\kappa})} \left(\frac{1}{1 + \kappa\mu_{ij}}\right)^{\frac{1}{\kappa}} \left(\frac{\kappa\mu_{ij}}{1 + \kappa\mu_{ij}}\right)^{y_{ij}} \quad (2.7)$$

with μ_{ij} as defined in 2.1. The log-likelihood function of equation 2.7 becomes:

$$\ln L(\kappa, \beta) = \sum_{i,j} \left\{ y_{ij} \ln \kappa - y_{ij} \mu_{ij} - \left(y_{ij} + \frac{1}{\kappa}\right) \ln(1 + \kappa e^{\mu_{ij}}) + \ln \Gamma(y_{ij} + \kappa) - \ln \Gamma(y_{ij} + 1) - \ln \Gamma\left(\frac{1}{\kappa}\right) \right\} \quad (2.8)$$

Equation 2.8 is maximised and the values of κ and β that maximise $L(\kappa, \beta)$ are the maximum likelihood estimates. The p-values of β estimates are obtained with the same procedures explained in subsection 2.1.2.

2.1.5 Smoothness of Poisson and Negative Binomial Models

We exploit smoothness of Poisson and Negative binomial regression models described in subsections 2.1.1 and 2.1.3 by specifying polynomial terms with degrees p and q in the regression, for the explanatory variables age and year respectively. Hence from the model, $\log(u_{ij}) = \log(n_{ij}) + (x_{ij})^T \beta$ where $u_{ij} = E(y_{ij})$, the ordinary polynomial regression becomes;

$$\log(u_{ij}) = \log(n_{ij}) + \beta_0 + \beta_1 i + \dots + \beta_p i^p + \beta_{p+1} j + \dots + \beta_{p+q} j^q$$

Where

$$(x_{ij})^T \beta = \beta_0 + \beta_1 i + \dots + \beta_p i^p + \beta_{p+1} j + \dots + \beta_{p+q} j^q \quad (2.9)$$

$i = 0, 1, \dots, 110$ and $j = 1950, \dots, 1999$ as usual and β 's are beta coefficients over age and year in the regression.

2.1.6 Model Selection Criteria

We select the best model with Akaike information criteria (AIC), since the parameters are maximum likelihood estimates described in the previous sections and the mortality data has high dimensions. If L is the maximised value of the likelihood function of the model and q is the number of estimated parameters of the model, then the AIC value of the model is;

$$AIC = -2\ln(L) + 2q \quad (2.10)$$

With given number of candidate models for the mortality data, the preferred model is the one with the least AIC value.

2.1.7 Fitting the models

In this subsection, we apply the previous subsections to fit the "optimal" Poisson and Negative binomial regression models to the training set in subsection 1.1.3.

R notes: We face multicollinearity and numerical problems with the large polynomials terms in equation 2.9. Multicollinearity involves inclusion of highly correlated independent explanatory variables in the regression. This can inflate estimates for β 's with high standard errors. We address these problems by using orthogonal polynomials for age i and year j explanatory variables and standardising year j . Thus, equation 2.9 can be basically written in the orthogonal form as;

$$(x_{ij})^T \beta = \beta_0^* + \beta_1^* \nu_{1i} + \dots + \beta_p^* \nu_{pi}^p + \beta_{p+1}^* \nu_{1j} + \dots + \beta_{p+q}^* \nu_{qj}^q \quad (2.11)$$

Where,

$$\nu_{ki} = c_{k,k} + c_{k,k-1} i_p + \dots + c_{k,1} i_p^{k-1} + i_p^k, k = 1, 2, \dots, p$$

And

$$\nu_{rj} = c_{r,r} + c_{r,r-1} j_q + \dots + c_{r,1} j_q^{r-1} + j_q^r, r = 1, 2, \dots, q$$

We have,

$$\log(u_{ij}) = \log(n_{ij}) + \beta_0^* + \beta_1^* \nu_{1i} + \dots + \beta_p^* \nu_{pi}^p + \beta_{p+1}^* \nu_{1j} + \dots + \beta_{p+q}^* \nu_{qj}^q \quad (2.12)$$

Standardising the explanatory variable year j with a , implies $j - a = s$. Hence,

$$\log(u_{is}) = \log(n_{is}) + \beta_0^* + \beta_1^* \nu_{1i} + \dots + \beta_p^* \nu_{pi}^p + \beta_{p+1}^* \nu_{1s} + \dots + \beta_{p+q}^* \nu_{qs}^q \quad (2.13)$$

ν_{ki} and ν_{rs} are orthogonal polynomials of degrees k in i_p and r in s_q respectively. The $\nu_k = (\nu_{ki}), k = 1, 2, \dots, p$ and $\nu_r = (\nu_{rs}), r = 1, 2, \dots, q$ represent the k^{th} and r^{th} orthogonal coefficients, i.e they satisfy the following relationships;

$$\sum_i \nu_{ki} = 0, \sum_s \nu_{rs} = 0, k = 1, 2, \dots, p, r = 1, 2, \dots, q \quad (2.14)$$

And,

$$\sum_i \nu_{ki} \nu_{gi} = 0, \sum_s \nu_{rs} \nu_{hs} = 0, k \neq g = 1, 2, \dots, p, r \neq h = 1, 2, \dots, q \quad (2.15)$$

Equations 2.14 and 2.15 are used to obtain systems of equations from which the orthogonal coefficients ν_k and ν_r are derived. These computational procedures are implemented in the `glm` and `glm.nb` R packages for Poisson and Negative binomial regression models respectively. We use the concept of variance inflation factor to check the multicollinearity in the regression. This is also implemented in R with "vif()" function of the "car" package.

Fitting Poisson regression model

The AIC values for selection of "optimal" Poisson regression model are infinite. This is an indication of presence of overdispersion in the mortality data. This is checked and confirmed in subsection 2.1.8. Hence we use the same polynomial terms $p=25$ and $q=4$ of the "optimal" Negative binomial regression model below to fit the Poisson regression with the same procedures for comparison purposes.

Fitting negative binomial regression model

We fit our preferred Negative binomial regression model to the training set of number of deaths in subsection 1.1.3 based on the results displayed in the tables below.

Polynomial Degrees	Akaike information criteria (AICs)
p=22 and q=1	43,163
p=23 and q=2	42,629
p=24 and q=3	42,593
p=25 and q=4	42,578

Table 2.1: Comparison of selection criteria for p and q

Polynomial Degrees	Akaike information criteria (AICs)
p=25 and q=1	43,117
p=25 and q=2	42,621
p=25 and q=3	42,591
p=25 and q=4	42,578

Table 2.2: Comparison of p=25 with different q

Table 2.1 shows the various AICs values for different values of polynomial orders p and q. Table 2.2 compares AICs values of polynomial order p=25 with different values of polynomial orders q. It is noted from the two tables that p=25 and q=4 yields the least AIC value, and hence a model with 30 parameters. The AIC for this "optimal" model is 42578. Where $i = 0, 1, \dots, 110$ and $s = -25, \dots, 24$, since we standardised j with 1975.

Fitted Poisson Regression Model

Call: glm(formula = ddths ~ poly(age, 25) + poly(yearc, 4) + offset(lexpa), family = "poisson")				
Deviance Residuals:				
Min	1Q	Median	3Q	Max
-14.3413	220	0.67	10 20.5	587
Coefficients:				
Estimate	Std. Error	z value	Pr(> z)	
(Intercept) -4.35189	0.01146	-379.663	< 2e-16	***
poly(age, 25)1 180.61817	1.42959	126.343	< 2e-16	***
poly(age, 25)2 28.51783	1.74674	16.326	< 2e-16	***
poly(age, 25)3 -23.55485	1.92424	-12.241	< 2e-16	***
poly(age, 25)4 6.58117	2.0097	3.275	0.001058	**
poly(age, 25)5 -7.10560	2.03761	-3.487	0.000488	***
poly(age, 25)6 13.23086	2.02978	6.518	7.11E-11	***
poly(age, 25)7 -11.10281	2.0007	-5.549	2.87E-08	***
poly(age, 25)8 12.60807	1.95693	6.443	1.17E-10	***
poly(age, 25)9 -2.19781	1.90147	-1.156	0.247743	
poly(age, 25)10 1.28832	1.83611	0.702	0.482891	
poly(age, 25)11 1.41675	1.76212	0.804	0.421393	
poly(age, 25)12 1.24164	1.68638	0.736	0.461562	
poly(age, 25)13 -4.07564	1.61013	-2.531	0.011366	*
poly(age, 25)14 4.19352	1.53308	2.735	0.006231	**
poly(age, 25)15 -6.34910	1.45698	-4.358	1.31E-05	***
poly(age, 25)16 1.84701	1.36865	1.35	0.177173	
poly(age, 25)17 -3.04714	1.26781	-2.403	0.01624	*
poly(age, 25)18 -2.26866	1.15475	-1.965	0.049456	*
poly(age, 25)19 -0.37199	1.00941	-0.369	0.712484	
poly(age, 25)20 -2.44590	0.8658	-2.825	0.004728	**
poly(age, 25)21 -1.40383	0.68875	-2.038	0.041527	*
poly(age, 25)22 0.01178	0.54378	0.022	0.982711	
poly(age, 25)23 -1.86142	0.38473	-4.838	1.31E-06	***
poly(age, 25)24 0.48211	0.23167	2.081	0.037436	*
poly(age, 25)25 -0.49846	0.20242	-2.463	0.013797	*
poly(yearc, 4)1 -4.00652	0.07781	-51.492	< 2e-16	***
poly(yearc, 4)2 -3.59545	0.07672	-46.865	< 2e-16	***
poly(yearc, 4)3 -0.60348	0.07633	-7.906	2.65E-15	***
poly(yearc, 4)4 0.37410	0.07566	4.944	7.64E-07	***
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				
(Dispersion parameter for poisson family taken to be 1)				
Null deviance: 2715240 on 5549 degrees of freedom				
Residual deviance: 13079 on 5520 degrees of freedom				
AIC: InfNumber of Fisher Scoring iterations: 9				

Table 2.3: Fitted Poisson regression Model with GLM

Table 2.3 above shows Poisson regression model fitted with R "glm" package. The deviance is 13,079 on 5520 degrees of freedom. Also displayed are regression coefficients against standard errors and p-values. Most of the coefficients are statistically significant.

Fitted Negative Binomial Regression Model

Call: glm.nb(formula = ddths ~ poly(age, 25) + poly(yearc, 4) + offset(lexpa), init.theta = 122.9929209, link = log)						
	Min	1Q	Median	3Q	Max	
			-7.6709	0.5038	9.2752	
Coefficients:						
Estimate				Std. Error	z value	Pr(> z)
(Intercept)			-4.35285	0.01152	-377.725	< 2e-16 ***
poly(age, 25)1	180.75788			1.43201	126.227	< 2e-16 ***
poly(age, 25)2	28.53522			1.74808	16.324	< 2e-16 ***
poly(age, 25)3	-23.24968			1.92472	-12.079	< 2e-16 ***
poly(age, 25)4	6.43388			2.00928	3.202	0.001364 **
poly(age, 25)5	-7.09929			2.03623	-3.486	0.000489 ***
poly(age, 25)6	13.17405			2.02761	6.497	8.18E-11 ***
poly(age, 25)7	-11.02722			1.99776	-5.52	3.39E-08 ***
poly(age, 25)8	12.60309			1.95362	6.451	1.11E-10 ***
poly(age, 25)9	-2.09122			1.89831	-1.102	0.270625
poly(age, 25)10	1.32722			1.83365	0.724	0.469181
poly(age, 25)11	1.55193			1.76151	0.881	0.378305
poly(age, 25)12	1.32073			1.68863	0.782	0.434137
poly(age, 25)13	-3.98959			1.61653	-2.468	0.013587 *
poly(age, 25)14	4.23537			1.54576	2.74	0.006144 **
poly(age, 25)15	-6.24619			1.47687	-4.229	2.34E-05 ***
poly(age, 25)16	1.85131			1.39788	1.324	0.18538
poly(age, 25)17	-2.97180			1.30853	-2.271	0.023141 *
poly(age, 25)18	-2.23821			1.20572	-1.856	0.063407 .
poly(age, 25)19	-0.33900			1.07307	-0.316	0.752065
poly(age, 25)20	-2.39980			0.93802	-2.558	0.010516 *
poly(age, 25)21	-1.41764			0.76985	-1.841	0.065556 .
poly(age, 25)22	0.05174			0.62401	0.083	0.933917
poly(age, 25)23	-1.84551			0.46201	-3.995	6.48E-05 ***
poly(age, 25)24	0.48498			0.31318	1.549	0.121487
poly(age, 25)25	-0.48590			0.26348	-1.844	0.065156 .
poly(yearc, 4)1	-5.60739			0.14247	-39.358	< 2e-16 ***
poly(yearc, 4)2	-3.37736			0.14184	-23.811	< 2e-16 ***
poly(yearc, 4)3	-0.81175			0.14185	-5.723	1.05E-08 ***
poly(yearc, 4)4	0.55143			0.14169	3.892	9.95E-05 ***
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1						
(Dispersion parameter for Negative Binomial(122.9929) family taken to be 1)						
Null deviance: 1184265.2 on 5549 degrees of freedom						
Residual deviance: 6047.5 on 5520 degrees of freedom						
AIC: 42578						
Number of Fisher Scoring iterations: 1						
Theta: 122.99						
Std. Err.: 4.89						
2 x log-likelihood: -42516.17						

Table 2.4: Fitted Negative Binomial Regression Model

Displayed in table 2.4 above is the "optimal" negative binomial model fitted with R "glm.nb" package. The deviance of the fit is 6047.5 on 5520 degrees of freedom. Displayed are also the fitted regression coefficients with their respective standard errors and p-values. Both results displayed in tables 2.3 and 2.3 are based on orthogonal polynomials of the explanatory variables age i and year s which we explained earlier.

2.1.8 Checking for model assumptions

In this subsection, we check assumptions underlying Poisson and Negative binomial models fitted in subsections 2.1.7 to determine which of them best fits the mortality data. Poisson distribution assumes that the mean is the same as the variance. Overdispersion then occurs when the data show extra variation that is greater than the mean. Negative binomial regression is more flexible in modeling overdispersed data. Thus we check for overdispersion in the mortality data by computing for scalar parameter ϕ . for Poisson and testing the significance of the dispersion parameter of the fitted Negative binomial regression model with Wald test.

The scalar parameter ϕ defines how spread the mortality data is and is held 1 in Poisson, for the mean and variance to be equal. There is overdispersion when $\phi > 1$. ϕ is derived as the mean square of the Pearson residuals, implemented in R as:

```
Pearson <- -residuals(model, "pearson")
Squared <- -sum(Pearson^2); Squared
[1] 10991.57
phi <- -sum(Squared)/df.residual(model); phi
[1] phi = 1.991587
```

From the result above, the variance of Poisson is 99% larger than the mean. Also, Negative binomial regression model becomes Poisson regression model when the overdispersion parameter $\kappa = 0$. We can check statistical significance of the overdispersion parameter by testing the hypothesis $H_0 : \kappa = 0$ against $H_1 : \kappa > 0$ with Wald or likelihood ratio test. The value of $\hat{\theta}$ for the fitted negative binomial model displayed in table 2.4 is 122.99 with standard error of 4.89. Since the dispersion parameter $\kappa = \frac{1}{\theta}$ means $\hat{\kappa} = 0.0081 > 0$. The test statistic of MLE $\hat{\kappa}$ for the Wald test is approximated to $N(0, 1)$. The test statistic is implemented in R as;

```
hat_kappa = 0.0081
kappa_o = 0
std.error = 4.89
Wald_statistic = (hat_kappa - kappa_o)/std.error ;
Wald_statistic = 0.001656442
```

$0.001656442 < Z_{\frac{\alpha}{2}} = 1.96$. Hence the null hypothesis $H_0 : \kappa = 0$ is rejected at the 5% level of significance. Both results indicate presence of overdispersion in the mortality data. In this regard, Negative binomial regression yields a better fit compared to Poisson and would be used henceforth.

2.1.9 Goodness of fit

We note the fitted Negative Binomial model gives good fit to the mortality compared to Poisson due to overdispersion in subsection 2.1.8. In this subsection, we validate the fitted negative binomial model with residual analysis.

Residual deviance diagnostics

The residual deviance terms which are elements of variation unexplained by the fitted negative binomial model are random and normally distributed with zero mean and constant variance. We check for the randomness and normality of these residuals with three dimensional and Q-Q plots .

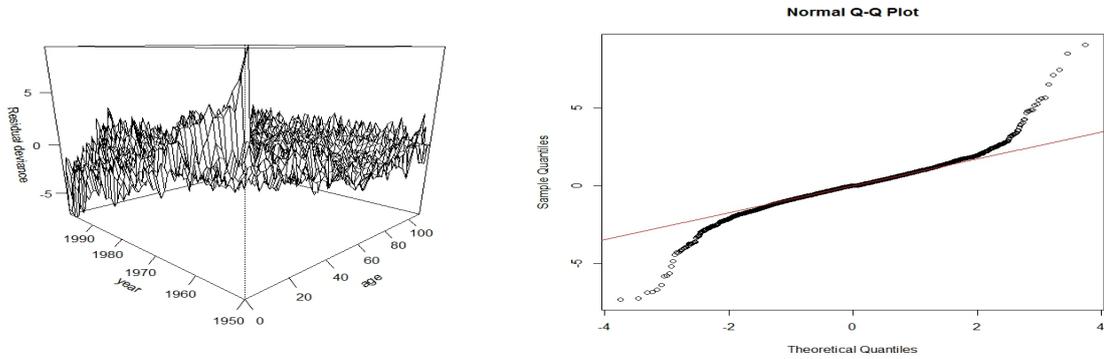


Figure 2.1: Residual deviance of log number of deaths

Figure 2.3 shows three dimensional and Q-Q plots for the residual deviance of the fitted negative binomial model for the male mortality. There is a clear structured pattern of residuals for the young age groups in 1950-60's with the rest of the residuals randomly distributed in the three dimensional plot. Some of the residuals depart from normality at the ends of the Q-Q plot, with majority roughly normal. Thus, adequacy of normality is probably indicated.

