

**An Extended Linear Difference Model for Mortality Projection,
with Applications to Japan**

by

Futoshi Ishii

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Demography

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Kenneth W. Wachter, Co-Chair
Professor John R. Wilmoth, Co-Chair
Professor Bin Yu

Fall 2014

UMI Number: 3686335

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI 3686335

Published by ProQuest LLC (2015). Copyright in the Dissertation held by the Author.

Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against unauthorized copying under Title 17, United States Code



ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

**An Extended Linear Difference Model for Mortality Projection,
with Applications to Japan**

Copyright 2014
by
Futoshi Ishii

Abstract

An Extended Linear Difference Model for Mortality Projection,
with Applications to Japan

by

Futoshi Ishii

Doctor of Philosophy in Demography

University of California, Berkeley

Professor Kenneth W. Wachter, Co-Chair

Professor John R. Wilmoth, Co-Chair

In this dissertation, we propose the Tangent Vector Field (TVF) model for Japanese mortality projection, which is an extended Linear Difference (LD) model, and show its applications.

In the two chapters following the introduction, we describe the mortality trends in Japan and review the mortality projection models for Japan.

Then, in the following two chapters, we describe the data and methods for the mortality models, show the results of fitting, and discuss them with special emphasis on the LD model. We describe the mathematical formulations for *decline*-type and *shift*-type models, and discuss the inverse function of log mortality and differential forms of mortality models. We discuss five models: two *decline*-type models (the Proportional Hazard (PH) and Lee-Carter (LC) models), and three *shift*-type models (the Horizontal Shifting (HS), Horizontal Lee-Carter (HL), and LD models). In particular, we compare the LC and LD models from a statistical viewpoint. The result guides better construction of a mortality projection model, namely, a blended model with LC properties in youth and LD properties in older age.

In the last chapter, we propose the TVF model applying the idea of tangent vector fields on the log mortality surface. We show a fully specified example of the projection procedure of the TVF model with all constants and coefficients applied for Japanese mortality projection. Then, we compare the TVF and LC models' results of mortality projection. From the observation of the relative mortality rates, we see that the LC model expresses mortality improvement only in a vertical direction, whereas the TVF model succeeds in expressing a shifting of mortality improvement in the direction of older ages that are observed in the actual mortality. In addition, we compare the projected m_x curves. The m_x by the LC model exhibits an unnatural pattern because the slope of the curve diminishes once around the age of 60 years and becomes much steeper after 80 years. The curve of the TVF

model is more plausible. As a whole, we observe that the TVF model has many advantages for Japanese mortality projection compared with the LC model.

We show that the TVF model proposed in this dissertation is not only quite useful for Japanese mortality projection but also has various applicability. At this point in time, there may be few countries with such strong shifting features for old age mortality as Japan. However, some countries are likely to experience the same mortality situation as Japan in the future through the extension of life expectancy. Thus, the TVF model will be a useful tool for projections in such situations.

Contents

List of Figures	iii
List of Tables	v
1 Introduction	1
2 Mortality Trends in Japan	13
2.1 Trends in Life Expectancy (e_0)	13
2.2 The Causes of Increases in Life Expectancy	14
2.2.1 Epidemiologic Transition Theory	14
2.2.2 Causes of Deaths	17
3 Mortality Projection Models for Japan	19
3.1 Modeling Age Patterns of Mortality	19
3.1.1 Mathematical Representations	19
3.1.2 Tabular Representations	21
3.1.3 Relational Models	22
3.2 Review of the Official Mortality Projection for Japan	22
3.3 The Lee–Carter Model and its Application to Japan	25
4 Data and Methods	28
4.1 Data	28
4.2 Methods	29
4.2.1 Formulation of Decline–type and Shift–type Models	29
4.3 Descriptions of Decline-Type and Shift-Type Models	33
4.3.1 Decline-Type Mortality Models	33
4.3.2 Shift-Type Mortality Models	34
4.3.3 The Linear Difference (LD) Model	34
4.4 Methods for Parameter Estimations	38
5 Linear Difference Model	41
5.1 Fitting the Mortality Models	41
5.1.1 Fitting Decline-Type Mortality Models	41

5.1.2	Fitting Shift-Type Mortality Models	48
5.2	Comparison of the Models from a Statistical Viewpoint	61
5.3	Application to Analysis of the Trends of Modal Age at Death	70
5.3.1	Decomposition of the Change in Modal Age using the LD Model	70
5.3.2	Results of Decomposition	73
6	Tangent Vector Field Approach to Mortality Projection	76
6.1	Tangent Vector Field Approach to Mortality Projection	76
6.1.1	Building an Entire Age Model Using Tangent Vector Fields	76
6.1.2	Application to Japanese Mortality Projection	80
6.1.3	Discussion of Mortality Projection	99
7	Summary and Conclusion	101
A	Supplement Tables	108

List of Figures

1.1	Stylized Example of Mortality Improvement (Gompertz Case)	2
1.2	Stylized Example of Mortality Improvement (General Case)	3
1.3	Log Mortality Rates(Japan, Female, 1970-2010)	4
1.4	Change in the Mortality Curves	5
1.5	Stylized Example of Log Mortality Surface	6
2.1	Trends of Life Expectancy (Sweden)	14
2.2	Trends of Life Expectancy (Japan)	15
2.3	Trends of Life Expectancy (Japan versus Other Countries)	16
2.4	Trends of l_x curves (Females, Japan)	17
2.5	Age-Standardized Death Rates by Major Causes (Males, Japan)	18
3.1	Log Mortality Rates (Coale and Demeny Model Life Tables)	21
3.2	Log Mortality Rates (Coale and Demeny Model Life Tables))	21
3.3	Comparison of Life Expectancy in the Official Projections	24
4.1	Log Mortality Rates (Female Japan)	29
4.2	Inverse Log Mortality Rates (Female Japan))	29
4.3	Log Mortality Surface and Two Differential Functions	32
4.4	Stylized example of the LD model	36
4.5	Stylized example of the Effect of Change in S_t and g_t	37
5.1	Mortality Rates (Actual and Model, PH)	42
5.2	Difference of Mortality Rates (Actual - Model, PH)	42
5.3	Mortality Rates (Actual and Model, LC)	44
5.4	Difference of Mortality Rates (Actual - Model, LC)	44
5.5	Comparison of Mortality Improvement Rates (1975–1990)	46
5.6	Comparison of Mortality Improvement Rates (1991–2006)	47
5.7	Inverse Mortality Rates (Actual and Model, HS)	49
5.8	Difference of Inverse Mortality Rates (Actual and Model, HS)	49
5.9	Inverse Mortality Rates (Actual and Model, HL)	51
5.10	Difference of Inverse Mortality Rates (Actual and Model, HL)	51
5.11	Estimated Parameters (LD)	52

5.12	Inverse Mortality Rates (Actual and Model, LD)	54
5.13	Difference of Inverse Mortality Rates (Actual and Model, LD)	54
5.14	Comparison of the Force of Age Increase by Age (1975–1990)	56
5.15	Comparison of the Force of Age Increase by Age (1991–2006)	57
5.16	Example of the LD model ($g'_0 < 0$)	60
5.17	Example of the LD model ($g'_0 > 0$)	60
5.18	Difference of Log Actual Rate and CI against Model (LC)(1970–1984, critical value = 0.01%)	62
5.19	Difference of Log Actual Rate and CI against Model (LC)(1985–1999, critical value = 0.01%)	63
5.20	Difference of Log Actual Rate and CI against Model (LC)(2000–2010, critical value = 0.01%)	64
5.21	Difference of Log Actual Rate and CI against Model (LD)(1970–1984, critical value = 0.01%)	65
5.22	Difference of Log Actual Rate and CI against Model (LD)(1985–1999, critical value = 0.01%)	66
5.23	Difference of Log Actual Rate and CI against Model (LD)(2000–2010, critical value = 0.01%)	67
5.24	Proportion of Log Actual Values That Are Outside of CI (critical value = 0.01%)	69
5.25	Trends of the Modal Age at Death (Actual and LD, Females, Japan) .	74
5.26	Decomposition of the change of the Modal Age at Death (Females Japan)	75
6.1	Change in the Mortality Curves	77
6.2	Tangent Vectors on the Log Mortality Surface	78
6.3	Example of Construction of a Tangent Vector Field	79
6.4	Projection of k_t	82
6.5	Projection of g_t and S_t	84
6.6	Projection of $w(x, t)$	89
6.7	Projected Life Expectancy	97
6.8	Relative Mortality Rates (Actual)	97
6.9	Relative Mortality Rates (LC)	97
6.10	Relative Mortality Rates (TVF)	97
6.11	Comparison of m_x curves (LC and TVF)	98

List of Tables

3.1	Examples of Mathematical Representations	20
6.1	Estimated a_x and b_x	81
6.2	Estimated k_t	83
6.3	Projected k_t	83
6.4	Estimated g_t and S_t	85
6.5	Projected g_t and S_t	85
6.6	Projected f_t	87
6.7	Projected $x_1(t)$	90
A.1	Estimated a_x and b_x	109
A.2	Projected Log Mortality Rates by the LC model	110
A.3	Projected Log Mortality Rates by the TVF model	111

Acknowledgements

This dissertation would not have been made possible without the support of many people.

First, I thank my committee members, Ken Wachter, John Wilmoth and Bin Yu. To Ken Wachter, I thank you for your mentoring, and providing me invaluable comments in particular from the viewpoint of demographic methods and mathematical demography. The idea of using the inverse function of log mortality was originally motivated by your comment at the committee meeting. I could not complete this work without your guidance and support. To John Wilmoth, I thank you for your insightful feedback to my mortality models. The tangent vector field approach was inspired by the figure that you drew on the blackboard in the demography seminar room. I am also grateful for giving me opportunity to work for the HMD project. Working for the project was a precious experience for me and was also fruitful for this study. To Bin Yu, I thank you for your valuable comments. I was deeply impressed by your lecture at the Department of Statistics. The methods and applications for statistical modeling that I learned in your class was extremely useful for developing models in this study.

I thank the Demography community at Berkeley. I am thankful to all of the faculty members, staffs and students at the Department of Demography. Especially, I thank Liz Ozselcuk and Monique Verrier for their great support and help.

I thank Shiro Horiuchi for giving me continuous insightful comments on my study. My model was greatly improved by his feedback. I also thank current and former colleagues at the National Institute of Population and Social Security Research. I thank Shigesato Takahashi for inviting me to the academic world of demography and giving me an opportunity to study mortality. I am grateful to Ryuichi Kaneko for giving me thoughtful suggestions about mortality modeling, in particular shift-type models. I also thank Toru Suzuki for helping me as an alumnus when I applied to Demography at Berkeley. I am grateful to all the members of the population projection team at the institute, especially Miho Iwasawa, who has always inspired me through the discussion about population projections.

I am also grateful to my family. I thank my parents, Taro and Aiko Ishii, and my brother Kiyoshi Ishii. Finally, I am especially thankful to my wife, Taiko Ishii. She is always with me, and supported me when we were in Berkeley as well as in Japan. I could not have done this work without her.

Chapter 1

Introduction

This dissertation proposes and examines a novel model for Japanese mortality projection, which is an extended Linear Difference (LD) model, namely, the Tangent Vector Field (TVF) model. This model was originally developed for the 2012 official population projection for Japan conducted by the National Institute of Population and Social Security Research (NIPSSR 2012).