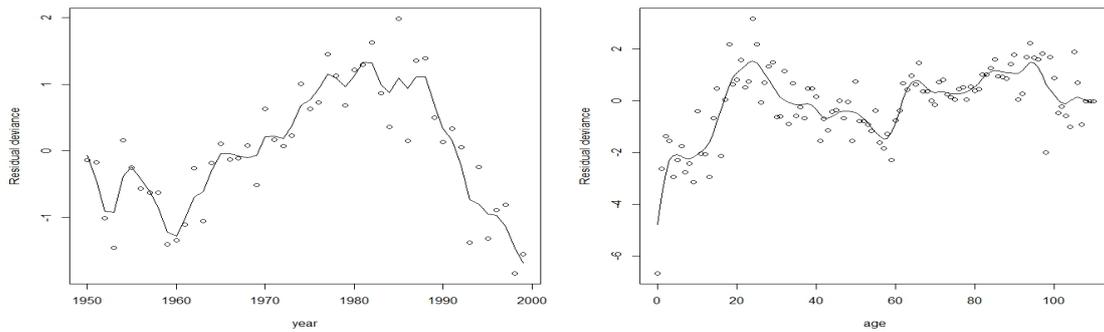


Figure 2.2: Residual plots of log mortality (at age 50 and year 1999)

Figure 2.4 displays two dimensional plots of the residuals for the 50 age group in the year 1999, smoothed with splines of 25 degrees of freedom. Few large values for the 50 age group are not well captured in the fitted model. Apart from that, both plots show random structure of the residuals.

Fitted against observed number of deaths

We assess how well the fitted Negative binomial regression model captures the historical trend of the observed number of deaths.

Plot of fitted number of deaths

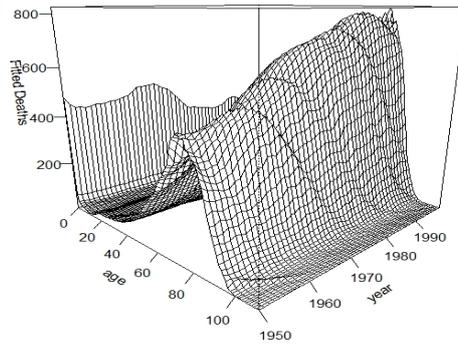


Figure 2.3: Fitted Norwegian male deaths using Negative Binomial Model

Figure 2.5 displays the smoothed plot of the training set of the observed number of deaths displayed in figure 1.2. The plot shows similar distribution pattern of the observed data. Figure 1.2 covers large number of deaths as compared to 2.5.

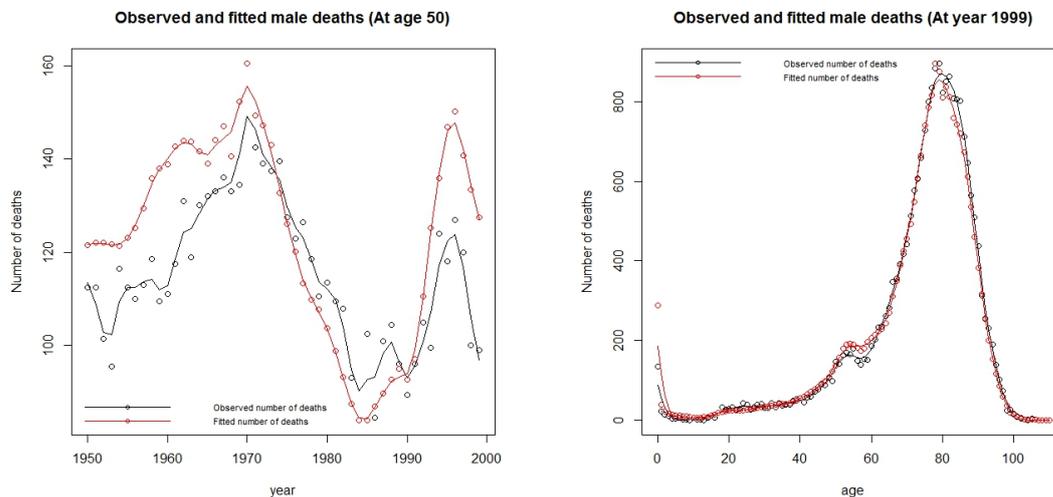


Figure 2.4: Fitted Norwegian male deaths (at age 50 and year 1999)

Figure 2.2 displays two dimensional plots of the fitted deaths superimposed on the observed data for the 50 age group in the year 1999. The fitted number of deaths depart considerably from the historical trend through time but fit well through age.

2.2 Modeling of risks of exposure

The number of risk exposure is also modeled using Negative binomial regression model. Having good fit to risks of exposure will yield consistent prediction of mortality rates, since mortality rates are highly influenced by risks of exposure.

2.2.1 Basic Concepts

Let x_{ij}^T be categorical variables for both year and age with the number of risk exposure as the response variable n_{ij} .

$$n_{ij} \sim \text{negb}(u_{ij}) \text{ and } u_{ij} = E(n_{ij})$$

then,

$$\begin{aligned} \log(u_{ij}) &= \log((x_{ij}^T \beta) \\ \implies u_{ij} &= \exp\{(x_{ij}^T \beta)\} \end{aligned} \quad (2.16)$$

Where $\log(u_{ij})$ is the natural log link. Also, $i = 0, 1, \dots, 110$ and $j = 1950, \dots, 1999$ are indicators of explanatory variables x_{ij} . The vector x_{ij} contains the values of the explanatory variables for each age i , year j combination and β , the regression coefficients. Inserting equation 2.16 into equation 2.5, we get pmf for risk exposure as;

$$P(N_{ij} = n_{ij}) = \frac{\Gamma(n_{ij} + \frac{1}{\kappa})}{\Gamma(n_{ij} + 1)\Gamma(\frac{1}{\kappa})} \left(\frac{1}{1 + \kappa(\exp\{(x_{ij}^T \beta)\})} \right)^{\frac{1}{\kappa}} \left(\frac{\kappa(\exp\{(x_{ij}^T \beta)\})}{1 + \kappa(\exp\{(x_{ij}^T \beta)\})} \right)^{n_{ij}} \quad (2.17)$$

With μ_{ij} as defined in equation 2.16, the likelihood function of equation 2.17 is:

$$L(\kappa, \beta) = \prod_{i,j} p(n_{ij}) = \prod_{i,j} \frac{\Gamma(n_{ij} + \frac{1}{\kappa})}{n_{ij}! \Gamma(\frac{1}{\kappa})} \left(\frac{1}{1 + \kappa \mu_{ij}} \right)^{\frac{1}{\kappa}} \left(\frac{\kappa \mu_{ij}}{1 + \kappa \mu_{ij}} \right)^{n_{ij}} \quad (2.18)$$

The log-likelihood function of equation 2.18 becomes:

$$\ln L(\kappa, \beta) = \sum_{i,j} \left\{ n_{ij} \ln \kappa - n_{ij} \mu_{ij} - (n_{ij} + \frac{1}{\kappa}) \ln(1 + \kappa e^{\mu_{ij}}) + \ln \Gamma(n_{ij} + \kappa) - \ln \Gamma(n_{ij} + 1) - \ln \Gamma(\frac{1}{\kappa}) \right\} \quad (2.19)$$

The dispersion parameter κ and β regression coefficients of equation 2.19 are estimated by maximum likelihood as explained in subsection 2.1.4. $i = 0, 1, \dots, 110$ and $j = 1950, \dots, 1999$ as usual. Smoothness of the Negative binomial regression model for the risks of exposure is also exploited by specifying polynomial orders p and q in the regression, for the age and year covariates as explained in subsection 2.1.5. Model selection for risk exposure is also based on AIC criteria described in subsection 2.1.6. Risks of exposure are not integers and so are rounded off.

2.2.2 Fitting Negative binomial regression model

We also face multicollinearity and numerical problems in fitting Negative binomial regression to the risks of exposure. We also address these by using orthogonal polynomials for age i and year j and standardising year j with 1975. These are based on the same procedures explained in subsection 2.1.7.

Fitted negative binomial regression Model

Call: glm.nb(formula = expo ~ poly(age, 10) + poly(yearc, 3), link = "log")					
Deviance Residuals:					
	Min	1Q	Median	3Q	Max
	-3.032	-0.8949	0.1176	0.6648	3.7778
Coefficients:					
	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	8.424e+00	0.00793	1062.89000	< 2e-16	***
poly(age, 10)1	-1.816e+02	0.95790	-189.61800	< 2e-16	***
poly(age, 10)2	-1.351e+02	1.13100	-119.51100	< 2e-16	***
poly(age, 10)3	-7.557e+01	1.17600	-64.23900	< 2e-16	***
poly(age, 10)4	-2.971e+01	1.12700	-26.36100	< 2e-16	***
poly(age, 10)5	-5.246e+00	1.01100	-5.19000	0.00000	***
poly(age, 10)6	3.236e+00	0.85400	3.78900	0.00015	***
poly(age, 10)7	3.999e+00	0.68120	5.87000	0.00000	***
poly(age, 10)8	1.889e+00	0.51310	3.68200	0.00023	***
poly(age, 10)9	9.892e-01	0.36750	2.69200	0.00710	**
poly(age, 10)10	4.056e-01	0.25900	1.56600	0.11739	
poly(yearc, 3)1	1.035e+01	0.16330	63.35000	< 2e-16	***
poly(yearc, 3)2	-5.392e-01	0.16330	-3.30200	0.00096	***
poly(yearc, 3)3	3.801e-01	0.16330	2.32800	0.01993	*
--- Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 (Dispersion parameter for Negative Binomial(42.3659) family taken to be 1) Null deviance: 491534.5 on 5549 degrees of freedom Residual deviance: 5748.7 on 5536 degrees of freedom AIC: 90895 Number of Fisher Scoring iterations: 1 Theta: 42.366 Std. Err.: 0.880 2 x log-likelihood: -90865.302					

Table 2.5: Fitted Negative Binomial Model

Table 2.5 below is the fitted Negative binomial regression model for the Norwegian male exposed to risk. The deviance of the model is 5748.7 on 5536 degree of freedom. Displayed also are the regression coefficients with their respective standard errors and p-values.

2.2.3 Goodness of fit

Since risks of exposure are also modeled with Negative binomial regression model, we apply the same measures of fit validation described earlier in subsection 2.1.9. We check for randomness and normality of residual deviance with three dimensional and Q-Q plots respectively.

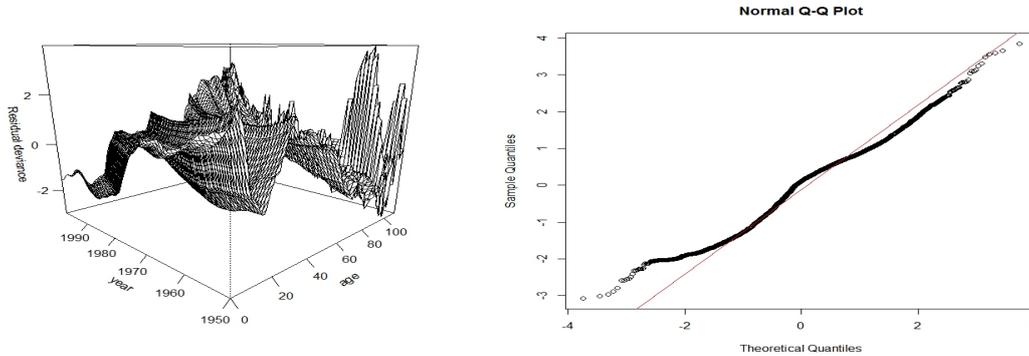


Figure 2.5: Plots of residuals

Figure 2.8 depicts three dimensional and Q-Q plots of residual deviance of the fitted model. The structure of the three dimensional plot shows no clear pattern of randomness of the residuals. The Q-Q plot shows considerable variation of the residuals from normality.

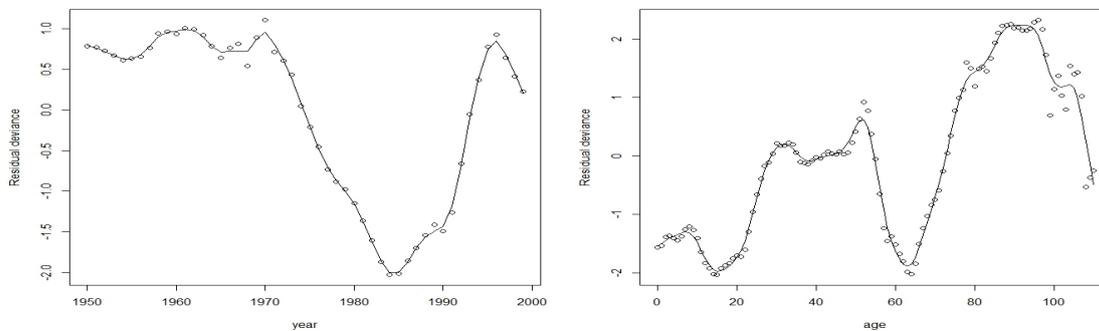


Figure 2.6: Plots of residuals (at age 50 and year 1999)

Displayed in figure 2.9 are two dimensional plots of residuals for the 50 age group in 1999, smoothed with splines of 25 degrees of freedom. Both plots show systematic downward and upward trends of residuals through time and age. They are well smoothed with no clear pattern of randomness. This is an indication of departure from normality, hence suggesting that the historical trend of risks of exposure are not sufficiently captured by the fitted Negative binomial regression model for the 50 age group in the year 1999.

Fitted against observed risks of exposure

In this subsection, we assess how well the fitted Negative binomial regression model captures the historical trends of the observed risks of exposure over age and time with two and three dimensional plots. The model can give good prediction if it fits well on the historical paths of the observed data.

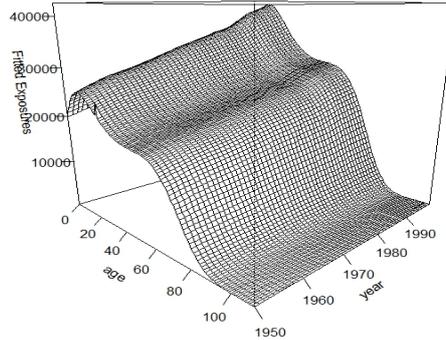


Figure 2.7: Plot of fitted male risk exposure, 1950 to 1999

Figure 2.10 shows smoothed plot of the training set of observed risks of exposure displayed in figure 1.7. There is considerable discrepancy between the structure of the two plots which is captured in residuals in subsection 2.2.3.

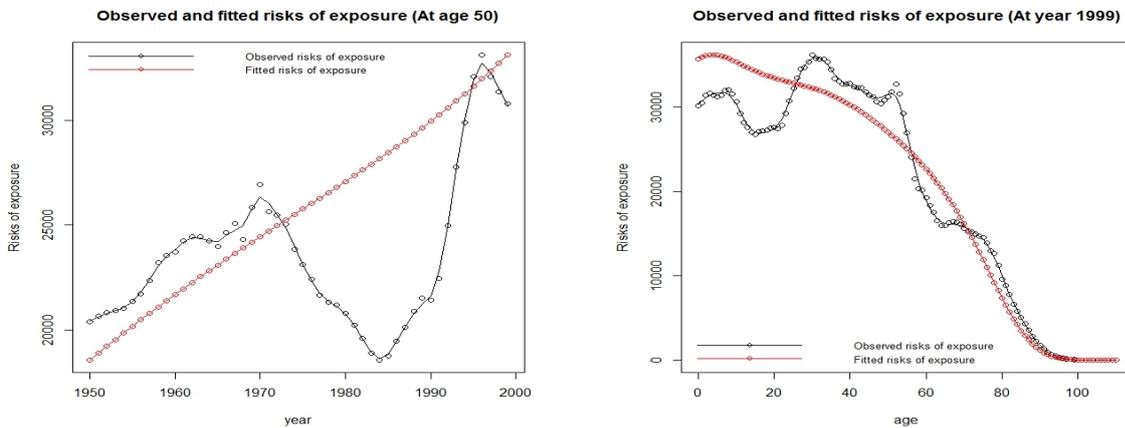


Figure 2.8: Fitted risk exposure (at year 1999 and age 50)

Displayed in figure 2.11 are two dimensional plots of fitted and observed risks of exposure for the 50 age group in the year 1999. The plots show trends of distribution of the fitted and observed risks of exposure in figures 2.10 and 1.5 respectively. The fitted risks of exposure depart significantly from the observed trend through time

for the 50 age group but fit roughly to the observed trend through age in the year 1999. Thus we can conclude that the fitted model does not provide adequate fit to the observed data for the 50 age group in the year 1999.

2.3 Estimation of mortality rates

Mortality rate is estimated as the ratio of number of fitted deaths and risk of exposure. That is, if \hat{y}_{ij} and \hat{n}_{ij} are the number of fitted deaths and risk of exposure for $i=0,1,\dots,110$ and $j=1950,\dots,1999$ as usual, then the nonparametric estimate of mortality rate \hat{m}_{ij} for age i and year j is expressed as:

$$\hat{m}_{ij} = \frac{\hat{y}_{ij}}{\hat{n}_{ij}} = \exp((x_{ij})^T \hat{\beta})$$

$$\implies \log(\hat{m}_{ij}) = \log\left(\frac{\hat{y}_{ij}}{\hat{n}_{ij}}\right) = (x_{ij})^T \hat{\beta} \quad (2.20)$$

The vector x_{ij} contains the values of the explanatory variables for each age i , year j combination and the $\hat{\beta}$ are the regression coefficients.

2.3.1 Goodness of fit

We assessed both year and age-specific patterns of trend of the observed log mortality rates in figure 1.9. These patterns of trend are expected to be fully captured by the fitted Negative binomial model for consistent forecasting. Hence in this subsection, we check for consistency of the fitted trends on the observed trends through age and year.

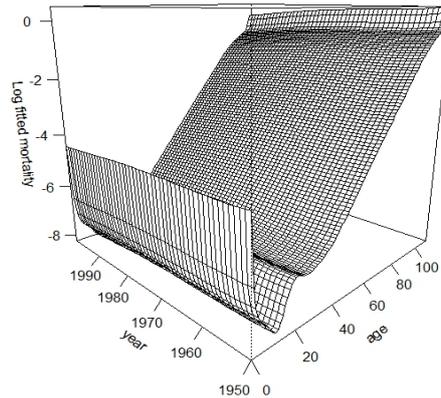


Figure 2.9: Fitted log male mortality from 1950 to 1999

Figure 2.12 displays three dimensional plot of fitted log mortality rates of Norwegian male from the calendar year period 1950 to 1999. This is the smoothed plot

of the observed log mortality rates graphed in figure 1.8. Both plots show similar pattern of distribution of the mortality rates over the intervals of age and year.