In Japan, the increase in life expectancy in the 20th century was remarkable as with other developed countries. Moreover, the pace of the extension of life expectancy was noticeable. Japanese life expectancy was at the lowest level among developed countries in 1950. However, Japan caught up rapidly, overtook other countries, and has since continued to increase.

However, these unique characteristics of Japanese mortality pose huge challenges for modeling and projecting mortality. Existing mortality models often cannot capture the peculiarities of Japanese mortality and neither can the Lee-Carter (LC) model (Lee and Carter 1992), which is now regarded internationally as a standard model. We started our research originally to overcome this problem.

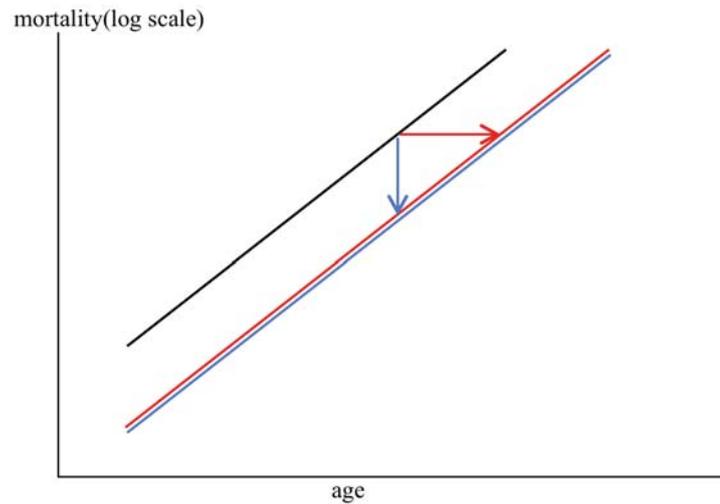
In this introduction, we overview our model, including a description of some important concepts and the substantial rationale for the model. Then, we provide an example of the projection procedure using the TVF model. Lastly, we describe advantages of the TVF model.

First, we introduce some necessary concepts. There are *three* important ideas infrequently used in general that we prefer to use in this dissertation:

- (a) *decline*-type and *shift*-type models, and the inverse function of log mortality;
- (b) differential forms of mortality models; and
- (c) tangent vector fields on the log mortality surface.

These ideas are discussed in detail later. Here, we briefly describe them and show the direction in which this dissertation proceeds.

Fig. 1.1: Stylized Example of Mortality Improvement (Gompertz Case)

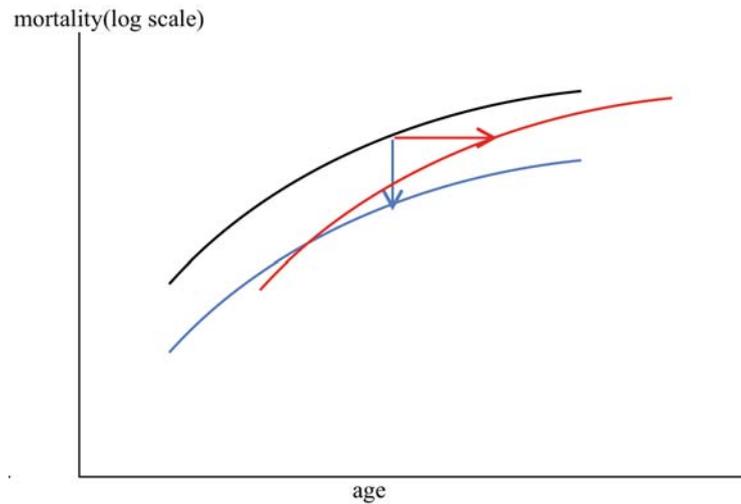


Regarding the first idea, let us observe Fig. 1.1 that shows a stylized example of mortality improvement. The black straight line is the log mortality at a base point of time. This is a Gompertz model because the log mortality is a linear function of age. Assume that the mortality *improved* during a period of time and the log mortality line moved to the colored (blue and red) lines. Usually, we recognize the improvement, as shown by the blue arrow, that is, the value of log mortality *declines* over time for a fixed age. However, this is also understood as improvement as shown by the red arrow. This means that the black line *shifted* toward older ages. In this case, the mortality improvement is recognized as a delay of the age that attains to a fixed value of the log mortality. In the Gompertz model, they are unidentifiable, making the blue and red line coincide.

In general, however, *decline* and *shift* do not yield the same results. Fig. 1.2 shows the generalized version of Fig. 1.1. The log mortality at the base point of time shown by the black line is no longer straight. For simplicity, we consider only the case in which the improvement occurs uniformly, that is, the black line moves parallel. We see that the blue and red lines are different in this case, and therefore, *decline* and *shift* exhibit different types of mortality improvement. We show later that *shift*-type mortality improvement is preferable for understanding Japanese old age mortality.

The inverse function of log mortality plays an important role in treating *shift*-type models efficiently. Recall that we recognized an increase of the age that corresponds to a fixed value of the log mortality in the *shift*-type mortality improvement. Here, the function that expresses the age that attains to a fixed value of the log mortality is the *inverse* function of the log mortality rates. Therefore, shifting features are modeled well by the inverse functions. Using this concept, we con-

Fig. 1.2: Stylized Example of Mortality Improvement (General Case)



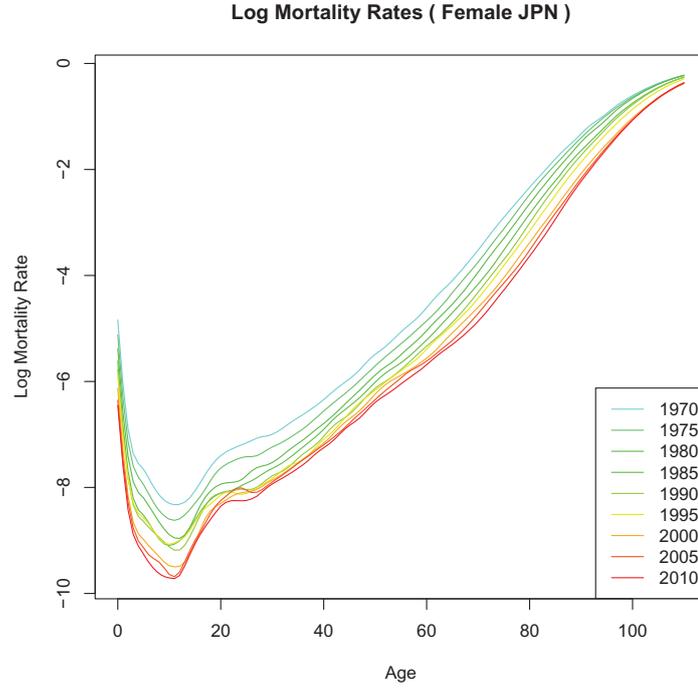
struct the LD model, which is a *shift*-type model. In addition, this enables us to treat more general types of shift. As the LC model handles a different amount of mortality decline with each age, we can treat a different amount of shifting with each log mortality level using the inverse functions.

The second idea, *differential forms* of mortality models, relates to targets of modeling. In general, the target of the mortality model is the age pattern of the log mortality. We start this research by developing a mortality projection model, and our models always include time as a variable. Therefore, we can model the mortality with a baseline mortality pattern and a change of mortality over time. Thus, we notice the differential of mortality function by time, which we call *differential forms*.

Actually, when relational models, which express an arbitrary mortality pattern by a standard age pattern and the differences from the standard pattern with some parameters, are used for time series modeling of mortality rates, the differences from the standard pattern mean changes of the mortality over time. In this case, the changes are expressed by the differential of mortality by time. Therefore, estimation of relational models for time series is equivalent to modeling the differential of the mortality. In this dissertation, differential forms are used more actively, and play a central role in defining the LD model.

The third idea concerns practical implementation of the mortality projection model. One of the reasons why *shift*-type models are not used frequently for mortality projection is the difficulty in building an entire age model. Fig. 1.3 shows the log mortality rates for Japanese females since 1970. We see that juvenile mortality is never recognized as shifting. Therefore, *shift*-type models are applicable to adult mortality only. However, we need an entire age model for mortality projection and

Fig. 1.3: Log Mortality Rates(Japan, Female, 1970-2010)



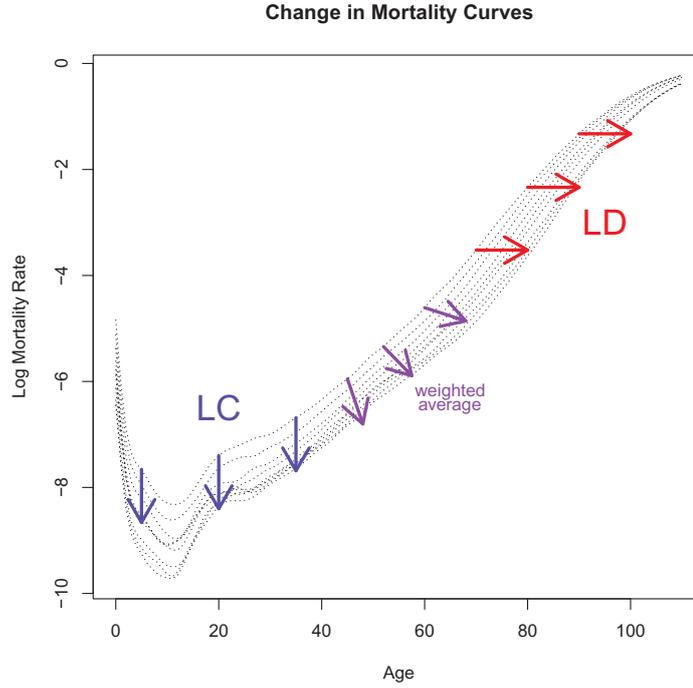
juvenile mortality is modeled only by *decline*-type models. This fact is considered as a chief obstacle to applying *shift*-type models to mortality projections in practice.

Fig. 1.4 shows a stylized example of the change in mortality curves. If we use the LD model, which is a *shift*-type adult mortality model, the direction of mortality improvements is expressed by the age increases shown by the red arrows. On the other hand, the mortality improvements in juvenile mortality are modeled well by the *decline*-type models, such as the LC model, whose mortality improvements are shown by the blue arrows. The arrows express the directions in which the points on the log mortality curves are heading. Mathematically, these arrows are formulated using *tangent vector fields* on the log mortality surface.

Fig. 1.5 shows an example of the log mortality surface S that is expressed in the blue grid. The log mortality surface is a set of points in $(x, t, y) \in \mathbb{R}^3$, where x is age, t is time, and $y = \log \mu_{x,t}$ is log mortality.

Recall from the second idea that modeling the differential of the mortality function by time is equivalent to mortality modeling in time-series analysis. For *decline*-type models, the mortality improvement rates $\rho_{x,t}$, which are the minus of the differential of the log mortality, become the target of modeling. Equivalently, the differential of the inverse of the log mortality is the target in the *shift*-type models.

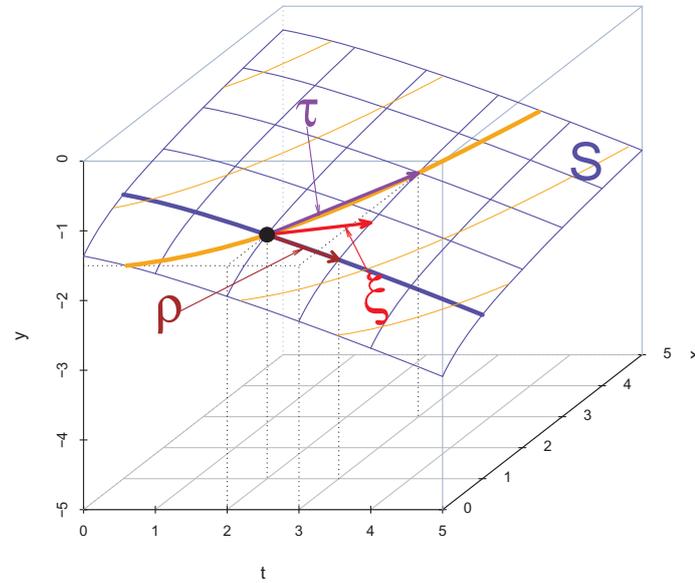
Fig. 1.4: Change in the Mortality Curves



Let us consider what they express on the log mortality surface. We begin with the black point in Fig. 1.5 and consider how mortality improvement is expressed for the two types of models. In *decline*-type models, the mortality improvement is recognized as a decline of log mortality with a fixed age, which means we travel along the broad blue line on S in Fig.1.5. Here, the mortality improvement rates $\rho_{x,t}$ is the minus of the slope of the broad blue line at the black point. Therefore, the vector $\rho = (0, 1, -\rho_{x,t})$ is a tangent vector on S and the set of the vectors for all points form a tangent vector field on S . This is interpreted as follows: modeling the *decline*-type models are equivalent to considering the flows at each point on the log mortality surface in the direction of age fixed with tangent vector ρ . The same discussion holds for the *shift*-type models, except that we consider the flows in the direction of the log mortality fixed with the broad orange line, and the corresponding tangent vector is τ in Fig. 1.5.

Considering the flows on the log mortality surface enables us to construct an entire age model in a natural way. We can consider an intermediate tangent vector ξ by a weighted average of ρ and τ . Then, we can construct flows that coincide *decline*-type flows in youth and *shift*-type flows in old ages, which produces an entire age model. In this way, the tangent vector fields dissociate from the difficulties of applying *shift*-type models to mortality projection.

Fig. 1.5: Stylized Example of Log Mortality Surface



Furthermore, considering the flows has substantive meaning. Usually, mortality models are expressed as a function of age that defines a parameter to pick out points on each curve. However, the flow on the log mortality surface means that each point $P(0)$ is associated with a specific point $P(t)$ on the curve for time t , which allows more general types of parameterization. In the LD model, the level of the mortality is used for the parameterization. This feature has an advantage when a change in log mortality depends more on mortality level than age. Therefore, this is useful for modeling old age mortality in developed countries, especially Japan, when “the aging of mortality decline” plays an important role in mortality improvement.

However, in this setting, we no longer have a formula for a mortality model as a function of age. To show how the projection procedure is carried out, here, we provide an example of the procedure of the projection by vector approach. The complete worked example appears in Chapter 6.

Before showing the example, we briefly describe some notations and definitions related to our models, even though they are discussed in detail in the following chapters. In this study, the log hazard function of mortality is expressed by $y = \lambda_{x,t} = \log \mu_{x,t}$, and $v_{y,t}$ exhibits the inverse function of $\lambda_{x,t}$ if defined. We use the following two differential functions by time t : (1) $\rho_{x,t}$: the mortality improve-

ment rate and (2) $\tau_{y,t}$: the force of age increase defined as:

$$\rho_{x,t} \stackrel{\text{def}}{=} -\frac{\partial \lambda_{x,t}}{\partial t} = -\frac{\partial \log \mu_{x,t}}{\partial t}$$

$$\tau_{y,t} \stackrel{\text{def}}{=} \frac{\partial v_{y,t}}{\partial t}$$

Using these notations, the LC model is expressed as follows¹:

$$\lambda_{x,t} = a_x + k_t b_x$$

Here, a_x is a baseline log mortality, k_t is interpreted as an indicator for the level of mortality at time t , and b_x is a set of sensitivity constants that stands for the amount of mortality improvement at age x for the unit change of k_t .

The LD model is a shift-type model and is described using the force of age increase $\tau_{y,t}$ that is a differential of the inverse function of log mortality. We define the LD model satisfying the property that $\tau_{y,t}$ is a linear function of t for each t :

$$\tau_{y,t} = f'_t + g'_t x$$

The parameters f'_t and g'_t exhibit the intercept and the slope of the line that corresponds to a linear relationship between $\tau_{y,t}$ and x . We can obtain f_t and g_t by integrating f'_t and g'_t with t .

f_t depends on the value of the slope and is not easy to interpret. Therefore, we introduce another variable S_t as a location of the mortality curve instead of f'_t or f_t , defined as the age that the mortality rate equals to 0.5 at time t . In theory, we can always convert from S_t to f_t using the value of g_t .

The TVF model is a blended mortality model that has the LC property in youth and the LD property in older age. Therefore, to project the mortality by the TVF model, we need all the parameters for the LC and the LD model, as well as a weight function that combines the tangent vectors that correspond to the two models. They are listed as follows:

- a baseline log mortality a_x (estimated in the LC model)
- a set of Lee–Carter b_x sensitivity constants
- estimated and projected values of k_t for the LC model
- estimated and projected values of g_t for the LD model
- estimated and projected values of S_t for the LD model (and the converted values of f_t)

¹We consider $m_{x,t}$ in the discrete form is an approximation of $\mu_{x+0.5,t+0.5}$ in the continuous form. Moreover, we drop the error term from models. See Chapter 4 for details.

- a weight function $w(x, t)$

Suppose we have already estimated the above parameters for now. We begin with the vector field approach for the LC model. For the LC model, we can directly project $\lambda_{x,t}$ (log mortality rates) from a_x , b_x , and k_t^2 . For example, $\lambda_{x,2010}$ is projected by the formula:

$$\lambda_{x,2010} = a_x + k_{2010}b_x$$

In the vector field approach, we carry out the projection from the mortality rates in year t to those for $t + 1$ recursively, starting from the baseline mortality. This is just like a population projection procedure by the cohort component method. With this approach, we project $\lambda_{x,2010}$ from $\lambda_{x,2009}$, adding $-\rho_{x,2009} = (k_{2010} - k_{2009})b_x$.

This procedure is expressed in the following diagram. The left box shows the coordinates $(x, y) = (x, \lambda_{x,2009})$ by the LC model for $t = 2009$. The center box is the vector ρ that indicates the change in each point on the mortality curve for the LC model. The right box, which shows the coordinates $(x, y) = (x, \lambda_{x,2010})$ is obtained by adding the center box to the left one.

t = 2009		Change		t = 2010	
x	y	x	y	x	y
0	-6.41880	0	-0.02575	0	-6.44456
5	-9.20639	0	-0.02444	5	-9.23082
10	-9.67716	0	-0.02119	10	-9.69835
15	-9.09892	0	-0.01506	15	-9.11398
20	-8.28018	0	-0.01034	20	-8.29052
25	-8.12543	0	-0.01109	25	-8.13652
30	-7.94453	0	-0.01197	30	-7.95650
35	-7.64278	0	-0.01230	35	-7.65508
40	-7.24150	0	-0.01196	40	-7.25346
45	-6.82626	0	-0.01243	45	-6.83869
50	-6.39262	0	-0.01275	50	-6.40537
55	-6.03363	0	-0.01365	55	-6.04728
60	-5.66740	0	-0.01532	60	-5.68272
65	-5.30739	0	-0.01887	65	-5.32626
70	-4.82596	0	-0.02087	70	-4.84683
75	-4.24970	0	-0.02203	75	-4.27173
80	-3.61874	0	-0.02236	80	-3.64110
85	-2.91721	0	-0.01995	85	-2.93715
90	-2.23439	0	-0.01643	90	-2.25082
95	-1.62468	0	-0.01277	95	-1.63746
100	-1.09103	0	-0.00910	100	-1.10012
105	-0.67125	0	-0.00565	105	-0.67690
110	-0.38047	0	-0.00307	110	-0.38354

²Note that when we work on the discrete form, the age x actually represents the age interval $[x, x + 1)$. In this example, we treat non-integer values for x , which should be understood not as the exact ages but as the age intervals.

Next, we consider the projection procedure for the LD model with a similar diagram as that for the LC.

The first box from the left is $(x, y) = (x, \lambda_{x,2009})$ by the LD model. The second box expresses the vector τ of change for the LD model. Note that the values of change $\tau_{y,t}$ are plugged in the "x" (left) column here, which corresponds to the direction of the flow for the LD model. We obtain the third box by adding the first and second ones. It shows the relationship between x and y . However, this is not a normal representation because the values of x are not integers. Therefore, we "standardize" it by linear interpolations and obtain a normal representation, as shown in the fourth box.

t = 2009		Change		t = 2010		t = 2010	
<i>x</i>	<i>y</i>	<i>x</i>	<i>y</i>	<i>x</i>	<i>y</i>	<i>x</i>	<i>y</i>
0	-6.41880	0.30780	0	0.30780	-6.41880	0	
5	-9.20639	0.29822	0	5.29822	-9.20639	5	-9.17441
10	-9.67716	0.28865	0	10.28865	-9.67716	10	-9.65753
15	-9.09892	0.27908	0	15.27908	-9.09892	15	-9.16408
20	-8.28018	0.26950	0	20.26950	-8.28018	20	-8.31101
25	-8.12543	0.25993	0	25.25993	-8.12543	25	-8.13085
30	-7.94453	0.25035	0	30.25035	-7.94453	30	-7.95642
35	-7.64278	0.24078	0	35.24078	-7.64278	35	-7.66127
40	-7.24150	0.23121	0	40.23121	-7.24150	40	-7.25807
45	-6.82327	0.22163	0	45.22163	-6.82327	45	-6.84214
50	-6.38706	0.21206	0	50.21206	-6.38706	50	-6.40529
55	-6.02569	0.20248	0	55.20248	-6.02569	55	-6.04032
60	-5.65690	0.19291	0	60.19291	-5.65690	60	-5.67083
65	-5.29181	0.18334	0	65.18334	-5.29181	65	-5.30685
70	-4.80574	0.17376	0	70.17376	-4.80574	70	-4.82438
75	-4.22945	0.16419	0	75.16419	-4.22945	75	-4.24906
80	-3.59925	0.15461	0	80.15461	-3.59925	80	-3.61961
85	-2.90073	0.14504	0	85.14504	-2.90073	85	-2.92162
90	-2.22207	0.13547	0	90.13547	-2.22207	90	-2.23970
95	-1.61629	0.12589	0	95.12589	-1.61629	95	-1.63079
100	-1.08622	0.11632	0	100.11632	-1.08622	100	-1.09765
105	-0.66927	0.10675	0	105.10675	-0.66927	105	-0.67709
110	-0.38016	0.09717	0	110.09717	-0.38016	110	-0.38483

In this case, we cannot obtain the value for age 0 by linear interpolation. However, this causes no problem because we do not use the projected mortality rates by the LD model for juvenile area.

Finally, we describe the TVF procedure that combines the two models. In the TVF model, we compute the vector of change ζ by the weighted average of the vectors ρ and τ . Using weight function $w(x, t)$, the vector of change ζ is constructed as shown in the following diagram.