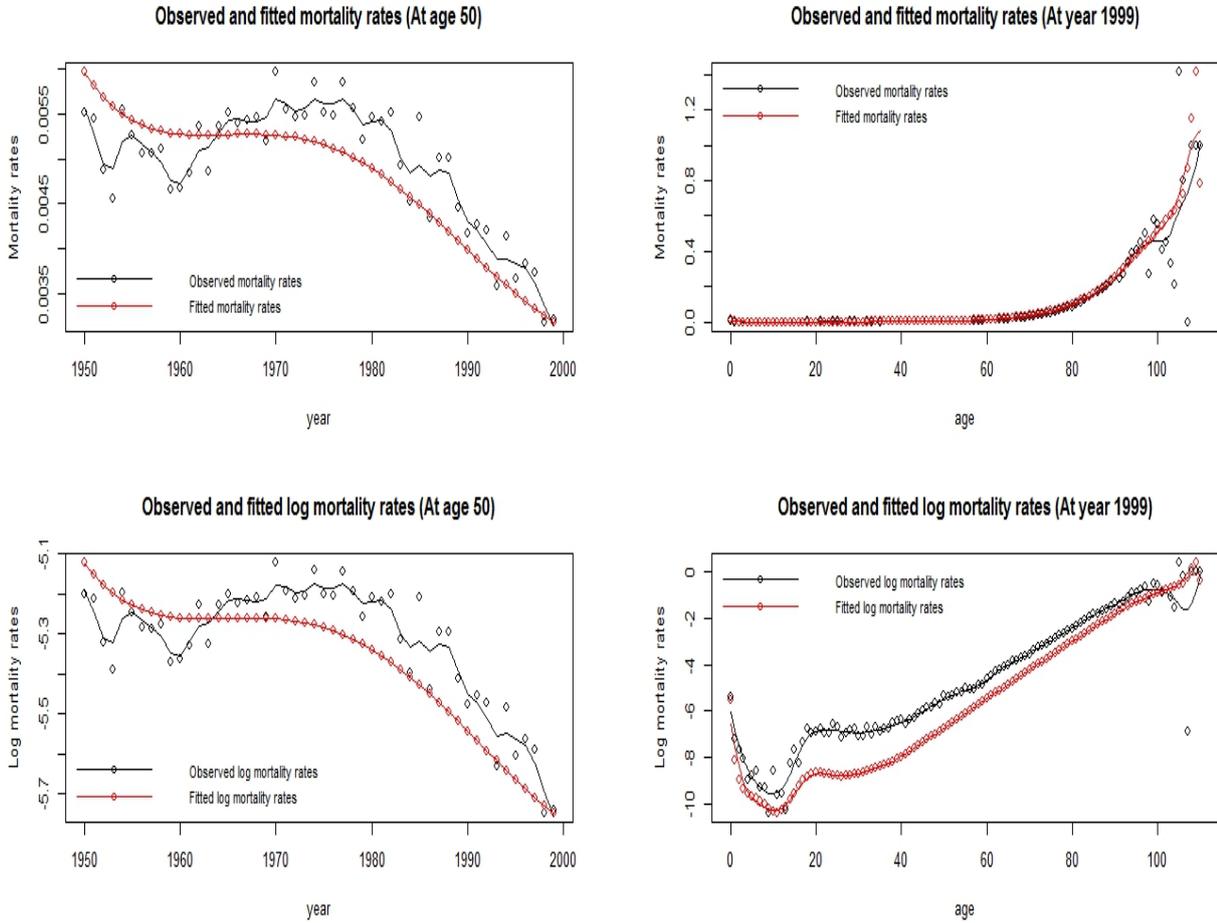


Figure 2.10: Observed and fitted mortality and log mortality

Figure 2.13 displays four plots of two dimensions of observed and fitted mortality and log mortality for the 50 age group in the calendar year 1999. The fitted rates of mortality follow similar year and age-specific patterns of the observed rates. We can see considerable departures of the fitted trend from the historical trend through age and year. These departures are clearly displayed on the logarithmic scale. The fitted and observed log mortality rates decline roughly downward with increasing years for the 50 age group and increase along with increasing age in the year 1999. The fitted trend departs from the observed trend from 1950's for the 50 age group and from 0 age group upto 100 year group in 1999. The model also fails to capture the few large rates after the 100 age group. Hence, we can say that the fitted Negative binomial regression model does not sufficiently capture the observed mortality trends through age and time for the 50 age group in the year 1999.

The Lee-Carter Model

In this chapter, Lee-Carter demographic model will be used to model patterns of trends and levels of the observed Norwegian male mortality data described in subsection 1.1.5. The model is primarily based on the past experience of the age-specific of the data.

3.1 Basic concepts

The model proposed by Ronald Lee and Lawrence Carter in 1992 for mortality forecasting assumes that the force of mortality $m_x(t)$ at age x in calendar year t is of the form $\exp\{a_x + b_x\kappa_t + \xi_x(t)\}$. Hence,

$$\log m_x(t) = a_x + b_x\kappa_t + \xi_x(t) \quad (3.1)$$

where, $t = t_1 + 1, t_1 + 2, \dots, t_n - t_1 + 1 = T$ and $x = x_1, x_2, x_3, \dots, x_t$. $m_x(t)$ is the probability of death at age x in year t . a_x is the age-specific parameter which describes the general pattern of mortality by age. b_x is the age-specific parameter which describes the time trend for the general mortality. κ_t is a time-varying mortality index. $\xi_x(t)$ are the error terms capturing particular age-specific historical influences not explained by the model. These errors are assumed $N(0, \sigma^2)$. To achieve a unique solution for the equation 3.1 the following restrictions are used;

$$a_x = \frac{1}{T} \sum_t \log m_x(t), \sum_x b_x = 1, \sum_t \kappa_t = 0 \quad (3.2)$$

3.2 Parameter Estimation

The nonparametric estimate of the age-specific death rates is given by the ratio:

$$\hat{m}_t(x) = \frac{D_{x,t}}{N_{x,t}} \quad (3.3)$$

where $D_{x,t}$, denotes the number of deaths at age x during year t from a corresponding exposure-to-risk $N_{x,t}$. The plot of $\hat{m}_t(x)$ is displayed in figure 1.8. a_x is estimated as the mean of $\log m_x(t)$ over time t . The estimation of b_x and κ_t

could not be solved explicitly and the model cannot be fit with ordinary regression methods. Lee and Carter use a two-stage estimation approach. The singular value decomposition is applied to the matrix $\{\log m_x(t) - a_x\}$ to estimate b_x and κ_t , and time series of κ_t is re-estimated in the second stage. They identified estimation of the parameters b_x and κ_t does not guarantee the observed total number of deaths to be equal to the fitted total number of deaths. This is because there are less parameters than the observations. They re-estimate κ_t to get the observed number of deaths equal the fitted number of deaths, i.e.

$$D_t = \sum_x \exp(a_x + b_x \kappa_t) N_{x,t} \quad (3.4)$$

where D_t is the total number of deaths in year t and $N_{x,t}$ is the population (exposure to risk) of age x in year t. No analytic solution is available so it can only be done by searching over a range of value of κ_t .

3.2.1 Singular Value Decomposition

The parameter vector a_x equation 3.1 at section 3.1 is computed as the mean of $\log \hat{m}_x(t)$ over time t. Hence, $\hat{a}_x = \frac{1}{T} \sum_t \log \hat{m}_x(t)$. Also from equation 3.1, $b_x \kappa_t = \log m_x(t) - \hat{a}_x$. Let $Y_x(t) = \log m_x(t) - \hat{a}_x$.

If Z denotes a p x q matrix of rank r, then there is p x p orthogonal matrix U, q x q orthogonal matrix V and p x q diagonal matrix d such that;

$$Z = U d V' \quad (3.5)$$

Where $V' = (v_{ji})$ is the inverse of $V = (v_{ij})$. For $m < n$, equation 5.4 is represented in the matrix form as;

$$A = \begin{bmatrix} u_{1,1} & \dots & \dots & u_{1,p} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ u_{p,1} & \dots & \dots & u_{p,p} \end{bmatrix} x \begin{bmatrix} d_1 & \dots & 0 & \dots & 0 \\ 0 & \dots & \dots & \dots & 0 \\ \dots & \dots & d_{p-1} & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & d_p \end{bmatrix} x \begin{bmatrix} v_{1,1} & \dots & v_{1,p} & \dots & v_{1,q} \\ \dots & \dots & \dots & \dots & \dots \\ v_{p,1} & \dots & v_{p,p} & \dots & v_{p,q} \\ \dots & \dots & \dots & \dots & \dots \\ v_{q,1} & \dots & \dots & \dots & v_{q,q} \end{bmatrix}$$

If we let $Z = Y_x(t)$, where $x = 1, 2, 3, \dots, X$ and $t = 1, 2, 3, \dots, T$, then the first rank approximation of $\hat{Y}_t^1(x) = d_1 U_1(x) V_1(t) = b_x^1 \kappa_t^1$ gives;

$$\hat{b}_x^1 = (u_{1,1} \ u_{1,2} \dots u_{1,x})'$$

and

$$\hat{\kappa}_t^1 = d_1 \times (v_{1,1} \ v_{2,1} \dots v_{T,1})$$

Using the Lee-Carter constraints 3.2, b_x and κ_t estimates become;

$$\hat{b}_x^1 = \left(\frac{1}{\sum_x u_{x,1}} \right) (u_{1,1} \ u_{1,2} \dots u_{1,X})', \quad \hat{\kappa}_t^1 = \left(\frac{1}{\sum_x u_{x,1}} \right) \times d_1 \times (v_{1,1} \ v_{2,1} \dots v_{T,1})$$

Where U(x) denotes the age, d the singular value and V(t) the time components respectively.

3.3 Fitting Lee-Carter Model

Lee Carter model is fitted by approximating new matrix $\hat{Y}_x(t)$ by the product of the estimated vector parameters \hat{b}_x and $\hat{\kappa}_t$ to obtain;

$$\hat{Y}_x(t) = \hat{b}_x \hat{\kappa}_t = \begin{bmatrix} \hat{Y}_x(t_1) \\ \hat{Y}_x(t_2) \\ \dots \\ \hat{Y}_x(t_n) \end{bmatrix} \hat{b}_x, \hat{\kappa}_x = \begin{bmatrix} \hat{Y}_x(t_1) \\ \hat{Y}_x(t_2) \\ \dots \\ \hat{Y}_x(t_n) \end{bmatrix} \quad (3.6)$$

From equation 3.6, the fitted logarithm of the central death rate then becomes;

$$\log \hat{m}_x(t) = \hat{a}_x + \begin{bmatrix} \hat{Y}_x(t_1) \\ \hat{Y}_x(t_2) \\ \dots \\ \hat{Y}_x(t_n) \end{bmatrix} \hat{b}_x \quad (3.7)$$

3.4 Applying Lee-Carter Model

In this section, we apply Lee Carter model (3.1) to the Norwegian male mortality data described in subsection 1.1.3. The parameters a_x , b_x and κ_t of the model are estimated by the methods described in 3.2.1 and analysed accordingly. The residuals of the model is later examined. Re-estimation of κ_t is not considered here.

3.4.1 Fitted Lee-Carter model

We fit Lee-Carter by employing R “demography” package by Hyndman, Booth, Tickle and Maindonald (2008), a program developed for the (1x1) data format of the Human Mortality Database. This is implemented in R as shown below:

```
deaths = read.table(file = "mot1950to1999.txt", header = T, sep = "")
exps = read.table(file = "expo1950to1999.txt", header = T, sep = "")
library(demography)
year = unique(deaths$Year); nC = length(year)
age = unique(deaths$Age); nL = length(age)
deathrate = matrix((deaths$Male)/(exps$Male), nL, nC)
expos = matrix(exps$Male, nL, nC)
# We use the demogdata format
demo <- demogdata(data = deathrate, pop = expos, ages = age,
years = year, type = "mortality", label = "Norway", name = "Male", lambda = 1)
# Then we fit the Lee-Carter model
Lca <- lca(demo, max.age = 110, interpolate = T, adjust = "none")
# We extract values of  $a_x$ ,  $b_x$  and  $k_t$ 

a_x <- -Lca$ax
b_x <- -Lca$bx
k_t <- -Lca$kt
```

¹Some of the r codes for implementing the “demography” package are based on the examples provided in Charpentier (2012) and Charpentier and Dutang (2013) online manual.

age	ax	bx	age	ax	bx	Year against kt	
0	-4.435796	0.0350643	56	-4.6332	0.0079	year	k(t)
1	-6.5595	0.0428	57	-4.5324	0.0062	1950	18.6470
2	-7.1998	0.0412	58	-4.4425	0.0056	1951	4.9659
3	-7.4067	0.0400	59	-4.3424	0.0048	1952	5.3165
4	-7.6408	0.0438	60	-4.2438	0.0050	1953	4.2962
5	-7.7018	0.0420	61	-4.1394	0.0042	1954	4.6368
6	-7.7190	0.0356	62	-4.0411	0.0042	1955	0.4991
7	-7.8705	0.0397	63	-3.9363	0.0037	1956	4.2280
8	-8.0101	0.0409	64	-3.8437	0.0043	1957	4.1196
9	-8.1396	0.0425	65	-3.7518	0.0043	1958	7.4318
10	-8.1084	0.0352	66	-3.6590	0.0040	1959	6.4169
11	-8.0980	0.0278	67	-3.5671	0.0037	1960	8.7349
12	-8.1047	0.0307	68	-3.4687	0.0027	1961	12.2159
13	-7.9460	0.0268	69	-3.3757	0.0026	1962	14.8952
14	-7.7661	0.0175	70	-3.2701	0.0026	1963	21.9554
15	-7.5190	0.0175	71	-3.1671	0.0024	1964	13.0101
16	-7.1957	0.0117	72	-3.0845	0.0022	1965	14.7991
17	-6.9611	0.0112	73	-2.9884	0.0023	1966	14.6924
18	-6.7711	0.0035	74	-2.8895	0.0021	1967	14.2441
19	-6.7541	0.0059	75	-2.7994	0.0018	1968	17.1339
20	-6.7667	0.0087	76	-2.7011	0.0021	1969	22.9032
21	-6.8040	0.0117	77	-2.6047	0.0024	1970	20.8066
22	-6.7975	0.0066	78	-2.5043	0.0022	1971	18.2942
23	-6.7716	0.0053	79	-2.4107	0.0021	1972	17.3298
24	-6.7901	0.0076	80	-2.3112	0.0023	1973	18.5540
25	-6.8231	0.0082	81	-2.2130	0.0020	1974	12.5194
26	-6.8412	0.0097	82	-2.1186	0.0021	1975	13.7752
27	-6.8358	0.0085	83	-2.0173	0.0023	1976	11.6674
28	-6.8070	0.0061	84	-1.9258	0.0022	1977	7.1830
29	-6.7881	0.0064	85	-1.8403	0.0026	1978	7.6156
30	-6.7695	0.0096	86	-1.7467	0.0025	1979	10.5683
31	-6.7231	0.0076	87	-1.6458	0.0023	1980	6.9084
32	-6.6786	0.0073	88	-1.5554	0.0025	1981	6.7810
33	-6.6129	0.0081	89	-1.4768	0.0023	1982	2.1469
34	-6.5776	0.0080	90	-1.3909	0.0025	1983	2.2975
35	-6.5362	0.0089	91	-1.2840	0.0026	1984	0.6120
36	-6.4491	0.0086	92	-1.2174	0.0028	1985	6.4417
37	-6.3807	0.0084	93	-1.1322	0.0023	1986	0.4960
38	-6.3419	0.0093	94	-1.0321	0.0014	1987	3.2424
39	-6.2561	0.0067	95	-0.9458	0.0040	1988	-2.0951
40	-6.1878	0.0065	96	-0.8992	0.0035	1989	-6.9698
41	-6.1060	0.0074	97	-0.8485	0.0017	1990	-7.4413
42	-6.0013	0.0088	98	-0.8042	0.0018	1991	-20.1433
43	-5.8953	0.0092	99	-0.7188	0.0008	1992	-23.4388
44	-5.7902	0.0082	100	-0.6184	-0.0003	1993	-17.0337
45	-5.6914	0.0065	101	-0.5921	0.0025	1994	-36.8693
46	-5.6137	0.0063	102	-0.5746	0.0034	1995	-30.9126
47	-5.5311	0.0062	103	-0.4723	0.0058	1996	-45.7036
48	-5.4258	0.0069	104	-0.1462	0.0260	1997	-46.4423
49	-5.3201	0.0077	105	0.2960	0.0284	1998	-52.2166
51	-5.1189	0.0066	106	0.4077	-0.0026	1999	-50.1434
52	-5.0239	0.0067	107	0.9328	-0.0333		
53	-4.9278	0.0073	108	0.7367	-0.0115		
54	-4.8303	0.0079	109	1.8104	-0.0011		
55	-4.7190	0.0071	110	0.6931	0.0000		

Table 3.1: Estimates of a_x , b_x and k_t

Displayed in table 3.1 above are estimated values of \hat{a}_x , \hat{b}_x and \hat{k}_t parameters from the R package “demography” by Hyndman, Booth, Tickle and Maindonald (2011). The package is based on singular value decomposition method of parameter estimation described in subsection 3.2.1. R source codes and updates of the “demography” package can be found at <https://github.com/robjhyndman/demography/blob/master/R/lca.R>.

The procedures for deriving \hat{a}_x , \hat{b}_x and \hat{k}_t estimates can also be basically implemented in R; create \hat{a}_x by taking logarithms of the mortality rates, $\text{rowMeans}(\text{logmx})$, center the results with the average log mortality at a given age, $\text{logmx}-\text{ax}$, and derive \hat{k}_t , a scaling eigenvalue, and \hat{b}_x from $u(1)$, $s(1)$, $v(1)$, where usv is the singular value decomposition of the centered log mortality rates, $\text{svd}(\text{logmx}-\text{ax})$.