ρ		$1 - w(x, t)$	τ		$w(x, t)$	ξ	
x	y		x	y		x	y
0	-0.02575	1.00000	0.30780	0	0.00000	0.00000	-0.02575
0	-0.02444	1.00000	0.29822	0	0.00000	0.00000	-0.02444
0	-0.02119	1.00000	0.28865	0	0.00000	0.00000	-0.02119
0	-0.01506	1.00000	0.27908	0	0.00000	0.00000	-0.01506
0	-0.01034	1.00000	0.26950	0	0.00000	0.00000	-0.01034
0	-0.01109	1.00000	0.25993	0	0.00000	0.00000	-0.01109
0	-0.01197	1.00000	0.25035	0	0.00000	0.00000	-0.01197
0	-0.01230	1.00000	0.24078	0	0.00000	0.00000	-0.01230
0	-0.01196	1.00000	0.23121	0	0.00000	0.00000	-0.01196
0	-0.01243	0.83526	0.22163	0	0.16474	0.03651	-0.01038
0	-0.01275	0.67052	0.21206	0	0.32948	0.06987	-0.00855
0	-0.01365	0.50578	0.20248	0	0.49422	0.10007	-0.00690
0	-0.01532	0.34104	0.19291	0	0.65896	0.12712	-0.00522
0	-0.01887	0.17630	0.18334	0	0.82370	0.15102	-0.00333
0	-0.02087	0.01155	0.17376	0	0.98845	0.17175	-0.00024
0	-0.02203	0.00000	0.16419	0	1.00000	0.16419	0.00000
0	-0.02236	0.00000	0.15461	0	1.00000	0.15461	0.00000
0	-0.01995	0.00000	0.14504	0	1.00000	0.14504	0.00000
0	-0.01643	0.00000	0.13547	0	1.00000	0.13547	0.00000
0	-0.01277	0.00000	0.12589	0	1.00000	0.12589	0.00000
0	-0.00910	0.00000	0.11632	0	1.00000	0.11632	0.00000
0	-0.00565	0.00000	0.10675	0	1.00000	0.10675	0.00000
0	-0.00307	0.00000	0.09717	0	1.00000	0.09717	0.00000

Then, we can perform the projection similarly to the procedure for the LD model, except we use ξ for the vector of change, as shown in the following diagram.

t = 2009		Change		t = 2010		t = 2010	
x	y	x	y	x	y	x	y
0	-6.41880	0.00000	-0.02575	0.00000	-6.44456	0	-6.44456
5	-9.20639	0.00000	-0.02444	5.00000	-9.23082	5	-9.23082
10	-9.67716	0.00000	-0.02119	10.00000	-9.69835	10	-9.69835
15	-9.09892	0.00000	-0.01506	15.00000	-9.11398	15	-9.11398
20	-8.28018	0.00000	-0.01034	20.00000	-8.29052	20	-8.29052
25	-8.12543	0.00000	-0.01109	25.00000	-8.13652	25	-8.13652
30	-7.94453	0.00000	-0.01197	30.00000	-7.95650	30	-7.95650
35	-7.64278	0.00000	-0.01230	35.00000	-7.65508	35	-7.65508
40	-7.24150	0.00000	-0.01196	40.00000	-7.25346	40	-7.25346
45	-6.82327	0.03651	-0.01038	45.03651	-6.83366	45	-6.83675
50	-6.38706	0.06987	-0.00855	50.06987	-6.39560	50	-6.40159
55	-6.02569	0.10007	-0.00690	55.10007	-6.03260	55	-6.03981
60	-5.65690	0.12712	-0.00522	60.12712	-5.66212	60	-5.67128
65	-5.29181	0.15102	-0.00333	65.15102	-5.29514	65	-5.30752
70	-4.80574	0.17175	-0.00024	70.17175	-4.80598	70	-4.82441
75	-4.22945	0.16419	0.00000	75.16419	-4.22945	75	-4.24906
80	-3.59925	0.15461	0.00000	80.15461	-3.59925	80	-3.61961
85	-2.90073	0.14504	0.00000	85.14504	-2.90073	85	-2.92162
90	-2.22207	0.13547	0.00000	90.13547	-2.22207	90	-2.23970
95	-1.61629	0.12589	0.00000	95.12589	-1.61629	95	-1.63079
100	-1.08622	0.11632	0.00000	100.11632	-1.08622	100	-1.09765
105	-0.66927	0.10675	0.00000	105.10675	-0.66927	105	-0.67709
110	-0.38016	0.09717	0.00000	110.09717	-0.38016	110	-0.38483

In Chapter 6, we show a fully specified example of the projection procedure of the TVF model with all constants and coefficients.

Lastly, we describe advantages of the TVF model. We will compare the LC and LD models from a statistical viewpoint in Section 5.2, and see that LD's performance is better than LC's performance over 75 years of age. This is brought by the shift feature of the LD model that can express the "aging of mortality decline" observed in the recent mortality improvement in Japan. Moreover, the *shift*-type models are applicable only for adult mortality and the juvenile mortality should be modeled by the *decline*-type models. Therefore, the TVF model that has the LC property in youth and the LD property in older age works quite well for mortality projection for Japan.

Moreover, it is often mentioned that the projected mortality curve by the LC model leads to an unnatural pattern especially in a long range projection. This is caused by the fact that the parameter b_x is fixed in the LC model, even though the age distribution of the mortality improvement rates is not fixed actually. On the other hand, we can obtain a plausible age pattern of mortality by the TVF model, which is considered another advantage of our model.

The LD model that composes the older age part of the TVF model is useful not only for projection but also analyses of mortality. We propose a decomposition

method for the modal age at death in Section 5.3. Since the LD model is originally developed for mortality projections, the number of parameters were reduced in terms of parsimony. This might be a restriction in terms of a flexible expression for various types of mortality situation. However, this feature brings another possibility to derive simple analytical formula. The decomposition that we propose is easy to apply when the mortality curves are modeled by the LD model, and has a clear interpretation composed by shifting, compression and other parts. Thus, we can say that the model that we propose is not only useful for Japanese mortality projection but also has various applicability, as we discuss in the following chapters.

This dissertation is organized as follows.

Chapter 2 describes Japan's mortality trends. We discuss the unique characteristics of Japanese mortality with the causes of extended life expectancy.

Chapter 3 reviews the mortality projection models for Japan. First, we review the general modeling for mortality that includes a discussion of the relational model described above. Then, we review the official mortality projection for Japan that is the starting point of this research. Lastly, we review the LC model and its applications to Japan.

Chapter 4 describes the data and methods for the study of the new mortality models. Following the data description, we describe the method, starting with the mathematical formulations for *decline*-type and *shift*-type models, and discuss the inverse function of log mortality and differential forms of mortality models. Then, we discuss five models: two *decline*-type models, which are the Proportional Hazard (PH) and LC models, and three *shift*-type models, which are the Horizontal Shifting (HS), Horizontal Lee-Carter (HL), and LD models.

Chapter 5 describes the results of fitting of the five models, with special focus on the comparison of the LC and LD models. First, we fit the five mortality models to Japanese female old age mortality and examine which model is appropriate for the projection of Japanese mortality. Then, we compare the LC and LD models from a statistical viewpoint to examine whether it is more plausible to understand Japan's recent old age mortality as *decline* or *shift*. In addition, we show an application of the LD model that is a decomposition method for the modal age at death, and provide decomposition analyses for Japan with the method.

Chapter 6 describes a mortality projection by the TVF model. We discuss in detail the method for building an entire age model by applying the idea of tangent vector fields.

Chapter 2

Mortality Trends in Japan

2.1 Trends in Life Expectancy (e_0)

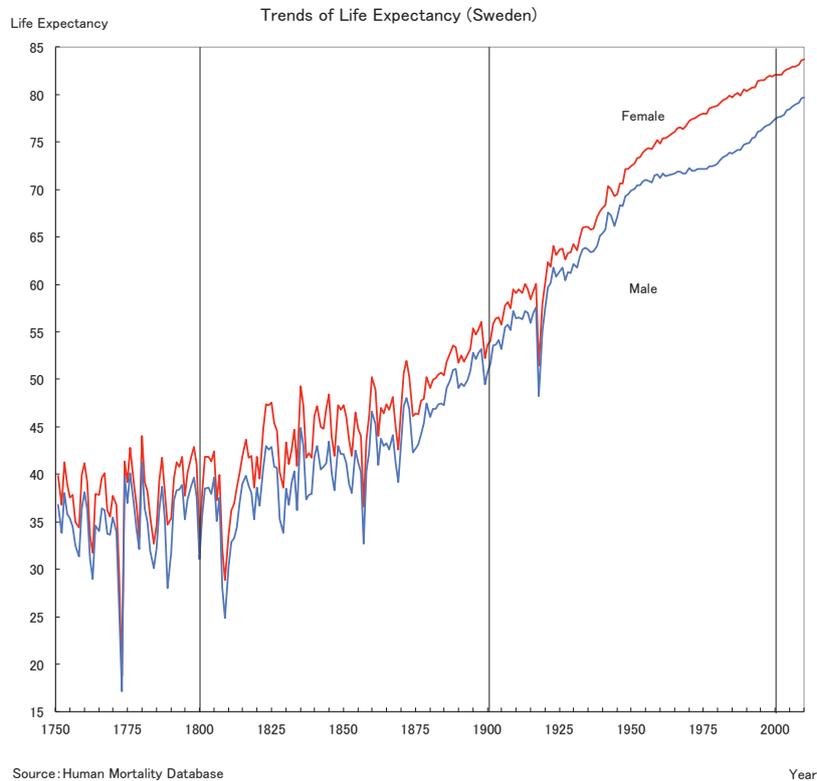
Longevity extension is one of the greatest achievements that human beings have made through history. Life expectancy (e_0) in the earliest human history is assumed to have been around 20–35 years, whereas that of developed countries in 1900 was 40–50 years. Nowadays, life expectancy has reached about 80 years in developed countries, having risen remarkably over the last few centuries, with a substantial part of the historical extension achieved recently.

Fig. 2.1 shows the historical trends in life expectancy in Sweden since the mid-18th century. This figure indicates that life expectancy was between about 30 and 40 years in 1750, whereas it is around 80 years now. In particular, we observe an enormous increase in life expectancy in the 20th century.

In Japan, the increase in life expectancy in the 20th century was also remarkable, in lines with other developed countries. Fig. 2.2 shows the trends in life expectancy in Japan since around 1920 by the official life tables. In 1921–1925, life expectancy was 42.06 and 43.20 years for males and females, respectively. However, it was 79.44 and 85.90 years for males and females, respectively, in 2011, having doubled in around 90 years.

Not only the amount but also the pace of the increase is noticeable for Japanese life expectancy. Fig. 2.3 compares the life expectancies of several developed countries in the second half of the 20th century. The black lines show the trajectories for Japan, whereas the colored ones show those for other countries. We observe that Japanese life expectancy was at its lowest level among developed countries in 1950. However, Japan rapidly caught up, overtook the other countries, and has continued to increase life expectancy. Therefore, the slopes of Japanese lines are much steeper than those for the other countries, exhibiting unique characteristics for Japanese mortality.

Fig. 2.1: Trends of Life Expectancy (Sweden)



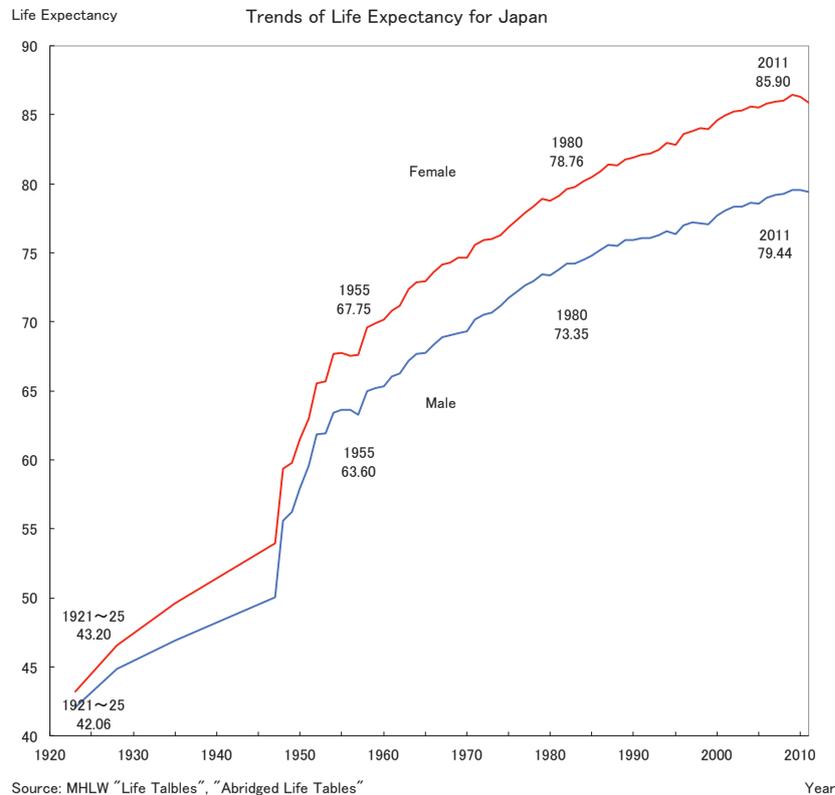
2.2 The Causes of Increases in Life Expectancy

2.2.1 Epidemiologic Transition Theory

The remarkable expansion of life expectancy in the first half of the 20th century was caused mainly by a fall in the death rates because of a decline in infectious disease mortality. This historical change in the structure of diseases and mortality is explained by the epidemiologic transition theory. Epidemiologic transition is a characteristic shift in the disease pattern of a population as mortality falls during the demographic transition. Through the transition, acute, infectious diseases are reduced, while chronic, degenerative diseases increase in prominence.

The epidemiologic transition theory was first introduced by Omran (1971). He argues that during the epidemiologic transition, a long-term shift occurs in mortality and disease patterns whereby pandemics of infection are gradually displaced

Fig. 2.2: Trends of Life Expectancy (Japan)



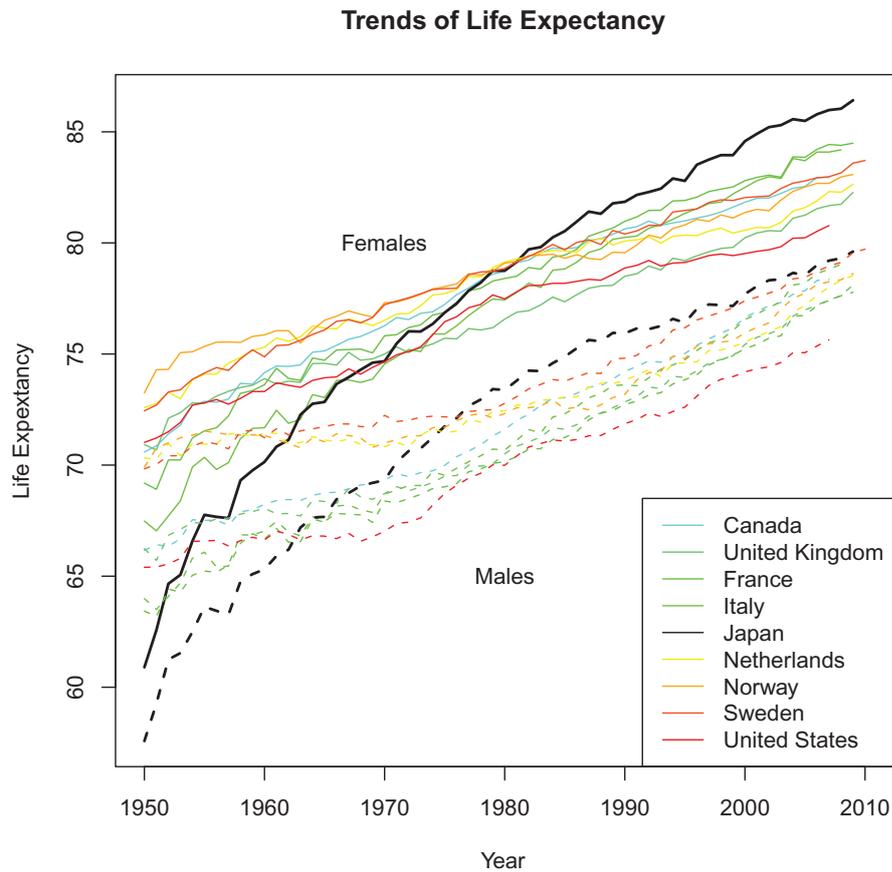
by degenerative and man-made diseases as the chief form of morbidity and primary cause of death. He distinguishes three stages:

- Age of pestilence and famine
- Age of receding pandemics
- Age of degenerative or "man-made" diseases

This transition causes a shift in the distribution of deaths from younger to older ages. However, there was no significant observed mortality improvement until 1970. This led to discussion about survival curves becoming more rectangularized as life expectancy increases assuming that the length of life is limited, and the improvement of life expectancy would slow down and eventually stop (Fries 1980).

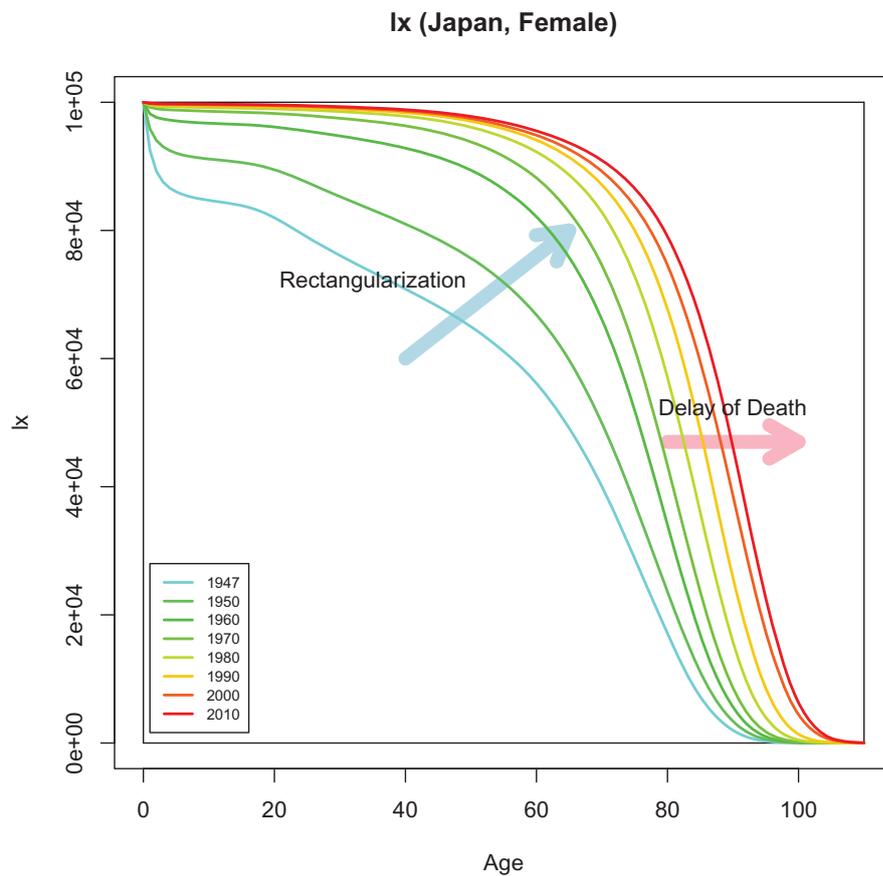
However, the increase of life expectancy continued after 1970 due to the improvement in old age mortality associated with the fall in death rates from de-

Fig. 2.3: Trends of Life Expectancy (Japan versus Other Countries)



generative diseases. Old-age mortality improvement also contributes to the recent prolongation of life expectancy in Japan. Olshansky and Ault (1986) argues that this change of mortality improvement patterns should be regarded as a different stage from the third one by Omran (1971), and proposes the fourth stage of the epidemiologic transition: the age of delayed degenerative diseases.

Fig. 2.4 shows the trends of l_x curves for Japanese females after World War II. We observe that mortality improvement is caused by rectangularization before 1970. However, the improvement after 1970 could be regarded as a “delay” of death rather than rectangularization. This would imply there is an advantage in introducing a shifting feature of mortality curves to express the recent Japanese mortality.

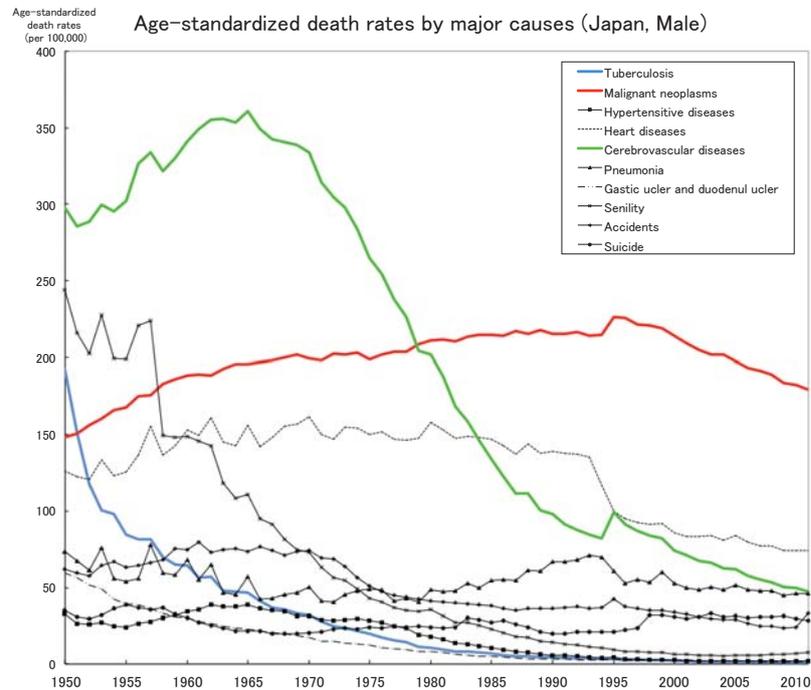
Fig. 2.4: Trends of l_x curves (Females, Japan)

2.2.2 Causes of Deaths

Fig. 2.5 shows the age standardized death rates for males by major causes since 1950 in Japan. The death rates from tuberculosis decreased remarkably in the 1950s–1960s. After the late 1960s, a decline in death rates from cerebrovascular diseases played a major role in increasing life expectancy. The death rates from malignant neoplasms increased gradually up to 1990 and have been decreasing since 1995.