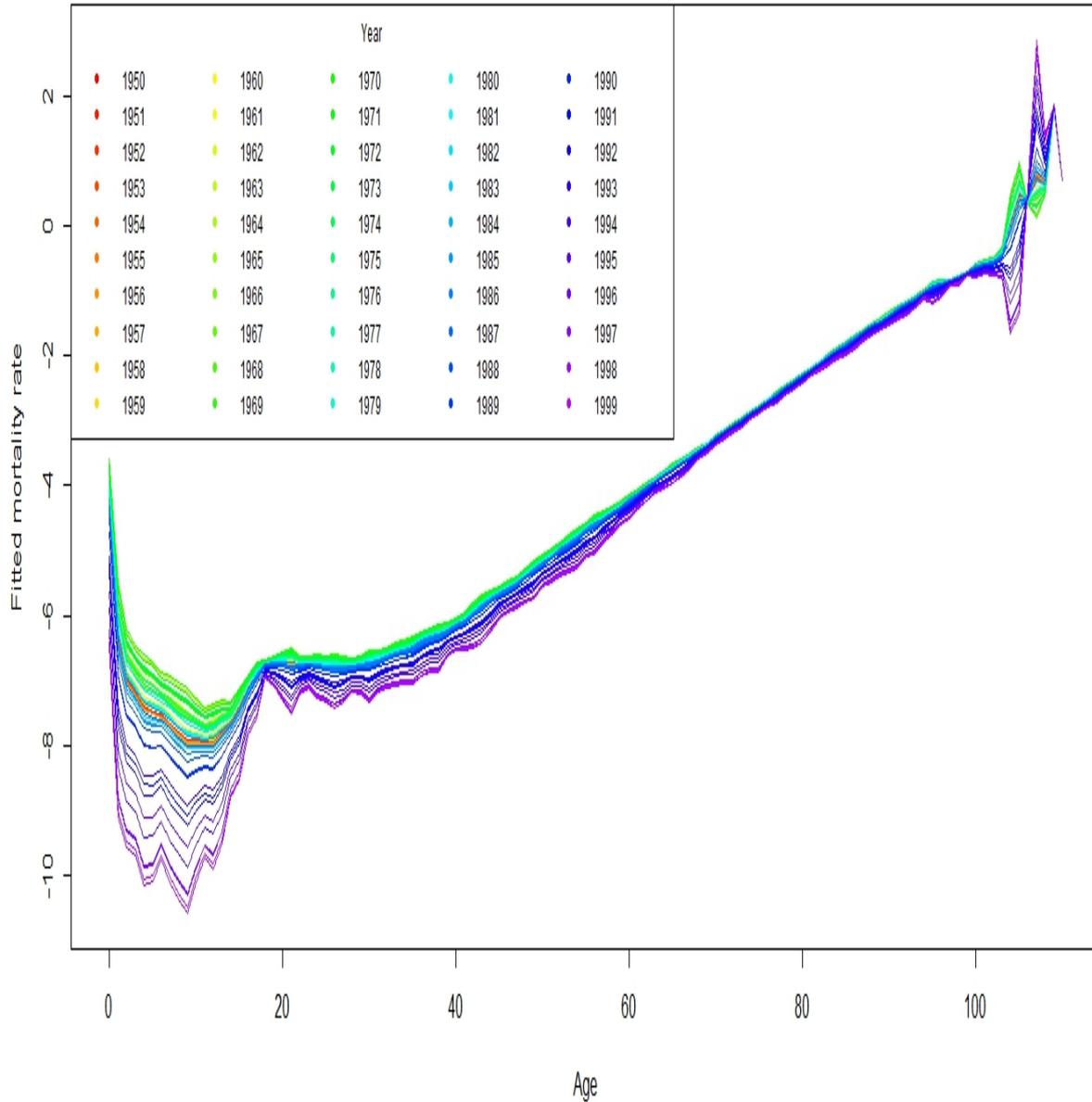


Figure 3.1: Plot of fitted mortality rate

Figure 3.1 above plots the fitted mortality rates from the year period 1950 to 1999. We see that the shape of the mortality index \hat{a}_x is well captured. The model yields a good fit on the data when $\hat{\kappa}_t$ is decreasing and also the estimated values of age-specific constant \hat{b}_x decrease with age. Predictive ability of the fitted model will be undermined if there are no continuous historical patterns of trends for these estimated parameters.

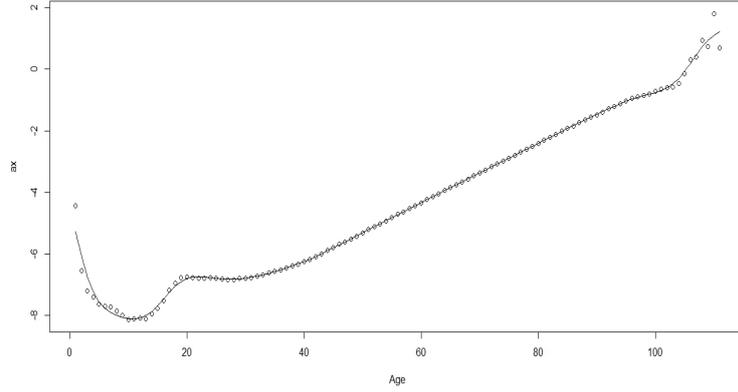


Figure 3.2: General pattern of the Norwegian male mortality (a_x)

As we expect the average mortality increases with increasing age. From table 3.1, \hat{a}_x shows a downward trend from the 0 age group at a value of (-4.4) to the 12 age group at a value of (-8.1), then rapidly increases to (-0.62) at the 100 age group and shows a relatively slight increase afterwards. Figure 3.2 plots estimated \hat{a}_x over the intervals of age, smoothed with splines of 25 degrees of freedom. We can see similar pattern for the observed mortality rates through age in the year 1999 in figure 1.9.

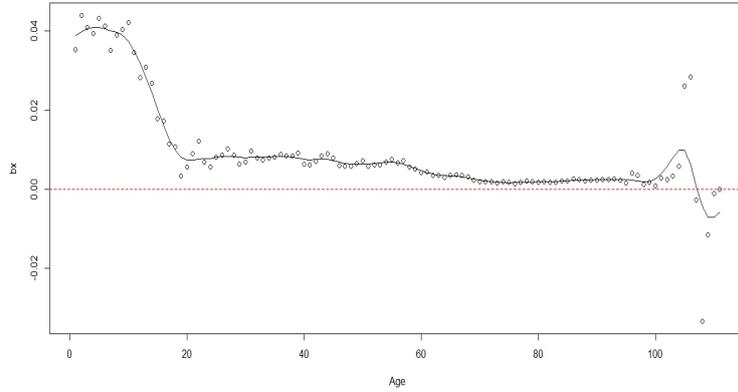


Figure 3.3: Age-specific constant(b_x) for Norwegian male mortality

The values for the estimated \hat{b}_x are shown in table 3.1. The estimated values are decreasing by age and roughly level out, which means that the mortality rates at younger ages decline more rapidly than for older ages. Higher values of \hat{b}_x for the younger ages means mortality varies significantly when $\hat{\kappa}_t$ changes. Also, lower values of this parameter for the older ages means slight variation of mortality in that period of time when $\hat{\kappa}_t$ changes. Negative values give an increasing mortality rate if the $\hat{\kappa}_t$ is negative and declining. Figure 3.3 plots the pattern of \hat{b}_x through age, smoothed with splines of 25 degrees of freedom.

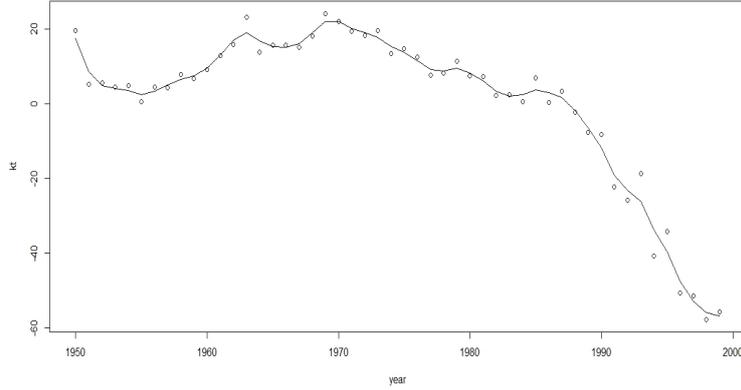


Figure 3.4: Mortality index (κ_t) for Norwegian male mortality

The values for the estimated time-varying parameter $\hat{\kappa}_t$ for Norwegian male mortality from the year period 1950 to 1999 are tabulated in table 3.1. Figure 3.4 plots these values, smoothed with splines of 25 degrees of freedom. $\hat{\kappa}_t$ captures the main time trend on the logarithmic scale of the mortality rates from the age period of 0 to 110. From the plot, $\hat{\kappa}_t$ of the mortality attains its maximum value in 1950. It then tends to fluctuate from 1951 to 1987 and declines roughly linearly from 1989 to 1999. We can see similar pattern for the observed mortality rates through time for the 50 age group in figure 1.9.

3.4.2 Goodness of fit

We validate the fitted demographic model described in the previous subsections by employing Pearson residual diagnostics. The residual terms measure unsystematic variability of the fitted model. Recalling from subsection 3.1: $\log m_x(t) = a_x + b_x \kappa_t + \xi_x(t)$, where $\xi_x(t)$ are the residual terms with zero mean and constant variance. Hence if \hat{a}_x , \hat{b}_x , $\hat{\kappa}_t$ are the estimated parameters of the fitted model, then the residual terms are estimated by: $\hat{\xi}_x(t) = \log m_x(t) - \hat{a}_x - \hat{b}_x \hat{\kappa}_t$.

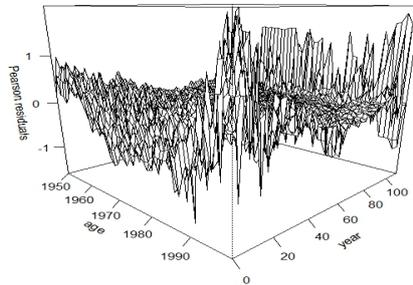


Figure 3.5: Residuals of log mortality rate using Lee-Carter

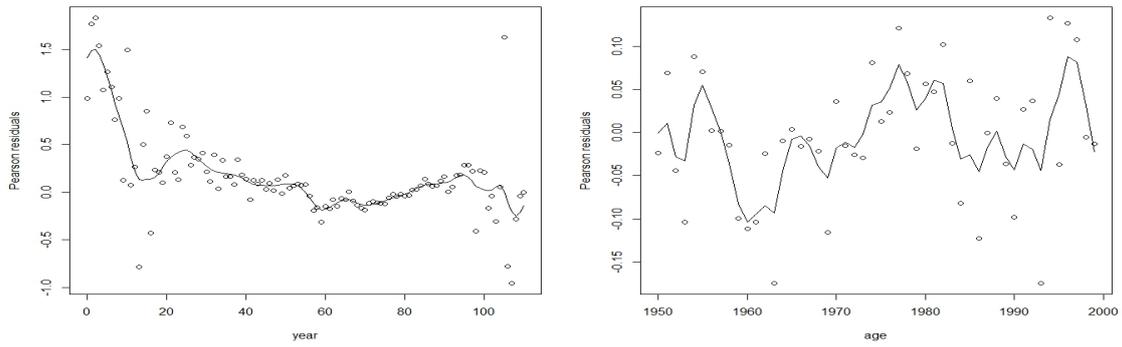


Figure 3.6: Residual plots of log mortality (at age 50 and year 1999)

Figure 3.5 shows three dimensional plots of the Pearson residual terms of the fitted Lee Carter model. Figure 3.6 depicts two dimensional plots of Figure 3.5 for 50 age group in 1999, smoothed with splines of 25 degrees of freedom. We can see figure 3.5 show more considerable clear patterns of residuals as compared to figure 2.3. In figure 3.6, residuals are randomly distributed in 1999 but decrease with time for the 50 age group. Hence, the assumption of random distribution of residual terms as expected is not entirely fulfilled. This indicates the fitted model does not adequately capture the entire historical information of the observed data.

Forecasting Techniques

The Norwegian male mortality will be forecast with the fitted Negative binomial and Lee-Carter models. The predictive performance of both models are evaluated by comparing the distributions of the observed and forecast values.

4.1 Forecasting with Negative Binomial Model

Here, the Norwegian number of deaths and their exposure to risk for the year period 2000 to 2009 described in subsection 1.1.3 are forecast with the fitted Negative binomial models in subsections 2.1.7 and 2.2.2 respectively. The forecast rates of mortality are then subsequently estimated and compared with the test set in subsection 1.1.3.

R note: Both fitted Negative binomial regression models for number of deaths and risks of exposure in subsections 2.1.7 and 2.2.2 are based on orthogonalised polynomials. As explained in subsection 2.1.5, we standardised and orthogonalised the explanatory variables age and year to eradicate multicollinearity and numerical problems. Hence we need a function in R that can create data frames of orthogonal arrays with unique attributes for the new categorical explanatory variables to enable us forecast with the `predict.glm` function using the fitted models. `expand.grid` function of the "base" R package has these features. `Expand.grid` creates the data frames from all combinations of the new categorical explanatory variables age and year with unique attributes which we use in the `predict.glm` function to get forecasts for risks of exposure and number of deaths from 2000 to 2009. Note that we also standardise the new categorical explanatory variable year j to obtain $s = 25, \dots, 34$.

4.1.1 Forecasting risks of exposure(2000 to 2009)

If n_{ij} is the number of males exposed to risk to be forecast and x_{ij} the new categorical variables age and year, then recalling from subsection 2.2.1, we obtain the equation for forecasting n_{ij} to be:

$$\log(u_{ij}) = (x_{ij})^T \hat{\beta} \implies u_{ij} = \exp\{(x_{ij})^T \hat{\beta}\}$$

$\tilde{n}_{ij} = E(n_{ij}) = u_{ij}$. $i = 0, 1, \dots, 110$ and $j = 2000, \dots, 2009$ are indicators of the vector x_{ij} for explanatory variables, for each i age and j year combination, from 2000 to 2009. $\hat{\beta}$'s are the regression coefficients.

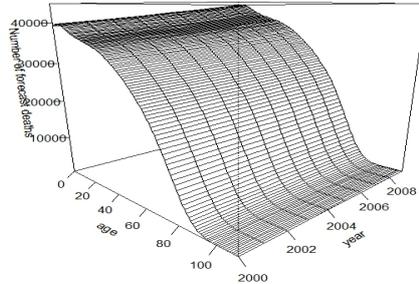


Figure 4.1: Forecasts of risk exposure (2000 to 2009)

Figure 4.1 shows smoothed plot of the observed risks of exposure displayed in figure 1.7. We can see slight variations in the structural appearances of both plots.

Goodness of forecast

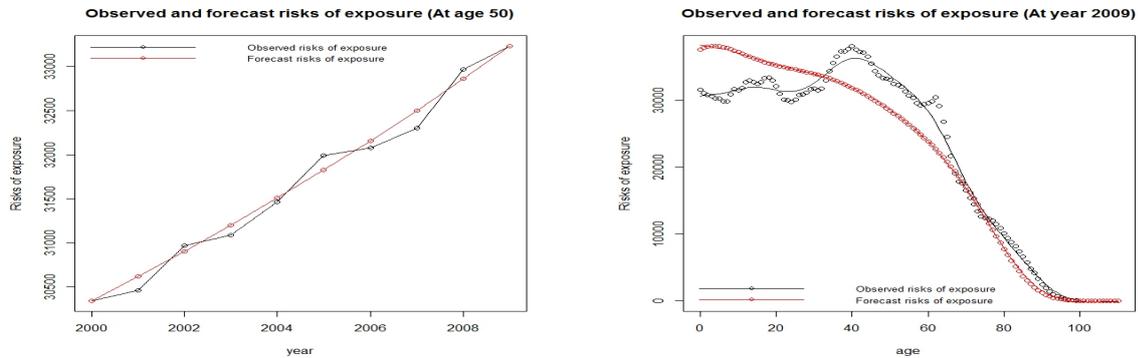


Figure 4.2: Forecast of number of deaths (at age 50 and year 2009)

Figure 4.2 shows two dimensional plots of figures 4.1 and 1.7 for the 50 age group in 2009. We can see forecast risks of exposure follow the same historical patterns through time and age for the 50 age group in 2009. Comparing this to figure 2.11, we can say the fitted model forecast well for shorter time periods but would likely not forecast well for longer time periods.

4.1.2 Forecasting number of deaths (2000 to 2009)

Let y_{ij} be the death counts of Norwegian male to be forecast and x_{ij} the new categorical variables age and year that have predictive power for y_{ij} . $i = 0, 1, \dots, 110$

and $j = 2000, \dots, 2009$ are indicators of the vector x_{ij} for explanatory variables, which is for each i age and j year combination, starting from 2000 calendar year through to 2009. Therefore recalling from subsection 2.1.1, we obtain the equations for forecasting the number of deaths y_{ij} for the year period 2000 to 2009 to be;

$$\log(u_{ij}) = \log(\tilde{n}_{ij}) + (x_{ij})^T \hat{\beta}$$

$$\implies u_{ij} = \exp \{ \log(\tilde{n}_{ij}) + (x_{ij})^T \hat{\beta} \}$$

Where the $\hat{\beta}$'s are the regression coefficients in subsection 2.1.7, based on orthogonalised polynomials. \tilde{n}_{ij} are the forecast values of risks of exposure in subsection 4.1.1. The expected forecast values are then obtained as, $\tilde{y}_{ij} = E(y_{ij}) = u_{ij}$.

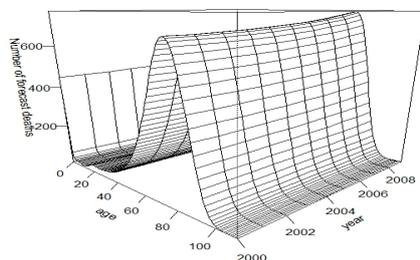


Figure 4.3: Forecast of number of deaths (2000 to 2009)

Displayed in figure 4.3 is the smoothed plot of the test set of number of deaths in figure 1.4. Structurally, both plots look alike, just that the forecast death counts have significantly few high values compared to the observed.

Goodness of forecast

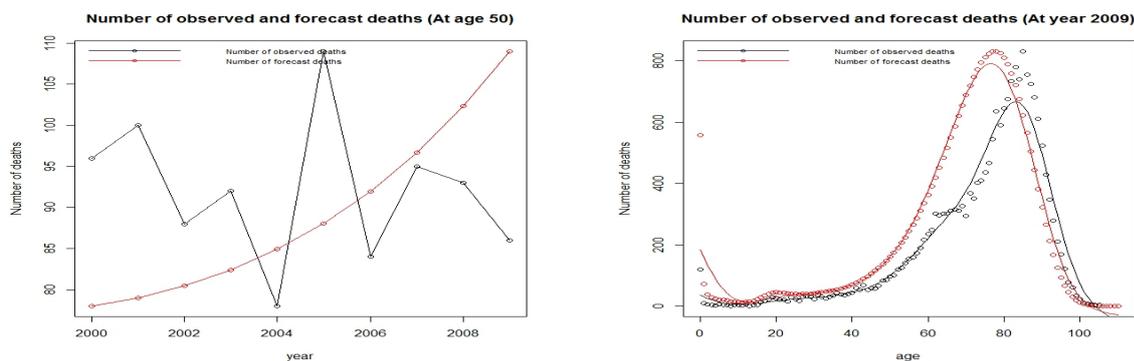


Figure 4.4: Forecast of number of deaths (at age 50 and year 2009)

Figure 4.2 depicts two dimensional plots of figures 1.4 and 4.1 for 50 age group in 2009. As we can see from the plot, number of forecast deaths depart significantly

from the observed over the time trend of the 50 age group and between 40 to 80 age groups in 2009. Too high mortality rates are forecast for the 70's age group. The time trend would likely get worse as time increases. These are departures from the fitted trends displayed in figure 2.6. Hence the fitted model gives poor predictive performance.

4.1.3 Forecasting rates of mortality (2000 to 2009)

If \tilde{y}_{ij} and \tilde{n}_{ij} are the forecast values for number of deaths and risks of exposure for the Norwegian males obtained in subsections 4.1.2 and 4.1.2 respectively, then the values for the forecast rates of mortality are fitted by the ratio:

$$\begin{aligned}\tilde{m}_{ij} &= \frac{\tilde{y}_{ij}}{\tilde{n}_{ij}} = \exp\{x_{ij}^T \hat{\beta}\} \\ \implies \log(\tilde{m}_{ij}) &= \log\left(\frac{\tilde{y}_{ij}}{\tilde{n}_{ij}}\right) = (x_{ij})^T \hat{\beta}\end{aligned}$$

Where i, j and x_{ij} are as defined in subsections 4.1.1 and 4.1.2. The $\hat{\beta}$'s are also the regression coefficients in table 2.4 based on orthogonalised polynomials.

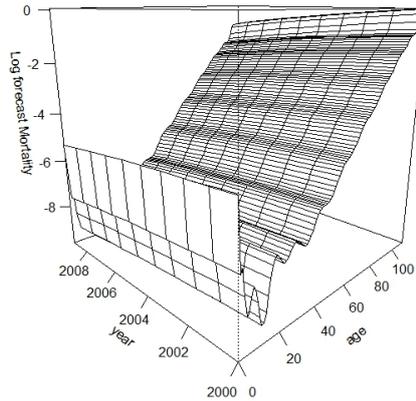


Figure 4.5: Forecast mortality rates(2000 to 2009)

Figure 4.5 displays the smooth plot of observed mortality rates in figure 1.10. Both plots follow similar patterns of distribution of the mortality rates against year and time.

Goodness of forecast

Figure 4.6 below shows four plots of two dimensions of observed and forecast mortality and log mortality rates for the 50 age group in the year 2009. As we can observe, the fitted model does not sufficiently relate the forecasts to the historical trend of the observed rates through time for the 50 age group. Similar problem

is noted for the fit in figure 2.13. The forecasts deviate considerably from the historical trend of the observed rates between the 5 and 70 age groups. The model also gives poor predictive performance for the few large mortality rates after the 100 age group. Similar problem is also noted in figure 2.13.

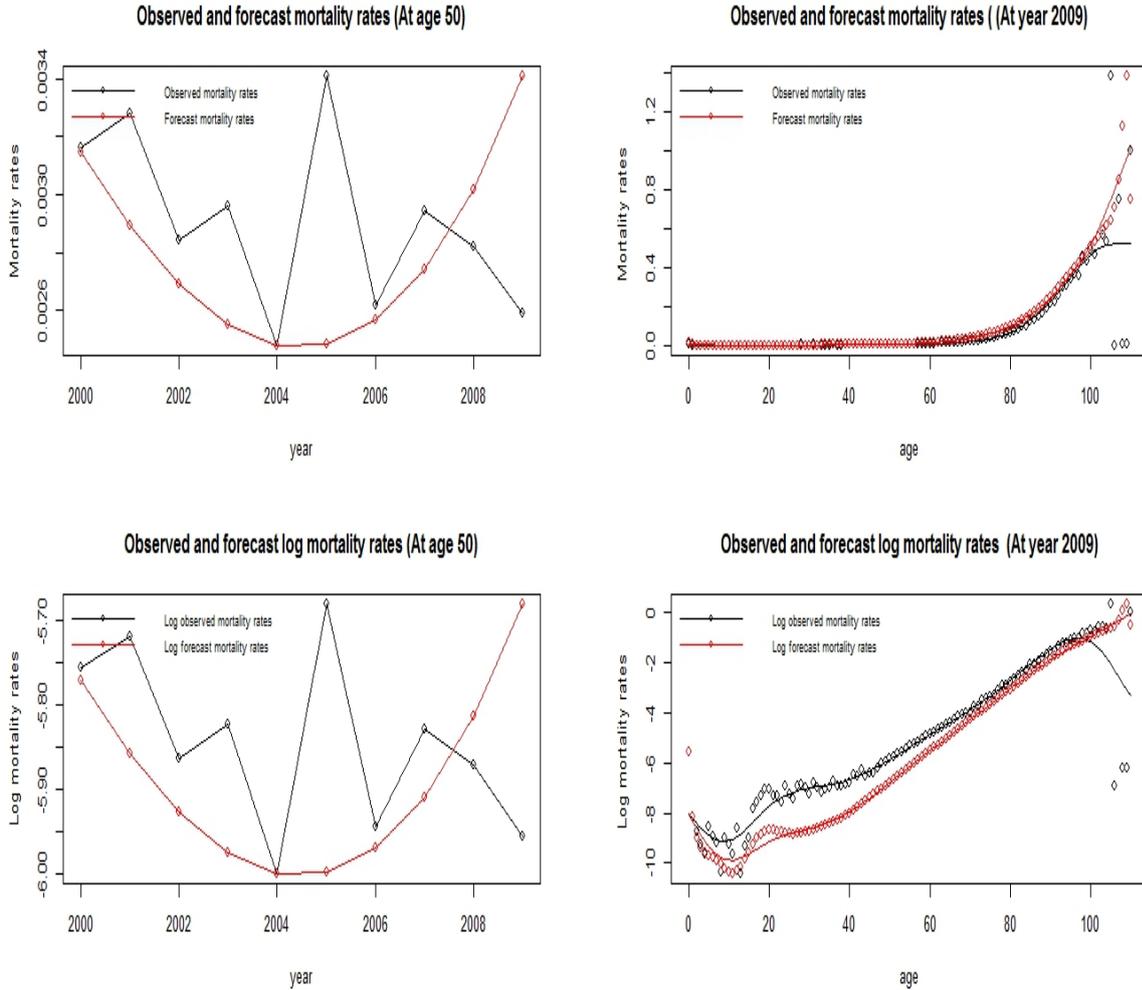


Figure 4.6: Observed and forecast mortality and log mortality

4.2 Forecasting with Lee-Carter Model

In this section, Lee-Carter demographic model fitted in chapter three is used to forecast the test set described in section 1.1.3 of chapter one. According to Lee and Carter (1992), modeling the dynamics of the fitted mortality index $\hat{\kappa}_t$ is key for the equation 3.3 in section 3.2 to hold in order to forecast. They proposed standard univariate ARIMA (0,1,0) time series model as appropriate model for modeling the dynamics of the index $\hat{\kappa}_t$. We will apply this model to the estimated mortality index, examine its quality of fit and make a forecast for $\hat{\kappa}_t$ from 2000 to 2009.

4.2.1 Basic concepts of standard ARIMA (0,1,0) model

The standard univariate ARIMA (0,1,0) time series model for κ_t is a random walk with a drift δ , and has its dynamics expressed as:

$$\kappa_t - \kappa_{t-1} = \delta + \xi_t \quad (4.1)$$

where $t = t_1 + 1, t_1 + 2, \dots, t_n - t_1 + 1 = T$. The drift parameter δ gives the mean annual change in κ_t , i.e $\delta = \nabla \kappa_t = \kappa_t - \kappa_{t-1}$. ξ_t are identically distributed error terms with $N(0, \sigma^2)$. From equation 4.1 implies:

$$\kappa_t = \kappa_{t-1} + \delta + \xi_t \quad (4.2)$$

To forecast κ_t two steps ahead, we obtain, $\kappa_t = (\kappa_{t-2} + \delta + \xi_{t-1}) + \delta + \xi_t$, which implies $\kappa_t = \kappa_{t-2} + 2\delta + (\xi_t + \xi_{t-1})$. Therefore to forecast κ_t h steps ahead, we obtain, $\kappa_{T+h} = \kappa_T + h\delta + (\xi_{T+1} + \dots + \xi_{T+h})$, with expectation, $E(\kappa_{T+h} / \kappa_T) = \kappa_T + h\delta$

Fitting of the ARIMA (0,1,0) model to $\hat{\kappa}_t$

Here, we fit the ARIMA (0,1,0) model to the fitted $\hat{\kappa}_t$ in table 3.1 with "astsa" R package by Shumway and Stoffer(2010). We obtain the drift term $\delta = -1.54$, with standard error 0.87, indicating downward trend as expected. Hence, equation 4.2 becomes $\kappa_t = \kappa_{t-1} - 1.54 + \xi_t = -1.54 + \sum_t \xi_t$

Goodness of fit

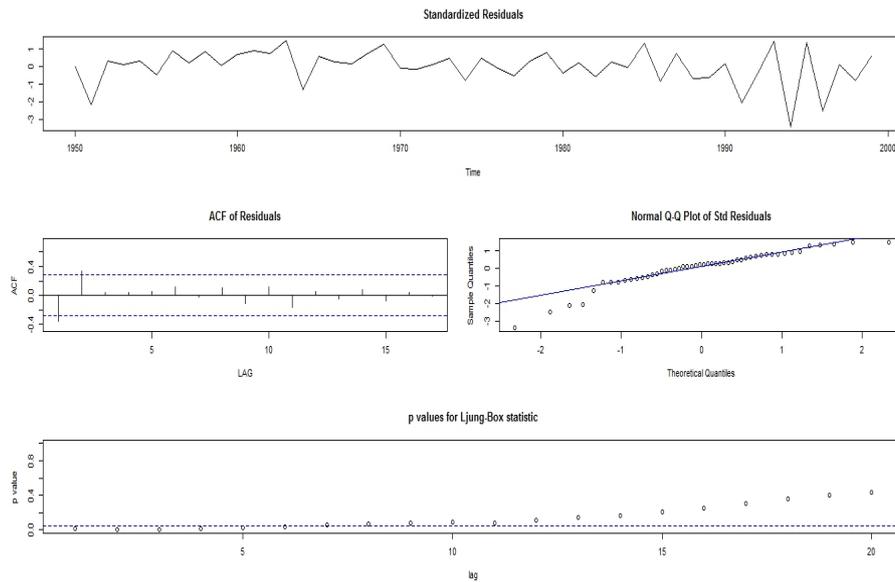


Figure 4.7: Diagnostic plots of refitted index κ_t

We validate the fitted ARIMA (0,1,0) model with diagnostic plots of residuals displayed in figure 4.7. Apparently, there is stationarity in the standardized residual

plot. The expected pattern of normality is deviated at the lower end of the Q-Q plot. In addition, there are presence of significant autocorrelations in the ACF plot and low p-values for Ljung Box statistic. This in effect indicates that ARIMA (0,1,0) model as proposed by Lee and Carter (1992) does not fully capture the historical dynamics of the Norwegian mortality index $\hat{\kappa}_t$.

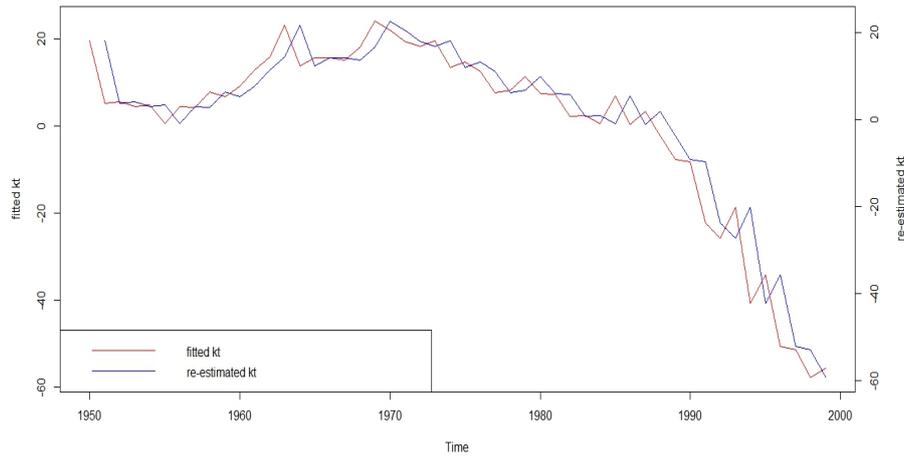


Figure 4.8: Plot of fitted and re-estimated κ_t with ARIMA(0,1,0)

Figure 4.8 plots fitted and re-estimated κ_t from 1950 to 1999. Both plots follow the same declining historical pattern. There are considerable variations in the trends and levels of both series.

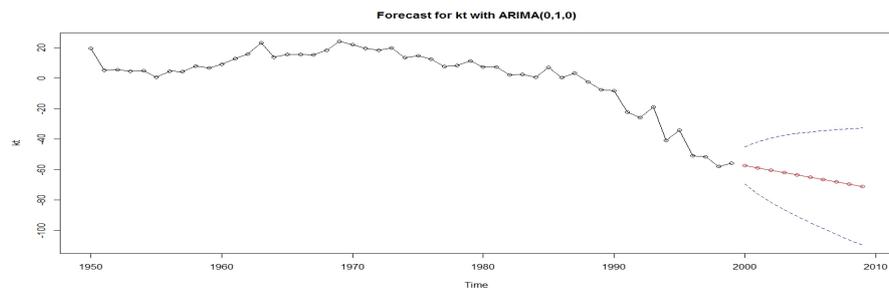


Figure 4.9: Forecast of κ_t from 2000 to 2009 with ARIMA(0,1,0)

We can see from figure 4.9 above that the trend of the forecast index $\hat{\kappa}_t$ is in line with the declining pattern of re-estimated index over the intervals of time. Also, the forecast $\hat{\kappa}_t$ is smoothed and linear through time with 95% prediction intervals. This linearity trend of $\hat{\kappa}_t$ is what we expect to forecast mortality rates with the Lee-Carter model.

4.2.2 Forecasting rates of mortality (2000 to 2009)

If \hat{a}_x and \hat{b}_x are the fitted values for the age-specific parameters tabulated in table 3.1 and $\hat{\kappa}_t$ denotes the forecast values with the ARIMA (0,1,0) model, then the forecast rates of mortality from the year period 2000 to 2009 are obtained from equation 3.1 by:

$$\begin{aligned}\tilde{m}_x(t) &= \exp\{\hat{a}_x + \hat{b}_x \hat{\kappa}_t\} \\ \implies \log \tilde{m}_x(t) &= \{\hat{a}_x + \hat{b}_x \hat{\kappa}_t\}\end{aligned}$$

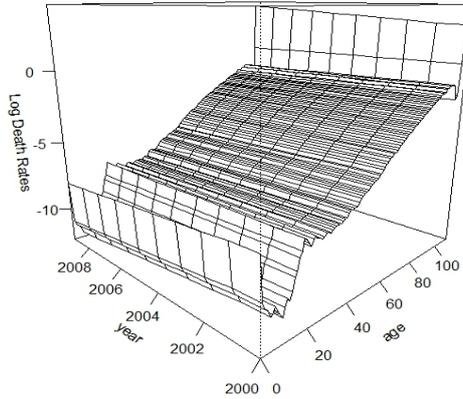


Figure 4.10: Forecast rates of mortality (2000 to 2009)

Displayed in figure 4.10 is the smoothed plot of observed mortality rates in figure 1.10. The plot shows similar distribution patterns of forecast log mortality rates in figure 4.5. Figure 4.10 covers more negative rates as compared to figure 4.5.

Goodness of forecast

Figure 4.11 depicts four plots of two dimensions of mortality and log mortality forecasts for the 50 age group in the year 2009 for the Lee-Carter and Negative binomial regression model. The three plots are smoothed with splines of 10 degrees of freedom. The downward linear time trend of the forecast for the 50 age group is due to the linearity of the re-estimated mortality index κ_t . We already noted statistical insignificance of κ_t in subsection 4.2.1. We observe polynomial time trend for the Negative binomial regression model, which we described in subsection 2.1.5. Both models give poor prediction through time for the 50 age group.

The forecasts for both models produce the general pattern of age-specific constant a_x displayed in figure 3.2. Lee-Carter model roughly captures the historical trends of the log observed rates from 10 to 40 age groups and departs significantly afterwards in 2009. On the other hand, Negative binomial model deviates from the historical

trend of the log observed rates between 5 and 70 age groups and after 100 age group in the same year period.

Both models in one way or the other, do not fully relate the historical patterns of trends and levels of the observed rates and log observed rates to their respective forecast rates and log forecast rates for the 50 age group in the year 2009. At this juncture, we cannot statistically draw conclusion on which of the models gives good forecast, since the comparison of the forecasts is based on only one age group and year period. But we can suggest the fitted Negative binomial model seems to outperform its fitted Lee-Carter competitor for the 50 age group and the year period 2009. Nevertheless, we will determine the best model that gives good forecast to the Norwegian male mortality by examining their forecasting accuracy through the distribution and measurement of their respective forecast errors.