Wilmoth (2011) argues that the longevity increase of the past two centuries is fundamentally a social phenomenon in which humans have *recognized* the causes of mortality, *reacted* by seeking means of averting or delaying such causes, and in this way *reduced* mortality rates across the age range. This pattern of *recognition/reaction/reduction* is an apt characterization of the process of mortality decline in various eras and in relation to various causes of death. Wilmoth calls it the

Fig. 2.5: Age-Standardized Death Rates by Major Causes (Males, Japan)



“triple R” theory of mortality improvement.

According to this theory, trends of the mortality rates by causes could be explained as follows: after the improvement of death rates from cerebrovascular diseases, malignant neoplasms were recognized as the next target, and people reacted to prevent or delay associated deaths, thus, finally succeeding in reducing mortality rates.

Therefore, it seems reasonable to expect that the Japanese mortality decline will continue into the future because our efforts should be similar to those in the past even though our focus may evolve.

Chapter 3

Mortality Projection Models for Japan

In this chapter, we review the mortality projection models for Japan. First, we review the general modeling for mortality in Section 3.1. Then, we review the official mortality projection for Japan in Section 3.2. The mortality model proposed in this dissertation is used in the official population projection in 2012. This is one of the reasons why we develop a new projections model. Lastly, Section 3.3 reviews the Lee–Carter model and its applications to Japan.

3.1 Modeling Age Patterns of Mortality

The intensity of mortality rates varies strongly with age. Generally, we can observe mortality rates more precisely if we disaggregate the rates into a single year of age. However, this produces more complexity because we must now observe a larger number of rates. It is convenient to simplify the life table functions by simple rules or a small number of tables, which are accomplished by modeling the age patterns of mortality.

There are three approaches for modeling mortality rates: (1) mathematical representations, (2) tabular representations, and (3) relational models. In this section, we review the mortality models for these three types.

3.1.1 Mathematical Representations

In the mathematical representation approach, some known mathematical functions are used to express the pattern of mortality rates as a function of age, also known as “the laws of mortality”. Table 3.1 shows some examples of mathematical representations. Although De Moivre proposed a simple function in 1725, the most well known is the Gompertz model that expresses an exponential function

for the force of mortality Gompertz (1825). Makeham (1860) adds a constant term to the Gompertz model. This model is used to obtain smoothed rates for old age mortality in the official life tables of Japan.

In old ages, the slopes of the $\log \mu(x)$ often taper and the exponential function overestimates the mortality rates. Perks (1932) and Beard (1971) propose logistic functions to improve fitness in old ages.

Table 3.1: Examples of Mathematical Representations

De Moivre (1725)	$\mu(x) = \frac{1}{\omega - x}$
Gompertz (1825)	$\mu(x) = B \cdot C^x$
Makeham (1860)	$\mu(x) = A + B \cdot C^x$
Thiele (1872)	$\mu(x) = a_1 e^{-b_1 x} + a_2 e^{-\frac{1}{2} b_2 (x-c)^2} + a_3 e^{b_3 x}$
Perks (1932)	$\mu(x) = \frac{A + B \cdot c^x}{1 + D \cdot c^x}$
Weibull (1951)	$\mu(x) = \alpha x^{\beta-1}$
Beard (1961)	$\mu(x) = \frac{B \cdot e^{ux}}{1 + D \cdot e^{ux}}$
Siler (1979)	$\mu(x) = a_1 e^{-b_1 t} + a_2 + a_3 e^{b_3 t}$
Heligman-Pollard (1980)	$\frac{q(x)}{p(x)} = A^{(x+B)^C} + D e^{-E(\log x - \log F)^2} + G H^x$
Rogers and Little (1993)	$y(x) = a_0 + m_1(x) + m_2(x) + m_3(x) + m_4(x)$

where

$$m_1(x) = a_1 \exp(-\alpha_1 x)$$

$$m_2(x) = a_2 \exp(-\alpha_2(x - \mu_2) - \exp(-\lambda_2(x - \mu_2)))$$

$$m_3(x) = a_3 \exp(-\alpha_3(x - \mu_3) - \exp(-\lambda_3(x - \mu_3)))$$

$$m_4(x) = a_4 \exp(\alpha_4 x)$$

$$y(x) = q(x), \frac{q(x)}{p(x)}, \mu(x)$$

Thatcher et al. (1998) compare the fitness of some models for old ages. They select six models: Gompertz; logistic (Perks); Kannisto; Weibull; Heligman and Pollard; and quadratic. They write the logistic model as

$$\mu(x) = c + \frac{a e^{bx}}{1 + \alpha e^{bx}}$$

and define the Kannisto model as

$$\mu(x) = \frac{ae^{bx}}{1 + ae^{bx}}$$

which is the simplest form of the logistic models noticed by Kannisto. The researchers find that the logistic model is most successful among the models, while the Kannisto model has a practical advantage in that it has only two parameters.

These models are considered in relation to old age mortality, whereas whole age models have also been proposed, such as Siler (1979), Heligman and Pollard (1980), and Rogers and Little (1994).

3.1.2 Tabular Representations

Tabular representations express arbitrary life tables by a set of tables that are induced from some empirical ones. The most familiar are the Coale and Demeny model life tables (Coale and Demeny 1983). This model has a parameter for mortality level labeled from 1 to 25 that corresponds to the women's e_0 , and a parameter for mortality shape labeled "north", "south", "east," and "west". Fig. 3.1 shows various levels of log mortality rates for the west model life tables. We observe that only the level of mortality changes and the shape of the mortality curve is kept. Fig. 3.2 shows the four shapes of the mortality rates in the model life tables for level 10.

Fig. 3.1: Log Mortality Rates (Coale and Demeny Model Life Tables)

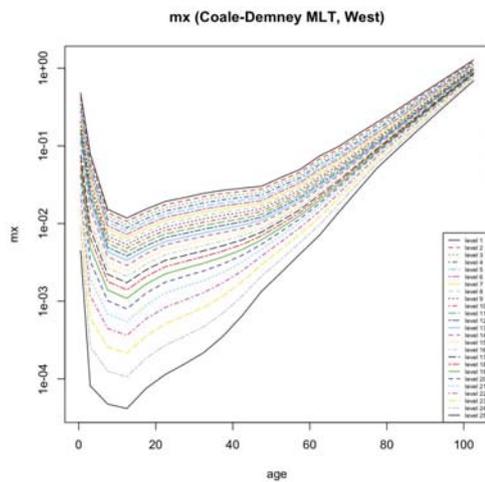
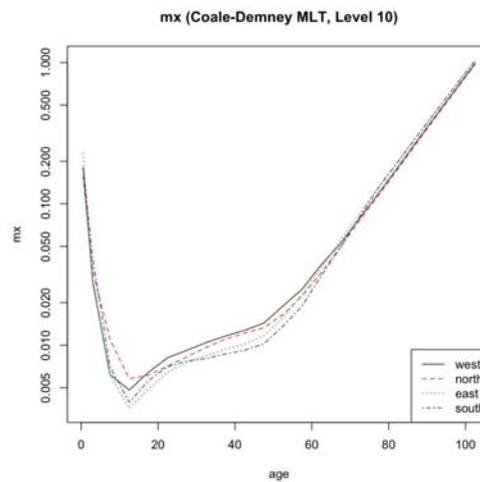


Fig. 3.2: Log Mortality Rates (Coale and Demeny Model Life Tables)



The model life tables have been used in the World Population Prospects. For

countries lacking recent or reliable information on age patterns of mortality, survival ratios are obtained from model life tables (United Nations 2006).

3.1.3 Relational Models

The mathematical representations have an advantage in that they have only a few parameters and simplify the life table functions, although the empirical life table functions are not generally expressed by some known mathematical functions. The tabular representations are free from this problem, whereas complexity increases when we require various levels and/or shapes of mortality.

The relational models could be considered to have advantages for both mathematical and tabular representations. The relational model expresses an arbitrary mortality pattern by a standard age pattern and the differences from the standard age pattern with some parameters.

Brass (1971) introduces the first relational model, known as the Brass Logit System, in which we define the function $Y_x = \log \left(\frac{l_x}{1-l_x} \right)$, that is, logit transformation of the l_x , and express Y_x^a , an arbitrary mortality pattern, as a linear combination $Y_x^a = \alpha + \beta Y_x^s$. Here, Y_x^s is the standard pattern, and α and β are the parameters for level and shape, respectively.

When relational models are used for time series of mortality rates, including mortality projections, the differences from the standard pattern mean changes of mortality over time. In this case, the changes are expressed by the differential of mortality by time. Therefore, the estimation of relational models for time series is equivalent to modeling the differential of mortality. This is one of the important concepts for building the LD model in the following discussions.

3.2 Review of the Official Mortality Projection for Japan

The National Institute of Population and Social Security Research (NIPSSR) prepares the official population projections for Japan. To project the cohort change of the population by deaths, future life tables are projected and set as assumptions for the projection.

The NIPSSR has been releasing population projections almost every five years since 1975. The methods for projecting mortality have changed over time and can be divided into three groups indicated in Fig. 3.3 as the 1st, 2nd, and 3rd Generations.

We assemble the mortality projection for 1981 and before into the 1st generation projection. The basic method used in the 1st generation involves projections by reference to optimal life tables. As we saw in Chapter 2, Japanese life expectancy in this period was the lowest of developed countries. Therefore, it was considered

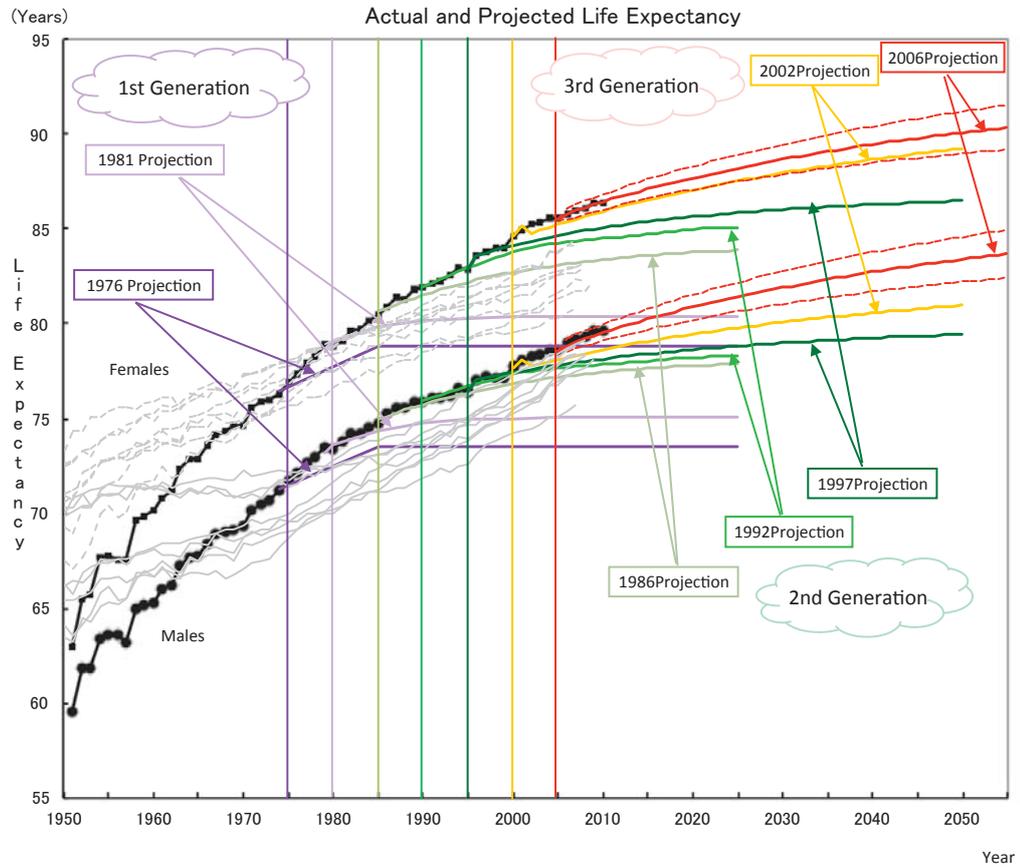
reasonable that the level of Japanese life expectancy would eventually approach that of the higher life expectancy countries.