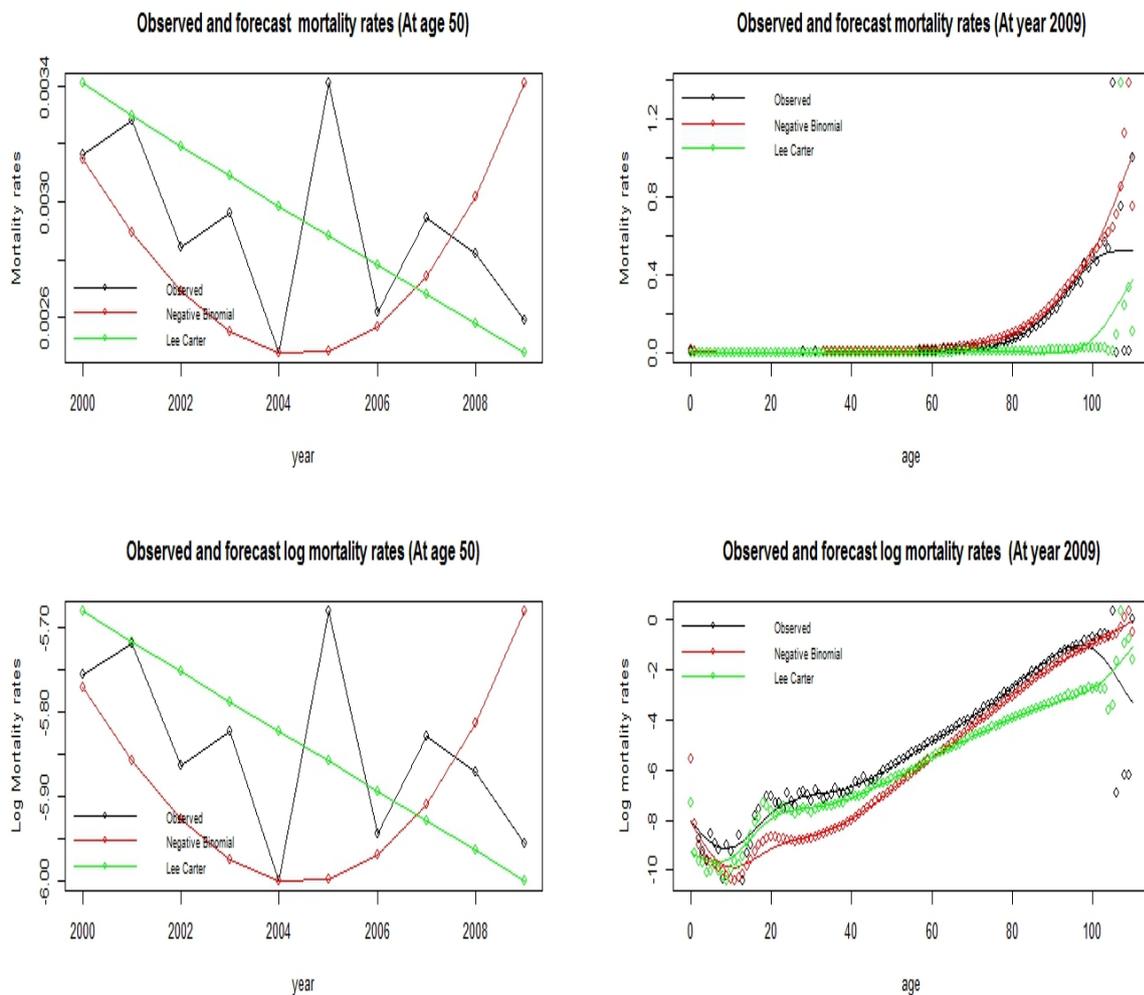


Figure 4.11: Observed and forecast mortality

Forecast errors

We used Negative binomial regression and Lee-Carter models fitted in chapter two to forecast Norwegian mortality rates from the year calendar year period 2000 to 2009 in chapter four. The question now is which of the two models gives high accuracy of forecasts? This is what this chapter seeks to address by comparing their forecast errors.

5.1 Basic concepts of forecast error

Forecast error is estimated as the difference between the observed value and the forecast value for the the corresponding year period. If \tilde{m}_{ij} is the value of forecast rate at age i in year j combination and m_{ij} is the corresponding observed mortality rate, then the forecast error e_{ij} is defined as:

$$e_{ij} = m_{ij} - \tilde{m}_{ij} \quad (5.1)$$

In this case $i = 0, 1, \dots, 110$ and $j = 2000, \dots, 2009$. We can conclude from equation 5.1 that, positive e_{ij} implies $m_{ij} > \tilde{m}_{ij}$, negative e_{ij} implies $m_{ij} < \tilde{m}_{ij}$ and neutral e_{ij} implies $m_{ij} = \tilde{m}_{ij}$. We will assess the distribution of error terms recorded by Negative binomial and Lee-Carter models and subsequently quantify these errors with Mean Squared (MSE) and Mean Absolute Error (MAE) measures.

5.2 Distribution of forecast errors

Figure 5.1 displays three dimensional plots of the forecast errors recorded over the intervals of age and year by the fitted Negative binomial regression and Lee-Carter demographic models respectively. Most of the forecast errors recorded by Negative binomial regression model are considerably positive. We can see large positive and few negative values after the 100 age group for Negative binomial regression model. On the contrary, Lee-Carter model records large negative and few positive values after the 100 age group. Comparably, there are higher concentration of forecast errors recorded by the fitted Lee-Carter model after the 100 age group than those recorded by the fitted Negative binomial model. Nevertheless, both models deviate

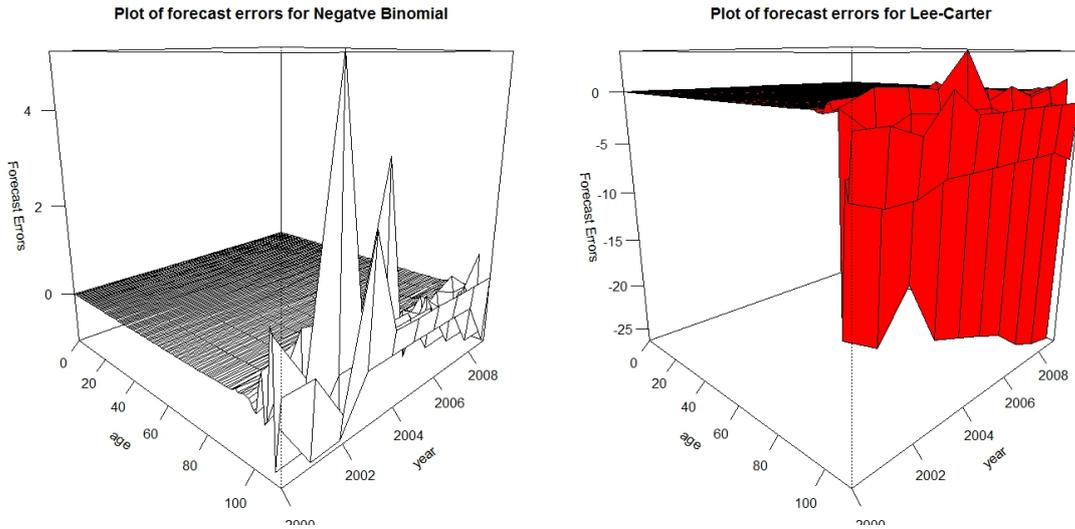


Figure 5.1: Distribution patterns of forecast errors

very significantly in forecasting mortality rates for the highest attained age groups of the Norwegian males.

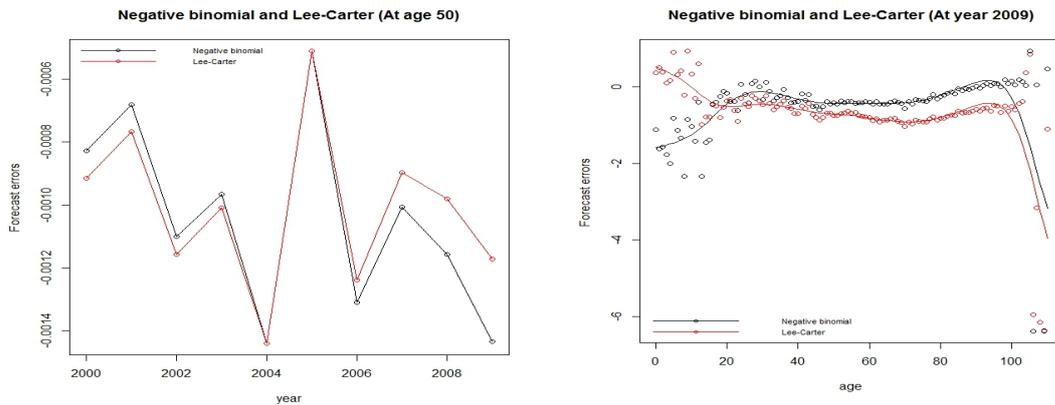


Figure 5.2: Distribution of forecast errors(at age 50 and year 2009)

Figure 5.2 compares trends of the estimated forecast errors of the Negative binomial regression and Lee- Carter models for the 50 age group in 2009. Both plots are smoothed with splines of 25 degrees of freedom. We can see variations in the trends of the forecast errors for the two models in the year period 2000 to 2003 and 2006 to 2009. Both models forecast almost with the same accuracy from 2003 to 2006. Most of the estimated forecast errors against age for the Negative binomial regression model are closer to zero as compared to those of the Lee Carter model. This indicates that the forecast values of Negative binomial are more closer to the observed data as compared to the Lee-Carter model. Thus we can say Negative binomial regression model gives better forecasting accuracy than the Lee-Carter model for the 50 age group in 2009.

5.3 Measurement of forecast errors

In this section, we use Mean Squared Error (MSE) and Mean Absolute Error (MAE) methods to quantify the forecast errors recorded by both Negative binomial regression and Lee-Carter models in figure 5.1.

5.3.1 Basic concepts of Mean Squared Error (MSE)

The mean squared error measures the mean of the squares of the forecast errors. Hence, from equation 5.1, we can define the mean squared error as:

$$MSE = \frac{1}{n} \sum_{i,j} e_{ij}^2 \quad (5.2)$$

5.3.2 Basic concepts of Mean Absolute Error (MAE)

The mean absolute error measures the mean of the absolute forecast errors. From equation 5.1, we can define the mean absolute error as:

$$MAE = \frac{1}{n} \sum_{i,j} |e_{ij}| \quad (5.3)$$

We can say from equations 5.2 and 5.3 that MAE is more sensitive to small forecast errors and much less sensitive to large forecast errors than MSE.

5.3.3 Comparison of Sum of Squares of forecast errors

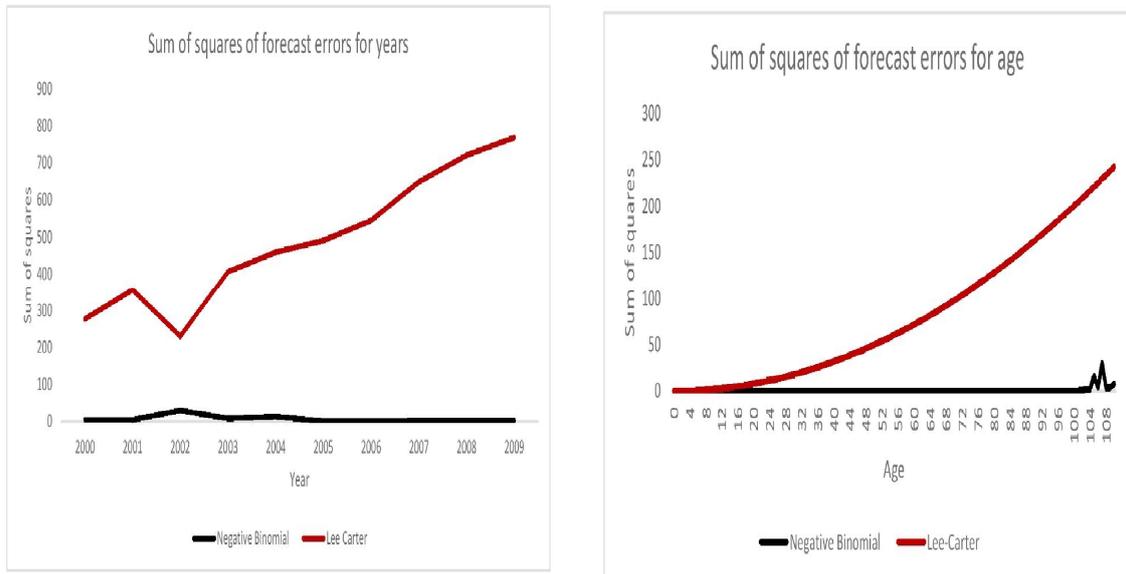


Figure 5.3: Sum of squares of forecast errors

Figure 5.3 displays sum of squares of forecast errors recorded by Negative and Lee Carter models for the year and age groups. We see Lee Carter records high values

of sum of squares for both year and age groups compared to Negative binomial regression model. This is the result of the high negative errors recorded by Lee Carter in figure 5.1, which inflate the sum of squares.

5.3.4 Comparison of MSE and MAE of forecast errors

Displayed in table 5.1 below are the estimated Mean Squared Error (MSE) and Mean Absolute Error (MAE) of the forecast errors recorded by the fitted Negative binomial and Lee-Carter models respectively, using R "metrics" package. Computations of MSE and MAE can also be implemented in R with the following codes:

$$mse <- \text{function}(e)\{\text{mean}(e^2)\}; mse(e)$$

$$mae <- \text{function}(e)\{\text{mean}(\text{abs}(e))\}; mae(e)$$

	MSE	MAE
Fitted Negative Binomial Model	0.064	0.052
Fitted Lee Carter model	2.415	0.318

Table 5.1: Summary of forecast errors

It is clear from table 5.1 that the forecasting accuracy of the Norwegian male mortality rates from 2000 to 2009 by the fitted Negative binomial regression model far outweighs that of the fitted Lee-Carter model.

Negative binomial model has 0.0642 for MSE whilst Lee-Carter model has 2.415, implying Lee-Carter model deviates from Negative binomial model by a difference of 2.35 in order to forecast with the same accuracy as the Negative binomial model. The high MSE estimated for Lee-Carter is due to the high values of the sum of squares of the forecast errors displayed in figure 5.3.

Lee-Carter has a low value of 0.318 for MAE as compared to the MSE. This is due to the absolute nature of the MAE, since Lee Carter recorded mostly negative forecast errors. On the contrary, Negative binomial has a lower value of 0.052 for MAE as compared to Lee-Carter. This produces a difference of 0.266 in MAE for both models, implying that the Lee-Carter model needs a reduction of 0.266 in its MAE to forecast with the same accuracy as Negative binomial regression model.

5.4 Other options of Negative Binomial Regression Model

This subsection seeks to assess whether some of the Negative binomial regression models specify in subsection 2.1.7 have better forecasting accuracy than the fitted Negative binomial regression model.

Fitted Negative Binomial Model	p=25	q=4	0.064	0.052
Negative Binomial Model	p=25	q=3	0.065	0.057
Negative Binomial Model	p=25	q=2	0.065	0.055
Negative Binomial Model	p=25	q=1	0.063	0.051

Table 5.2: Summary of forecast errors

Table 5.2 displays summary of MSE and MAE for some of the Negative binomial regression models specified in table 2.2 of subsection 2.1.7. Note that the polynomial orders of p and q for age and years specified above are used to model number of deaths only. Polynomial orders of p and q for age and years specified in the fitted negative binomial regression model for risks of exposure in subsection 2.2.2 still remain the same.

We see that Negative binomial model with polynomial order $p = 25$ for age and $q = 1$ for year has the lowest estimates for MAE and MSE when used for forecasting the mortality. This model outperforms the fitted Negative binomial model in subsection 2.1.7 although it records the highest AIC value in table 2.2. This is strange but true because "incorporating more polynomial terms with unnecessarily high orders of p and q for age and year permits an increasingly complicated response structure with fitted values close to the observations but low predictive strength", Piet and Gillian (2008). Hence, simple polynomial function is recommendable in modeling the time trend with Negative binomial regression model to obtain tradeoff between good fit and forecasts for the Norwegian male mortality.

Conclusion

In this chapter, we present summary and conclusion of the thesis. We also present weakness identified for future work regarding the topic.

6.1 Summary Remarks

This thesis sought to compare Poisson or Negative binomial regression and original Lee-Carter demographic models in forecasting Norwegian male mortality. We noted that Negative binomial regression model overruled its Poisson competitor due to rejection of equidispersion hypothesis to confirm overdispersion in the mortality data. We proceeded to fit Negative binomial regression model to the mortality and risks of exposure data to estimate mortality rates for the Norwegian males from the year period 1950 to 1999. We observed Negative binomial regression model tended to capture the historical trends of the observed data for the age groups but departed significantly from the time trends. Comparative evaluations of the observed and fitted mortality rates for the 50 age group in 1999 are of no exceptions. Lee Carter model on the other hand was noted to provide poor fit to the mortality data based on assessment of the distribution of residual terms. Validating the fits of both models with the 50 age group in the year 1999, we observed Negative binomial regression model indicated better fit as compared to the Lee-Carter model.

We used the fitted Negative binomial regression and Lee-Carter models to forecast the mortality rates for the year period 2000 to 2009. We noted that Negative binomial regression model indicated strong predictive power over the Lee-Carter model based on comparative assessment of the forecast and observed rates for the 50 age group in 2009. The poor predictive performance of the Lee-Carter model is basically attributable to inability of the reestimated mortality index to adequately project age-specific mortality rates for the forecasting period. This problem was moderately minimal in Negative binomial regression model because of the simple polynomial functions allowed for the period.

The comparative assessment of the distributions of forecast error terms of the two models indicated that Lee- Carter model under-predicted the mortality data. We

noted that the Lee-Carter model recorded on average Negative errors. This resulted in inflation of the mean squares leading to the high MSE estimate as compared to the lower MSE estimate for the Negative binomial regression model. However, we noted substantial reduction in the MSE estimate for Lee-Carter model using MAE measurement technique, which is based on absolute values.

Although both Lee-Carter and Negative binomial regression models deviated in many ways in capturing the future trends of the mortality data, at the end Negative binomial regression model outperforms its Lee-Carter competitor in forecasting the Norwegian male mortality.

6.2 Known Weakness and Suggestion

We noted that the Lee-Carter model gave poor predictive performance because of statistical insignificance of the mortality index. This suggests that Lee-Carter model could have predicted well if the dynamics of the mortality index had been modeled with any appropriate stochastic model proposed in any of the extensions of the Lee-Carter model rather than the random walk model.

We also noted that Negative binomial regression model gave high predictive accuracy when simple polynomial functions were allowed for the period. Hence we recommend that simple polynomial functions should be allowed in Negative binomial regression model for consistent forecasting of Norwegian male mortality.

Both models departed significantly in providing good fits and forecasts for the high attained age groups of the Norwegian males. This is due to the missing actual death counts within these age groups. This suggests having high quality of data within these groups could have improved the fits and predictive performance of the models. Hence we suggest models with strong fitting and predictive power for high thresholds and jump-offs for these groups.

Appendix

Provided below are the R codes used for the entire thesis.

```
# Uploading the data files to R platform
rm(list=ls())
ls()
setwd("C:\\Users\\Emmanuel\\Desktop\\serious\\odd")
deaths=read.table("mot1950to1999.txt",header=T)
setwd("C:\\Users\\Emmanuel\\Desktop\\serious\\odd")
death.09=read.table("mot2000to2009.txt",header=T)
setwd("C:\\Users\\Emmanuel\\Desktop\\serious\\odd")
exps=read.table("expo1950to1999.txt",header=T)
setwd("C:\\Users\\Emmanuel\\Desktop\\serious\\odd")
exps.09=read.table("expo2000to2009.txt",header=T)

CHAPTER ONE
# Three dimensional plot for Number of Deaths(1950-1999)
Deaths<-matrix(deaths$Male,111,50,byrow=F)
persp(seq(0,110),seq(1950,1999),Deaths,ticktype = "detailed",theta=45,xlab="age",ylab="year",zlab="Deaths")

# observed death counts (At age 50 and year 1999)
par(mfrow=c(1,2))
plot(seq(1950,1999),Deaths[50,],xlab="year",ylab="Deaths")
lines(smooth.spline(seq(1950,1999),Deaths[50,],df=25), col = "black")
plot(seq(0,110),Deaths[,50],xlab="age",ylab="Deaths")
lines(smooth.spline(seq(0,110),Deaths[,50],df=25), col = "black")

# Three dimensional plot for Number of Deaths (2000-2009)
Death.09<-matrix(death.09$Male,111,50,byrow=F)
persp(seq(0,110),seq(1950,1999),Death.09,ticktype = "detailed",theta=45,xlab="age",ylab="year",zlab="Deaths")

# Three dimensional plot for risks of exposure (1950-1999)
Exposures<-matrix(exps$Male,111,50,byrow=F)
persp(seq(0,110),seq(1950,1999),Exposures,ticktype = "detailed",theta=45,xlab="age",ylab="year",zlab="Exposures")

# observed risks of exposure (At age 50 and year 1999)
par(mfrow=c(1,2))
plot(seq(1950,1999),Exposures[50,],xlab="year",ylab="Deaths")
lines(smooth.spline(seq(1950,1999),Exposures[50,],df=25), col = "black")
plot(seq(0,110),Exposures[,50],xlab="age",ylab="Deaths")
lines(smooth.spline(seq(0,110),Exposures[,50],df=25), col = "black")

# Three dimensional plot for risks of exposure (2000-2009)
Exposure.09<-matrix(exps.09$Male,111,50,byrow=F)
persp(seq(0,110),seq(1950,1999),Exposure.09,ticktype = "detailed",theta=45,xlab="age",ylab="year",zlab="Exposures")

#Persp for mortality rates (1950-1999)
lmort<-matrix(log(deaths$Male/(exps$Male)),111,50,byrow=F)
persp(seq(0,110),seq(1950,1999),lmort,ticktype = "detailed",theta=-45,xlab="age",ylab="year",zlab="Log Death Rates")

## observed mortality rates (At age 50 and year 1999)
par(mfrow=c(1,2))
plot(seq(1950,1999),lmort[50,],xlab="year",ylab="Deaths")
lines(smooth.spline(seq(1950,1999),lmort[50,],df=25), col = "black")
plot(seq(0,110),lmort[,50],xlab="age",ylab="Deaths")
lines(smooth.spline(seq(0,110),lmort[,50],df=25), col = "black")

#Persp for observed mortality rates (2000-2009)
lmort.09<-matrix(log(death.09$Male/(exps.09$Male)),111,50,byrow=F)
persp(seq(0,110),seq(1950,1999),lmort.09,ticktype = "detailed",theta=-45,xlab="age",ylab="year",zlab="Log Death Rates")

CHAPTER TWO
#Fitted Poisson regression model for number of deaths
ddths<-deaths$Male
age1<-deaths$Age
year1<-deaths$Year
year1<-year1-1975
lexpa<-log(exps$Male)
library(MASS)
mod1<-glm(ddths~poly(age1,25)+ poly(year1,4)+lexpa,link="log")

#Fitted Negative Binomial regression model for number of deaths
ddths<-deaths$Male
age1<-deaths$Age
year1<-deaths$Year
year1<-year1-1975
lexpa<-log(exps$Male)
library(MASS)
mod1<-glm.nb(ddths~poly(age1,25)+ poly(year1,4)+lexpa,link="log")

#Residual and normality plots for fitted number of deaths
par(mfrow=c(1,2))
res<-matrix(resid(mod1,type="deviance"),111,50,byrow=F)
persp(seq(0,110),seq(1950,1999),res,ticktype = "detailed",theta=-45,xlab="age",ylab="year",zlab="Residual deviance")
qqnorm(residuals(mod1,type="deviance"));qqline(residuals(mod1,type="deviance"),col="red")

# Residual deviance (At age 50 and year 1999)
par(mfrow=c(1,2))
plot(seq(1950,1999),res[50,],xlab="year",ylab="Residual deviance")
```

```

lines(smooth.spline(seq(1950,1999),res[50,],df=25), col = "black")
plot(seq(0,110),res[,50],xlab="age",ylab="Residual deviance")
lines(smooth.spline(seq(0,110),res[,50],df=25), col = "black")

# Three dimensional plot for fitted deaths
fit<-matrix(mod1$fit,111,50,byrow=F)
persp(seq(0,110),seq(1950,1999),fit,ticktype = "detailed",theta=45,xlab="age",ylab="year",zlab="Fitted Deaths")

# observed and fitted deaths (At age 50 and year 1999)
obs.ddths<-matrix(deaths$Male,111,50,byrow=F)
par(mfrow=c(1,2))
par(mar=c(5,4,4,5)+.1)
plot(seq(1950,1999),obs.ddths[50,],xlab="year",ylab="Number of deaths")
lines(smooth.spline(seq(1950,1999),obs.ddths[50,],df=25), col = "black")
par(new=TRUE)
plot(seq(1950,1999),fit[50,],xaxt="n",yaxt="n",xlab="",ylab="",col = "red")
lines(smooth.spline(seq(1950,1999),fit[50,],df=25), col = "red")
legend("bottomleft",col=c("black","red"),legend=c("Observed number of deaths","Fitted number of deaths"),pch=1,lty=1, cex=0.7,bty= "n")
title("Observed and fitted male deaths (At age 50)")
par(mar=c(5,4,4,5)+.1)
plot(seq(0,110),obs.ddths[,50],xlab="age",ylab="Number of deaths")
lines(smooth.spline(seq(0,110),obs.ddths[,50],df=25), col = "black")
par(new=TRUE)
plot(seq(0,110),fit[,50],xaxt="n",yaxt="n",xlab="",ylab="",col = "red")
lines(smooth.spline(seq(0,110),fit[,50],df=25), col = "red")
legend("topleft",col=c("black","red"),legend=c("Observed number of deaths","Fitted number of deaths"),pch=1,lty=1, cex=0.7,bty= "n")
title("Observed and fitted male deaths (At year 1999)")

# Fitted Negative Binomial Regression for exposure
expo<-exps$Male
age<-exps$Age
year<-exps$Year
yearc<-year-1975
library(MASS)
mod2<-glm.nb(round(expo,0)-poly(age,10)+ poly(yearc,3),link="log")

#Residual and normality plots for fitted exposure
par(mfrow=c(1,2))
res2<-matrix(resid(mod2,type="deviance"),111,50,byrow=F)
persp(seq(0,110),seq(1950,1999),res2,ticktype = "detailed",theta=45,xlab="age",ylab="year",zlab="Residual deviance")
qqnorm(residuals(mod2,type="deviance"));qqline(residuals(mod2,type="deviance"),col="red")

# Residual deviance of fitted exposure (At age 50 and year 1999)
par(mfrow=c(1,2))
plot(seq(1950,1999),res2[50,],xlab="year",ylab="Residual deviance")
lines(smooth.spline(seq(1950,1999),res2[50,],df=25), col = "black")
plot(seq(0,110),res2[,50],xlab="age",ylab="Residual deviance")
lines(smooth.spline(seq(0,110),res2[,50],df=25), col = "black")

# Three dimensional plot for fitted exposure
fit2<-matrix(mod2$fit,111,50,byrow=F)
obs.exps<-matrix(exps$Male,111,50,byrow=F)
persp(seq(0,110),seq(1950,1999),fit2,ticktype = "detailed",theta=45,xlab="age",ylab="year",zlab="Fitted Exposure")

# observed and fitted exposure (At age 50 and year 1999)
par(mfrow=c(1,2))
par(mar=c(5,4,4,5)+.1)
plot(seq(1950,1999),obs.exps[50,],xlab="year",ylab="Risks of exposure")
lines(smooth.spline(seq(1950,1999),obs.exps[50,],df=25), col = "black")
par(new=TRUE)
plot(seq(1950,1999),fit2[50,],xaxt="n",yaxt="n",xlab="",ylab="",col = "red")
lines(smooth.spline(seq(1950,1999),fit2[50,],df=25), col = "red")
legend("topleft",col=c("black","red"),legend=c("Observed risks of exposure","Fitted risks of exposure"),pch=1,lty=1, cex=0.8,bty= "n")
title("Observed and fitted risks of exposure (At age 50)")
par(mar=c(5,4,4,5)+.1)
plot(seq(0,110),obs.exps[,50],xlab="age",ylab="Risks of exposure")
lines(smooth.spline(seq(0,110),obs.exps[,50],df=25), col = "black")
par(new=TRUE)
plot(seq(0,110),fit2[,50],xaxt="n",yaxt="n",xlab="",ylab="",col = "red")
lines(smooth.spline(seq(0,110),fit2[,50],df=25), col = "red")
legend("bottomleft",col=c("black","red"),legend=c("Observed risks of exposure","Fitted risks of exposure"),pch=1,lty=1, cex=0.8,bty= "n")
title("Observed and fitted risks of exposure (At year 1999)")

# Fitted mortality rates
fit<-matrix(mod1$fit,111,50,byrow=F)
obs.ddths<-matrix(deaths$Male,111,50,byrow=F)
fit2<-matrix(mod2$fit,111,50,byrow=F)
obs.exps<-matrix(exps$Male,111,50,byrow=F)
odeath.rate<-matrix((ddths/expo),111,50,byrow=F)
ldeath.rate<-matrix(log(ddths/expo),111,50,byrow=F)
fdeath.rate<-matrix((fit/fit2),111,50,byrow=F)
fldeath.rate<-matrix(log(fit/fit2),111,50,byrow=F)

# Three dimensional plot for fitted log mortality
persp(seq(0,110),seq(1950,1999),fldeath.rate,ticktype = "detailed",theta=45,xlab="age",ylab="year",zlab="Fitted log mortality")

# observed and fitted mortality and log mortality (At age 50 and year 1999)
par(mfrow=c(2,2))
par(mar=c(5,4,4,5)+.1)
plot(seq(1950,1999),odeath.rate[50,],xlab="year",ylab="Mortality rates")

```

```

lines(smooth.spline(seq(1950,1999),odeath.rate[50,],df=25), col = "black")
par(new=TRUE)
plot(seq(1950,1999),fdeath.rate[50,],xaxt="n",yaxt="n",xlab="",ylab="",col = "red")
lines(smooth.spline(seq(1950,1999),fdeath.rate[50,],df=25), col = "red")
title("Observed and fitted mortality rates (At age 50)")
legend("bottomleft",col=c("black","red"),legend=c("Observed mortality rates","Fitted mortality rates"),pch=1,lty=1, cex=0.8,bty= "n")
par(mar=c(5,4,4,5)+.1)
plot(seq(0,110),odeath.rate[,50],xlab="age",ylab="Mortality rates")
lines(smooth.spline(seq(0,110),odeath.rate[,50],df=25), col = "black")
par(new=TRUE)
plot(seq(0,110),fdeath.rate[,50],xaxt="n",yaxt="n",xlab="",ylab="",col = "red")
lines(smooth.spline(seq(0,110),fdeath.rate[,50],df=25), col = "red")
legend("topleft",col=c("black","red"),legend=c("Observed mortality rates","Fitted mortality rates"), pch=1,lty=1, cex=0.8,bty= "n")
title("Observed and fitted mortality rates (At year 1999)")
par(mar=c(5,4,4,5)+.1)
plot(seq(1950,1999),ldeath.rate[50,],xlab="year",ylab="Log mortality rates")
lines(smooth.spline(seq(1950,1999),ldeath.rate[50,],df=25), col = "black")
par(new=TRUE)
plot(seq(1950,1999),fldeath.rate[50,],xaxt="n",yaxt="n",xlab="",ylab="",col = "red")
lines(smooth.spline(seq(1950,1999),fldeath.rate[50,],df=25), col = "red")
title("Observed and fitted log mortality rates (At age 50)")
legend("bottomleft",col=c("black","red"),legend=c("Observed log mortality rates","Fitted log mortality rates"),pch=1,lty=1, cex=0.8,bty= "n")
par(mar=c(5,4,4,5)+.1)
plot(seq(0,110),ldeath.rate[,50],xlab="age",ylab="Log mortality rates")
lines(smooth.spline(seq(0,110),ldeath.rate[,50],df=25), col = "black")
par(new=TRUE)
plot(seq(0,110),fldeath.rate[,50],xaxt="n",yaxt="n",xlab="",ylab="",col = "red")
lines(smooth.spline(seq(0,110),fldeath.rate[,50],df=25), col = "red")
legend("topleft",col=c("black","red"),legend=c("Observed log mortality rates","Fitted log mortality rates"),pch=1,lty=1, cex=0.8,bty= "n")
title("Observed and fitted log mortality rates (At year 1999)")

CHAPTER THREE
#Fitting Lca From 1950-1999
library(demography)
year=unique(deaths$Year);nC=length(year)
age =unique(deaths$Age);nL=length(age)
deathrate=matrix(deaths$Male/exps$Male,nL,nC)
expos=matrix(exps$Male,nL,nC)
# Then we use the demogdata format
demo <- demogdata(data=deathrate, pop=expos, ages=age, years=year, type="mortality",label="Norway", name="Male", lambda=1)
Lca<-lca(demo,max.age=110,interpolate=T, adjust="none")

# Extracting ax, bx and kt parameters
ax<-Lca$ax
bx<-Lca$bx
kt<-Lca$kt

#Plotting the fitted Lee-Carter model
plot(Lca$fit)
legend("topleft",legend=unique(Lca$year),col=rainbow(length(Lca$year)*1.25), ncol=5, pch=19, title="Year", cex=0.5)

#Plotting ax
plot(ax,xlab="Age")
lines(smooth.spline(ax,df=25), xlab="Age",col = "black")

#Plotting bx
plot(bx,xlab="Age")
lines(smooth.spline(bx,df=25), xlab="Age",col = "black")
abline(h=0, lty=2, col="red")

#Plotting kt
plot(kt,xlab="Year")
lines(smooth.spline(kt,df=25), xlab="Year",col = "black")

# Three dimensional plot of Pearson residuals (Lee-Carter)
RES=residuals(Lca,"pearson")
persp(RES$x,RES$y,RES$z,ticktype = "detailed",theta=45,xlab="age",ylab="year",zlab="Pearson residuals")

## Plot of Pearson residuals (At age 50 and year 1999)
par(mfrow=c(1,2))
plot(RES$y,RES$z[50,],xlab="year",ylab="Pearson residuals")
lines(smooth.spline(RES$y,RES$z[50,],df=25), col = "black")
plot(RES$x,RES$z[,50],xlab="age",ylab="Pearson residuals")
lines(smooth.spline(RES$x,RES$z[,50],df=25), col = "black")

CHAPTER FOUR
#Forecasting risks of exposure (2000-2009)
agef<-exps.09$Age
yearf<-exps.09$Year
yearf<-yearf-1975
grid <- expand.grid(age=unique(agef), yearc=unique(yearf),KEEP.OUT.ATTRS = TRUE))
forecast2<-predict(mod2, newdata=grid)

# Three dimensional plot of forecast risks of exposure
exposures<-matrix((exps.09$Male),111,10,byrow=F)
for2<-matrix(exp(forecast2),111,10,byrow=F)
persp(seq(0,110),seq(2000,2009),for2,ticktype = "detailed",theta=-45,xlab="age",ylab="year",zlab="Log mortality")

#Observed and forecast risks of exposure (At age 50 and year 2009)
par(mfrow=c(1,2))