However, Japanese life expectancy was prolonged rapidly and attained almost similar levels to those of other developed countries at about the end of the period of the 1st generation. Therefore, it became impossible to use the optimal life table method for projections for Japan. At the same time, the structure of deaths by cause had been changing dramatically. In the 1970s, mortality by cerebrovascular diseases started to decline remarkably and this helped to raise life expectancy. Thereafter, projection by cause of deaths was adopted as the method for the 2nd generation projections.

As Wilmoth (1995) points out, the projection by cause of death tends to underestimate life expectancy. The 2nd generation projections by NIPSSR are not exceptions. We can observe from Fig. 3.3 that projected life expectancy is lower than actual life expectancy.

Then, the relational models for the mortality for all causes of deaths are introduced in the 3rd generation. They use the LC model with some modification to fit Japanese mortality. We review these in Section 3.3.

Fig. 3.3: Comparison of Life Expectancy in the Official Projections



Source: MHLW "Life Tables", NIPSSR "Population Projection for Japan", Human Mortality Database
 The grey lines (solid: males, dotted: females) are life expectancy for the following countries:
 Canada, France, Italy, Netherlands, Norway, Sweden, UK, USA

3.3 The Lee–Carter Model and its Application to Japan

There are various models for mortality projection. Among them, the Lee–Carter (LC) model (Lee and Carter 1992) is now regarded as a standard method internationally. Tuljapurkar et al. (2000) applies this model to mortality in Group of 7 countries and demonstrates its effectiveness. Here, we briefly review a historical background of this model.

The LC model has a standard age pattern of mortality and expresses deviations from it with parameters, which is considered as a characteristic of the relational models reviewed in Section 3.1. This method is based on the preceding works for construction of model life tables. One of these studies is the UN model life tables originated in United Nations (1956) and revised in United Nations (1982). In United Nations (1982), the age patterns of mortality are stratified into clusters, each cluster having a distinct average age pattern of mortality, and then a principal components model is fitted to the deviations of each age pattern of mortality from its own cluster average. Coale and Demeny model life tables (Coale and Demeny 1983) reviewed in Subsection 3.1.2 is constructed using a similar idea. They regress ${}_nq_x$ and $\log_{10}({}_nq_x)$ on e_{10} by least square method as:

$${}_nq_x = A_x + B_x e_{10}$$

$$\log_{10}({}_nq_x) = A'_x + B'_x e_{10}$$

The meanings of the parameters A'_x and B'_x are similar to the parameters a_x and b_x used in the LC model.

This model is also based on the works of time-series modeling, in particular the ARIMA (autoregressive integrated moving average) time-series model by Box et al. (2013), first published in 1970. To project mortality rates with this model, we need to project a parameter k_t . Lee and Carter (1992) seek for an appropriate ARIMA time-series model for the parameter k_t and find that a random walk with drift describes the parameter well, which shows that this forecasting is certainly rooted in time-series modeling. This historical background confirms that the LC model is built on a long tradition in demography and statistics.

Here, we provide the definition of the model in detail. Let $\ln(m_{x,t})$ be the natural logarithm of central death rates. Then, the LC model is expressed as follows.

$$\ln(m_{x,t}) = a_x + k_t b_x + \epsilon_{x,t}$$

where a_x is the standard mortality age pattern and $\epsilon_{x,t}$ represents an error term. To estimate the parameters b_x and k_t , we apply singular value decomposition (SVD) to the matrix $\ln(m_{x,t}) - a_x$:

$$\ln(m_{x,t}) - a_x = \sum_i u_{xi} q_i v_{ti} \quad (q_1 \geq q_2 \geq \dots)$$

Then, we observe the term relating to q_1 (the first singular value), and set

$$k_t = q_1 v_{t1}$$

$$b_x = u_{x1}$$

Here, the parameter k_t is interpreted as an indicator for the level of mortality at time t , and b_x is the amount of mortality improvement at age x for the unit change of k_t . This interpretation means that the LC model expresses mortality improvement as a decline of the mortality rate for each age x .

For mortality projections, first, the future values of k_t are projected, and then, the future mortality rates $\ln(m_{x,t})$ are projected using the projected k_t values.

There are many studies that apply the LC model to Japanese mortality. Wilmoth (1996) applies it to Japanese total mortality (Method I), and compares the projection by forcing its future trend to match the projected Swedish trend (Method II) and the projections by cause-specific mortality (Methods III and IV).

Komatsu (2002) studies and develops the projection procedure by applying the LC method for the past Japanese official population projection in 2002 (NIPSSR 2002), known as the “Komatsu procedure.” This is the first mortality model in the 3rd generation for the official projection. Moreover, Ogawa et al. (2002), Nanjo and Yoshinaga (2003), Kogure and Hasegawa (2005), Ozeki (2005), and Oikawa (2006) study the application of the LC model to Japanese mortality. Recently, Igawa (2013) studies the residual structure involved in the application of the LC model to Japanese death rates, and proposes the LC–Vector Autoregressive (LC–VAR) model, which is an extended LC model. Moreover, Li et al. (2013) propose an extension of the LC model named as the LC method extended with rotation (LC_ER). We discuss their model in Chapter 6.

The Komatsu procedure, which is used in the past projection in 2002, applies a LC model that is slightly modified to suit Japanese mortality projections. The main differences are: (1) a_x is the average of the most recent two years in the Komatsu procedure, which is the average of the whole term in the original LC model and (2) k_t is projected by non-linear curve fitting in the Komatsu procedure¹, not by time-series analysis as in the original LC model.

However, comparing the projected life expectancy in the 2002 projection to the actual one in Fig. 3.3, we observe that the projected level of e_0 is less than the actual level. Detailed observation shows that mortality rates for older ages tend to be higher than the actual mortality rates.

In Lee and Miller (2001), the performance of the LC model in mortality projections is evaluated using data in the US, Canada, Sweden, France, and Japan. The study shows that the projected life expectancy using the LC model tends to be lower, especially if the projected period is lengthening. In addition, the study

¹The fitted function is the average of the exponential and the logarithm function, which is supposed to fit well with the recent trend in Japanese mortality.

points out that this tendency is in some way related to the changing age pattern of decline.

Wilmoth (1997) and Horiuchi and Wilmoth (1995) argue that the recent mortality improvement in old age for developed countries could be recognized as "aging of mortality decline," which implies an increase in the age of the most pronounced decline observed in several countries including Japan. If aging of mortality decline is observed, then the LC model performs poorly because the distribution of mortality improvement rates moves to an older direction, although it is assumed to be constant in the LC model.

On this point, we saw in Section 2.2 that we can regard the recent mortality improvement in Japan as a "delay" of death rather than rectangularization, which is related to a shifting of the mortality curve in the direction of older people. Therefore, seeking a model that could express an age-shifting structure would improve mortality projections.

As a mortality model with an age shifting structure, Bongaarts (2005) proposes the "shifting logistic model," noticing that the slope parameter in the three parameter logistic curve, which is fitted to the mortality data in each country, is almost constant over time.

Ishii (2006) studies the LC model with an age shifting structure, and states that the model has an advantage in fitting with Japanese old age mortality. The mortality projection model developed in Ishii (2006) is used in the 2006 official projection, which is a LC model with an age-shifting structure applied with the shift amount in the shifting logistic model.

Following these studies, we propose a new mortality projection model in this dissertation that is a more sophisticated version of the model in Ishii (2006).

Chapter 4

Data and Methods

In this Chapter, we describe the data and methods for the study of the new mortality models.

4.1 Data

For the purpose of the mortality study in this dissertation, we use mortality data for Japan from the Japanese Mortality Database (JMD).

We use

$$m_{x,t}, \quad x = x_s (= 0), \dots, x_e (= 150) \quad \text{and} \quad t = t_s (= 1970), \dots, t_e (= 2010)$$

where t is a calendar year. We extrapolate the mortality rates above age 110 years by fitting the two parameter logistic model

$$m_{x,t} = \frac{\alpha_t \exp(\beta_t x)}{1 + \alpha_t \exp(\beta_t x)}$$

which is the same method used in the JMD and HMD (Human Mortality Database).

Note that $m_{x,t}$ in the discrete form here represents the central death rate of age $[x, x + 1)$ in a calendar year t whose notations are different from the continuous form in which x and t represent exact age and time, respectively. Thus, $m_{x,t}$ in the discrete form is considered to be an approximation of $\mu_{x+0.5,t+0.5}$ in the continuous form. We use $\mu_{x,t}$ in theoretical discussion, whereas $m_{x,t}$ is used in numerical computations.

4.2 Methods

4.2.1 Formulation of Decline-type and Shift-type Models

In Section 3.3, we show that the Lee-Carter model, which expresses mortality improvement as a decline of mortality rates, has some disadvantages for modeling Japanese old age mortality, and that a model that could express an age-shifting structure would improve the projections.

Here, we reconsider and compare the performance of *decline*-type and *shift*-type models for the purpose of modeling Japanese adult mortality improvements. To compare which model is better quantitatively, we need mathematical formulations for both models.

To regard the mortality improvement as a *decline* means that the value of log mortality declines over time for a fixed age. In this case, the mortality improvement rates for an age decrease, or equivalently, the differential of the mortality improvement rates by time is negative. On the other hand, the delay of the age that attains to a fixed value of the log mortality is considered as a *shift* of the mortality curve. In this case, the age that corresponds to a fixed value of the log mortality increases, or equivalently, the differential of the age by time is positive.

Actually, the function that expresses the age that attains to a fixed value of the log mortality is the *inverse* function of the log mortality rates. Therefore, considering a *shift*-type model for the log mortality rates is identical to considering a *decline*-type model for the inverse function of the log mortality rates (Figs. 4.1 and 4.2).

Fig. 4.1: Log Mortality Rates (Female Japan)

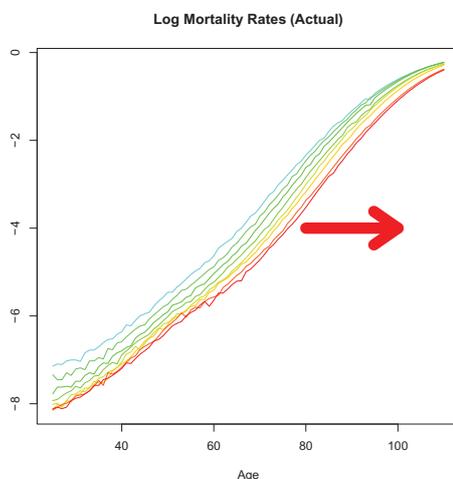
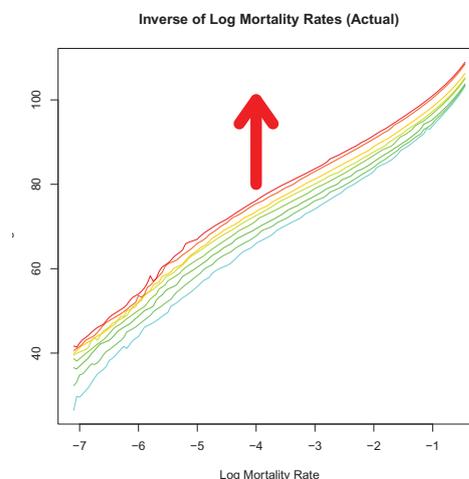


Fig. 4.2: Inverse Log Mortality Rates (Actual) for Female Japan



Based on these observations, we can state the mathematical formulations putting

special emphasis on the log mortality and its inverse functions, and the differential of them by time.