```

```

par(mar=c(5,4,4,5)+.1)
plot(seq(2000,2009),exposures[50,],xlab="year",ylab="Risks of exposure")
lines(smooth.spline(seq(2000,2009),exposures[50,],df=10), col = "black")
par(new=TRUE)
plot(seq(2000,2009),for2[50,],xaxt="n",yaxt="n",xlab="",ylab="",col = "red")
lines(smooth.spline(seq(2000,2009),for2[50,],df=10), col = "red")
legend("topleft",col=c("black","red"),legend=c("Observed risks of exposure","Forecast risks of exposure"),pch=1,lty=1, cex=0.8,bty= "n")
title("Observed and forecast risks of exposure (At age 50)")
par(mar=c(5,4,4,5)+.1)
plot(seq(0,110),exposures[,10],xlab="age",ylab="Risks of exposure")
lines(smooth.spline(seq(0,110),exposures[,10],df=10), col = "black")
par(new=TRUE)
plot(seq(0,110),for2[,10],xaxt="n",yaxt="n",xlab="",ylab="",col = "red")
lines(smooth.spline(seq(0,110),for2[,10],df=10), col = "red")
legend("bottomleft",col=c("black","red"),legend=c("Observed risks of exposure","Forecast risks of exposure"),pch=1,lty=1, cex=0.8,bty= "n")
title("Observed and forecast risks of exposure (At year 2009)")

#Prediction for deaths from 2000-2009
lexpo<-log(forecast2)
age2<-death.09$Age
yearp<-death.09$Year
year2<-yearp-1975

grid1 <- expand.grid(age1=unique(age2), year1=unique(year2),lexpa=unique(lexpo),KEEP.OUT.ATTRS = TRUE)
forecast1<-predict(mod1, newdata=grid1)

# Three dimensional plot of forecast deaths
for1<-matrix(exp(forecast1),111,10,byrow=F)
persp(seq(0,110),seq(2000,2009),for2,ticktype = "detailed",theta=45,xlab="age",ylab="year",zlab="Number of forecast deaths")

#Observed and forecast deaths (At age 50 and year 2009)
Dead<-matrix((death.09$Male),111,10,byrow=F)
par(mfrow=c(1,2))
par(mar=c(5,4,4,5)+.1)
plot(seq(2000,2009),Dead[50,],xlab="year",ylab="Number of deaths")
lines(smooth.spline(seq(2000,2009),Dead[50,],df=10), col = "black")
par(new=TRUE)
plot(seq(2000,2009),for1[50,],xaxt="n",yaxt="n",xlab="",ylab="",col = "red")
lines(smooth.spline(seq(2000,2009),for1[50,],df=10), col = "red")
legend("topleft",col=c("black","red"),legend=c("Number of observed deaths","Number of forecast deaths"),pch=1,lty=1, cex=0.7,bty= "n")
title("Number of observed and forecast deaths (At age 50)")
par(mar=c(5,4,4,5)+.1)
plot(seq(0,110),Dead[,10],xlab="age",ylab="Number of deaths")
lines(smooth.spline(seq(0,110),Dead[,10],df=10), col = "black")
par(new=TRUE)
plot(seq(0,110),for1[,10],xaxt="n",yaxt="n",xlab="",ylab="",col = "red")
lines(smooth.spline(seq(0,110),for1[,10],df=10), col = "red")
legend("topleft",col=c("black","red"),legend=c("Number of observed deaths","Number of forecast deaths"),pch=1,lty=1, cex=0.7,bty= "n")
title("Number of observed and forecast deaths (At year 2009)")

#Forecast mortality rates
est.mot<-matrix(exp(forecast1)/exp(forecast2),111,10,byrow=F)
obs.mot<-matrix((death.09$Male)/(exps.09$Male),111,10,byrow=F)
lest.mot<-matrix(log(exp(forecast1)/exp(forecast2)),111,10,byrow=F)
lobs.mot<-matrix(log((death.09$Male)/(exps.09$Male)),111,10,byrow=F)

#Observed and forecast mortality and log mortality (At age 50 and year 2009)
par(mfrow=c(2,2))
par(mar=c(5,4,4,5)+.1)
plot(seq(2000,2009),obs.mot[50,],xlab="year",ylab="Mortality rates")
lines(smooth.spline(seq(2000,2009),obs.mot[50,],df=10), col = "black")
par(new=TRUE)
plot(seq(2000,2009),est.mot[50,],xaxt="n",yaxt="n",xlab="",ylab="",col = "red")
lines(smooth.spline(seq(2000,2009),est.mot[50,],df=10), col = "red")
title("Observed and forecast mortality rates (At age 50)")
legend("topleft",col=c("black","red"),legend=c("Observed mortality rates","Forecast mortality rates"),pch=1,lty=1, cex=0.7,bty= "n")
par(mar=c(5,4,4,5)+.1)
plot(seq(0,110),obs.mot[,10],xlab="age",ylab="Mortality rates")
lines(smooth.spline(seq(0,110),obs.mot[,10],df=10), col = "black")
par(new=TRUE)
plot(seq(0,110),est.mot[,10],xaxt="n",yaxt="n",xlab="",ylab="",col = "red")
lines(smooth.spline(seq(0,110),est.mot[,10],df=10), col = "red")
legend("topleft",col=c("black","red"),legend=c("Observed mortality rates","Forecast mortality rates"),pch=1,lty=1, cex=0.7,bty= "n")
title("Observed and forecast mortality rates ( At year 2009)")
par(mar=c(5,4,4,5)+.1)
plot(seq(2000,2009),lobs.mot[50,],xlab="year",ylab="Log mortality rates")
lines(smooth.spline(seq(2000,2009),lobs.mot[50,],df=10), col = "black")
par(new=TRUE)
plot(seq(2000,2009),lest.mot[50,],xaxt="n",yaxt="n",xlab="",ylab="",col = "red")
lines(smooth.spline(seq(2000,2009),lest.mot[50,],df=10), col = "red")
title("Observed and forecast log mortality rates (At age 50)")
legend("topleft",col=c("black","red"),legend=c("Log observed mortality rates","Log forecast mortality rates"),pch=1,lty=1, cex=0.7,bty= "n")
par(mar=c(5,4,4,5)+.1)
plot(seq(0,110),lobs.mot[,10],xlab="age",ylab="Log mortality rates")
lines(smooth.spline(seq(0,110),lobs.mot[,10],df=10), col = "black")
par(new=TRUE)
plot(seq(0,110),lest.mot[,10],xaxt="n",yaxt="n",xlab="",ylab="",col = "red")
lines(smooth.spline(seq(0,110),lest.mot[,10],df=10), col = "red")
legend("topleft",col=c("black","red"),legend=c("Log observed mortality rates","Log forecast mortality rates"),pch=1,lty=1, cex=0.7,bty= "n")
title("Observed and forecast log mortality rates (At year 2009)")

```

```

#Forecasting with Lee-Carter Model

# Re-estimating kt
require(astsa)
sarima<-sarima(kt, p =0,d = 1, q = 0)
kxt<-sarima$fit
forecast<-sarima.for(kxt, n.ahead = 10, p =0,d = 1, q = 0)
title("Forecast for kt with ARIMA(0,1,0)")

#Random walk approach
rdw<-rwf(kt,drift=T, level=c(0,1,0))

#Plot of fitted and restimated kt
res_kt<-sarima$fitted.values
y1 <- kt
y2 <- res_kt
par(mar=c(5,4,4,5)+.1)
plot(year, y1,type="l",col="red",ylab="fitted kt")
par(new=TRUE)
plot(y2,type="l",col="blue",xaxt="n",yaxt="n",xlab="",ylab="")
axis(4)
mtext("re-estimated kt",side=4,line=3)
legend("bottomleft",col=c("red","blue"),lty=1,legend=c("fitted kt","re-estimated kt"))

# Deriving mortality rates
for_kt<-cbind(-57.23450,-58.77321,-60.31192,-61.85063,-63.38934,-64.92805,-66.46676,-68.00548,-69.54419,-71.08290)
u<-exp(ax+bx%*%for_kt)

#Three dimensional plot of log forecast mortality
lest.mot<-matrix(log(u),111,10,byrow=F)
persp(seq(0,110),seq(2000,2009),lest.mot,ticktype = "detailed",theta=45,xlab="age",ylab="year",zlab="Number of forecast deaths")

# all combined forecast mortality and log mortality (At age 50 and year 2009)
obs.mot<-matrix((death.09$Male+0.001)/(exps.09$Male+0.001),111,10,byrow=F)
lobs.mot<-matrix(log((death.09$Male)/(exps.09$Male)),111,10,byrow=F)
nest.mot<-matrix((exp(forecast1)/exp(forecast2)),111,10,byrow=F)
lnest.mot<-matrix(log(exp(forecast1+0.001)/exp(forecast2+0.001)),111,10,byrow=F)
est.mot<-matrix(u,111,10,byrow=F)
lest.mot<-matrix(log(u),111,10,byrow=F)

par(mfrow=c(2,2))
par(mar=c(5,4,4,5)+.1)
plot(seq(2000,2009),obs.mot[50,],xlab="year",ylab="Mortality rates")
lines(smooth.spline(seq(2000,2009),obs.mot[50,],df=10), col = "black")
par(new=TRUE)
plot(seq(2000,2009),nest.mot[50,],xaxt="n",yaxt="n",xlab="",ylab="", col = "red")
lines(smooth.spline(seq(2000,2009),nest.mot[50,],df=10), col = "red")
par(new=TRUE)
plot(seq(2000,2009),est.mot[50,],xaxt="n",yaxt="n",xlab="",ylab="",col = "green")
lines(smooth.spline(seq(2000,2009),est.mot[50,],df=10), col = "green")
title("Observed and forecast mortality rates (At age 50)")
legend("bottomleft",col=c("black","red","green"),legend=c("Observed","Negative Binomial","Lee Carter"),pch=1,lty=1, cex=0.7,bty= "n")
par(mar=c(5,4,4,5)+.1)
plot(seq(0,110),obs.mot[,10],xlab="age",ylab="Mortality rates",col = "black")
lines(smooth.spline(seq(0,110),obs.mot[,10],df=10), col = "black")
par(new=TRUE)
plot(seq(0,110),nest.mot[,10],xaxt="n",yaxt="n",xlab="",ylab="",col = "red")
lines(smooth.spline(seq(0,110),nest.mot[,10],df=10), col = "red")
par(new=TRUE)
plot(seq(0,110),est.mot[,10],xaxt="n",yaxt="n",xlab="",ylab="",col = "green")
lines(smooth.spline(seq(0,110),est.mot[,10],df=10), col = "green")
legend("topleft",col=c("black","red","green"),legend=c("Observed","Negative Binomial","Lee Carter"),pch=1,lty=1, cex=0.7,bty= "n")
title("Observed and forecast mortality rates (At year 2009)")
par(mar=c(5,4,4,5)+.1)
plot(seq(2000,2009),lobs.mot[50,],xlab="year",ylab="Log Mortality rates",col = "black")
lines(smooth.spline(seq(2000,2009),lobs.mot[50,],df=10), col = "black")
par(new=TRUE)
plot(seq(2000,2009),lnest.mot[50,],xaxt="n",yaxt="n",xlab="",ylab="",col = "red")
lines(smooth.spline(seq(2000,2009),lnest.mot[50,],df=10), col = "red")
par(new=TRUE)
plot(seq(2000,2009),lest.mot[50,],xaxt="n",yaxt="n",xlab="",ylab="",col = "green")
lines(smooth.spline(seq(2000,2009),lest.mot[50,],df=10), col = "green")
title("Observed and forecast log mortality rates (At age 50)")
legend("bottomleft",col=c("black","red","green"),legend=c("Observed","Negative Binomial","Lee Carter"),pch=1,lty=1, cex=0.7,bty= "n")
par(mar=c(5,4,4,5)+.1)
plot(seq(0,110),lobs.mot[,10],xlab="age",ylab="Log mortality rates",col = "black")
lines(smooth.spline(seq(0,110),lobs.mot[,10],df=10), col = "black")
par(new=TRUE)
plot(seq(0,110),lnest.mot[,10],xaxt="n",yaxt="n",xlab="",ylab="",col = "red")
lines(smooth.spline(seq(0,110),lnest.mot[,10],df=10), col = "red")
par(new=TRUE)
plot(seq(0,110),lest.mot[,10],xaxt="n",yaxt="n",xlab="",ylab="",col = "green")
lines(smooth.spline(seq(0,110),lest.mot[,10],df=10), col = "green")
legend("topleft",col=c("black","red","green"),legend=c("Observed","Negative Binomial","Lee Carter"),pch=1,lty=1, cex=0.7,bty= "n")
title("Observed and forecast log mortality rates (At year 2009)")

CHAPTER FIVE
# Forecast errors

```

```

robs.mot<-(death.09$Male)/(exp.09$Male)
rest.mot<-(exp(forecast1))/(exp(forecast2))
rest.lca<-list(u)
list.lca<- rapply(rest.lca,c)

neg.errors=robs.mot-rest.mot
lca.errors=robs.mot-list.lca

# Computing MSE and MAE
library(Metrics)
cbind(mse(robs.mot,rest.mot),mse(robs.mot,list.lca))
cbind(mae(robs.mot,rest.mot),mae(robs.mot,list.lca))

lneg.errors=log(robs.mot)-log(rest.mot)
llca.errors=log(robs.mot)-log(list.lca)

Mneg.err<-matrix(neg.errors,111,10,byrow=F)
Mlca.err<-matrix(lca.errors,111,10,byrow=F)
LMneg.err<-matrix(lneg.errors,111,10,byrow=F)
LMlca.err<-matrix(llca.errors,111,10,byrow=F)

# Plot of the distribution of forecast errors
par(mfrow=c(1,2))
par(mar=c(1,1,2,1)+.01)
persp(seq(0,110),seq(2000,2009),Mneg.err,ticktype = "detailed",theta =45,xlab="age",ylab="year",zlab="Forecast Errors")
title("Plot of forecast errors for Negative Binomial")
persp(seq(0,110),seq(2000,2009),Mlca.err,ticktype = "detailed",theta = 45,xlab="age",ylab="year",zlab="Forecast Errors",col = "red")
title("Plot of forecast errors for Lee-Carter")

# Plot of the distribution of log forecast errors
par(mar=c(1,1,2,1)+.2)
persp(seq(0,110),seq(2000,2009),LMneg.err,ticktype = "detailed",theta=45,xlab="age",ylab="year",zlab="log forecast Errors")
title("Plot of log forecast errors for Negative Binomial")
persp(seq(0,110),seq(2000,2009),LMlca.err,ticktype = "detailed",theta =45,xlab="age",ylab="year",zlab="Log forecast Errors",col = "red")
title("Plot of log forecast errors for Lee-Carter")

# forecast errors and log forecast errors (At age 50 and year 2009)
par(mfrow=c(2,2))
par(mar=c(5,4,4,5)+.1)
plot(seq(2000,2009),Mneg.err[50,],xlab="year",ylab="Forecast errors")
lines(smooth.spline(seq(2000,2009),Mneg.err[50,],df=10), col = "black")
par(new=TRUE)
plot(seq(2000,2009),Mlca.err[50,],xaxt="n",yaxt="n",xlab="",ylab="",col = "red")
lines(smooth.spline(seq(2000,2009),Mlca.err[50,],df=10), col = "red")
title("Negative binomial and Lee-Carter (At age 50)")
legend("topleft",col=c("black","red"),legend=c("Negative binomial","Lee-Carter"),pch=1,lty=1, cex=0.7,bty= "n")
par(mar=c(5,4,4,5)+.1)
plot(seq(0,110),Mneg.err[,10],xlab="age",ylab="Forecast errors ")
lines(smooth.spline(seq(0,110),Mneg.err[,10],df=10), col = "black")
par(new=TRUE)
plot(seq(0,110),Mlca.err[,10],xaxt="n",yaxt="n",xlab="",ylab="",col = "red")
lines(smooth.spline(seq(0,110),Mlca.err[,10],df=10), col = "red")
legend("bottomleft",col=c("black","red"),legend=c("Negative binomial","Lee-Carter"),pch=1,lty=1, cex=0.7,bty= "n")
title("Negative binomial and Lee-Carter (At year 2009)")

```

References

- [1] Human Mortality Database (maintained by the University of California, Berkeley (U.S.A), and the Max Planck Institute for Demographic Research (Germany)), www.mortality.org (data downloaded in January 2015).
- [2] Wilmoth, J.R. Andreev, K. Jdanvo, D. and Gleis, D.A. (2007) Methods Protocol for the Human Mortality Database [Online], Available at www.mortality.org (Accessed 28 March 2015).
- [3] Berkeley Mortality Database (BMD): The Lexis Diagram [Online], Available at <http://www.demog.berkeley.edu/~bmd/lexis.html> (accessed Feb, 2015)
- [4] R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.
- [5] Jong, P.de. and Heller, G.Z. (2008) Generalized Linear Model for Insurance Data, New York, Cambridge University Press.
- [6] Hilbe, J.M (2011) Negative Binomial Regression, 2nd Edition, New York, Cambridge University Press.
- [7] Zwilling, M.L. Negative Binomial Regression. The Mathematica Journal, [Online]. Available at <http://mathematica-journal.com/2013/06/> (Accessed April 2015)
- [8] Narula, S.C. (1979) Orthogonal Polynomial Regression, International Statistical Review, vol. 47, no. 1, pp. 31-36 [Online]. Available at <http://www.jstor.org/stable/1403204/> (Accessed March 2015)
- [9] Rodríguez, G. (2015) 4.A Models for Over-Dispersed Count Data [Online]. Available at <http://data.princeton.edu/wws509/r/overdispersion.html> (Accessed April 2015)
- [10] R Data Analysis Examples: Negative Binomial Regression. UCLA: Statistical Consulting Group. from <http://www.ats.ucla.edu/stat/r/dae/nbreg.htm> (accessed May 24, 2015)
- [11] Lee, R.D. Carter, L. Modelling and forecasting the time series of US mortality, Journal of the American Statistical Association 1992; 87:659–671
- [12] Koissi, M-C. Arnold, Shapiro, A.F. (2012), "The Lee-Carter model under the condition of variables age-specific parameters", [Online]. Available at <http://paa2012.princeton.edu/papers/122934> (Accessed 18 May 2015).
- [13] Delwarde, A. Denuit, M. Partrat, C. (2007) Negative binomial version of the Lee–Carter model for mortality forecasting, Wiley InterScience (www.interscience.wiley.com).

wiley.com), 23:385–401, DOI: 10.1002/asmb.679 (Accessed Feb 2015).

[14] Hyndman,R.J. Booth,H. Leonie,T. John,M. (2014). demography: Forecasting Mortality, Fertility, Migration and Population Data. R package version 1.18, Available at <http://CRAN.R-project.org/package=demography>

[15] Charpentier,A. Dutang,C. Goulet,V. (2012) Life insurance and actuarial models with R. Meielisalp 2012 Conference [Online]. Available at <http://freakonometrics.free.fr/slides-Rmetrics-2.pdf> (Accessed April 2015).

[16] Scherp,H. (2003) Applying the Lee-Carter model to countries in Eastern Europe and the former Soviet Union, Diplomarbete, http://aktuarieforeningen.se/filer/DA_Hanna_Scherp.doc (Accessed March 2015).

[17]Jenny,Z.W. (2007) Fitting and Forecasting Mortality for Sweden: Applying the Lee-Carter Model, Master thesis, Mathematical Statistics, Stockholm University.

[18] Robert,H.S. and David,S.S. (2011)Time Series Analysis and Its Applications, New York, Springer Science + Business Media.

[19] Li,S-H. & Chan, W-S. PhD, FSA (2007) The Lee-Carter Model for Forecasting Mortality, Revisited, North American Actuarial Journal, 11:1, 68-89, DOI: 10.1080/10920277.2007.10597438.

[20] Hyndman, R. and Koehler, A. (2005) Another look at measures of forecast accuracy, International Journal of Forecasting; Volume 22, Issue 4, Pages 679–688.