Let $X = [0, +\infty)$ be the space of age and $T = (-\infty, +\infty)$ be the space of time. In the following discussion for modeling mortality, we work on $\mu_{x,t}$, the hazard function for exact age $x \in X$ at time $t \in T$.

The log hazard function of mortality is expressed by $y = \lambda_{x,t} = \log \mu_{x,t}$, where $y \in Y = (-\infty, +\infty)$ is the value of the function. Then, the set $S = \{(x, t, y) | y = \lambda_{x,t}\}$ determines a surface in \mathbb{R}^3 , called the *log mortality surface*. This is a conventional representation of the log mortality surface. In this representation, $y = \lambda_{x,t}$ is considered as the height from the X - T plane in \mathbb{R}^3 .

Here, we consider another representation of the log mortality surface under a set of assumptions.

We assume that $\lambda_{x,t}$ is a smooth continuous function with respect to x and t defined on $X_0 \times T_0 = [0, \omega] \times [t_0, t_1] \subset X \times T$, where $\omega < +\infty$ is a finite maximum age for mortality models.

For the purpose of modeling *adult* mortality, we further assume that $\lambda_{x,t}$ exhibits a strictly monotonic increase with respect to x for each t and $x > x_0(t)$. Here, $x_0(t)$ represents the lower bound of x , above which $\lambda_{x,t}$ exhibits a strictly monotonic increase for each t . Then, for each t , the function $\lambda_t(x)$ defined by

$$\lambda_t : \tilde{X}_t \rightarrow Y, \quad \lambda_t(x) \stackrel{\text{def}}{=} \lambda_{x,t}$$

is an injective (one to one) function of x , where $\tilde{X}_t = [x_0(t), \omega]$. Let $\tilde{Y}_t = \lambda_t(\tilde{X}_t)$, then $\lambda_t(x) : \tilde{X}_t \rightarrow \tilde{Y}_t$ has an inverse function $\nu_t(y) : \tilde{Y}_t \rightarrow \tilde{X}_t$ defined on \tilde{Y}_t for each t .

Let us define Y_0 as follows:

$$Y_0 \stackrel{\text{def}}{=} [y_0, y_1], \quad \text{where } y_0 = \sup_{t \in T_0} \min \tilde{Y}_t, \quad y_1 = \inf_{t \in T_0} \max \tilde{Y}_t$$

Then, we can define $\nu_{y,t} : Y_0 \times T_0 \rightarrow X_0$ by

$$\nu_{y,t} \stackrel{\text{def}}{=} \nu_t(y)$$

$\nu_{y,t}$ gives the age x at which the value of the log hazard function is equivalent to a value y at time t .

Moreover, we define the following two differential functions by time t : (1) $\rho_{x,t}$: the mortality improvement rate and (2) $\tau_{y,t}$: the force of age increase.

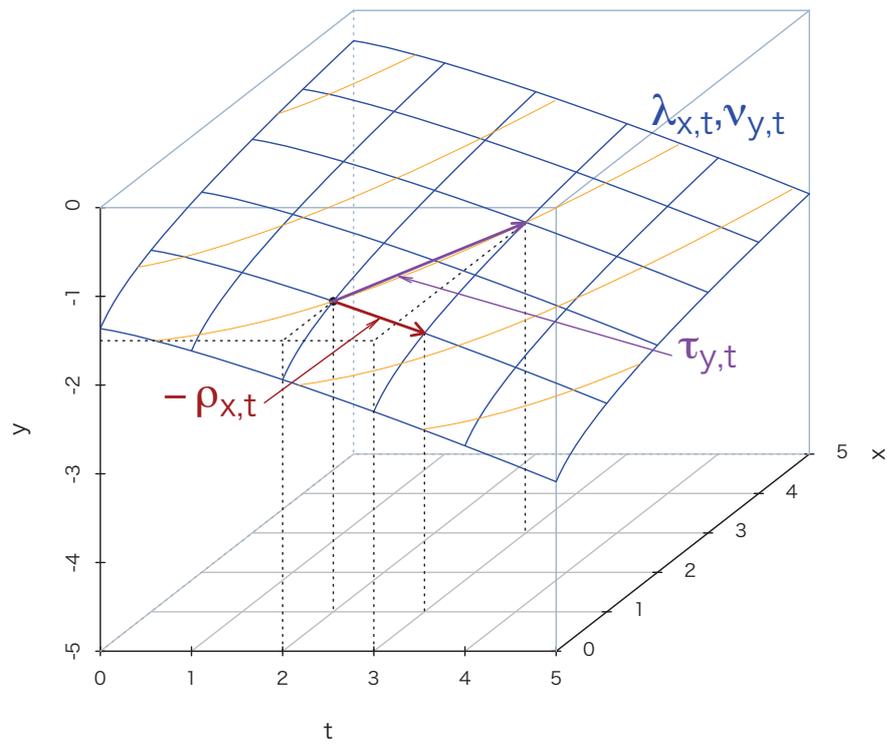
$$\rho_{x,t} \stackrel{\text{def}}{=} -\frac{\partial \lambda_{x,t}}{\partial t} = -\frac{\partial \log \mu_{x,t}}{\partial t}$$

$$\tau_{y,t} \stackrel{\text{def}}{=} \frac{\partial \nu_{y,t}}{\partial t}$$

Fig. 4.3 shows a stylized example of the log mortality surface and the two differential functions $\rho_{x,t}$ and $\tau_{y,t}$. The blue lines show the log mortality surface in the usual representation, that is, the height from the X - T plane which is determined by $\lambda_{x,t}$. The black point on the log mortality surface is $(x, t, y) = (1, 2, -1.5)$, which can be recognized also as the height from the X - T plane of -1.5 when $(x, t) = (1, 2)$. If we travel on the surface with x fixed, the height from the X - T plane will decrease to around -1.86 when $t = 3$, which is shown by the brown arrow. The difference between the two heights corresponds to $-\rho_{x,t}$.

On the other hand, the log mortality surface is also represented by the height from the Y - T plane, which is determined by $\nu_{y,t}$. From this viewpoint, the black point recognizes that the height from the Y - T plane is 1 when $(y, t) = (-1.5, 2)$. The orange lines on the surface show the contour with y fixed, so we proceed along these lines when traveling on the surface with y fixed. If we start again from the black point but this time keep y fixed, the height from the Y - T plane will be 3 when $t = 3$, which is shown in the purple arrow in Fig. 4.3. The difference between the two heights corresponds to $\tau_{x,t}$.

Fig. 4.3: Log Mortality Surface and Two Differential Functions



4.3 Descriptions of Decline-Type and Shift-Type Models

In this section, we describe the *decline*-type and *shift*-type adult mortality models. First, we describe the definitions of the proportional hazard model and the LC model, which are considered *decline*-type models. Then, we introduce the horizontal shifting model and the horizontal LC model, which are considered *shift*-type models and correspond to the two *decline*-type models. In this chapter, we propose a new *shift* type of adult mortality model that can favorably express recent Japanese old age mortality.

4.3.1 Decline-Type Mortality Models

The Proportional Hazard (PH) Model

The PH model is a simple model that expresses mortality improvement as *decline*. In the PH model, $\lambda_{x,t}$, the log hazard rate function at time t is expressed by

$$\lambda_{x,t} = \log \mu_{x,t} = a_x + k_t$$

where a_x represents the baseline logged hazard rates. We drop the error term hereafter.

In the PH model, $\rho_{x,t}$, which is the rate of mortality improvement, is expressed as follows.

$$\rho_{x,t} = -\frac{dk_t}{dt} = -k'_t$$

Therefore, it is constant with respect to age. This is the differential form for this model.

The Lee–Carter (LC) Model

The LC model is expressed by the following formula (Lee and Carter 1992) as shown in Chapter3.

$$\lambda_{x,t} = \log \mu_{x,t} = a_x + k_t b_x$$

where a_x is a standard age pattern of mortality.

Taking a partial derivative by time t , we obtain the following relationship.

$$\rho_{x,t} = -\frac{dk_t}{dt} b_x = -k'_t b_x \quad (4.1)$$

Equation 4.1 shows that the age distribution of $\rho_{x,t}$ is constant in the LC model. If we further assume that k_t is linear over time, then $\rho_{x,t}$ is constant over time. Therefore, the LC model works well when the age specific rate of mortality improvement

is considered to be constant over time, that is, the mortality improvement is considered as a *decline*.

4.3.2 Shift-Type Mortality Models

The Horizontal Shifting (HS) Model

Next, we discuss models that express mortality improvement through a *shift*. The simplest model for *shifting* would be one whereby the entire log hazard curve moves to the right-hand side. We can restate this model using the inverse function of log hazard mortality $\nu_{y,t}$, that is, the proportional hazard model for $\nu_{y,t}$.

This HS model is expressed formally as follows:

$$\nu_{y,t} = a_y + k_t$$

In the differential form,

$$\tau_{y,t} = \frac{dk_t}{dt} = k'_t$$

Parameter estimation for the HS model is completely identical to the PH models, except for adapting these procedures to $\nu_{y,t}$ instead of $\lambda_{x,t}$.

The Horizontal Lee–Carter (HL) Model

As we consider the LC model, which admits a different amount of decline by age and provides a more general framework compared to the PH model, we can also consider the LC model for $\nu_{y,t}$, which in turn supports a more general shifting feature. We call this the HL model.

$$\nu_{y,t} = a_y + k_t b_y$$

In the differential form,

$$\tau_{y,t} = \frac{dk_t}{dt} b_y = -k'_t b_y$$

4.3.3 The Linear Difference (LD) Model

First, we show the following property of $\tau_{y,t}$ for the two parameter logistic model. It is a theoretical foundation of the LD model.

Proposition 1. *For the two parameter logistic model*

$$y = \lambda_{x,t} = \log \frac{\alpha_t \exp(\beta_t x)}{1 + \alpha_t \exp(\beta_t x)} = \log \alpha_t + \beta_t x - \log(1 + \alpha_t \exp(\beta_t x))$$

$\tau_{y,t}$ is a linear function of x for each t , that is,

$$\tau_{y,t} = f'_t + g'_t x$$

Proof.

$$e^y = \frac{\alpha_t \exp(\beta_t x)}{1 + \alpha_t \exp(\beta_t x)}$$

$$\Leftrightarrow \alpha_t \exp(\beta_t x) = \frac{e^y}{1 - e^y}$$

By differentiating both sides by t with y fixed, we obtain

$$\alpha'_t \exp(\beta_t x) + \alpha_t \left(\beta'_t x + \beta_t \frac{\partial x}{\partial t} \right) = 0$$

$$\Leftrightarrow \frac{\partial x}{\partial t} = -\frac{\alpha'_t}{\alpha_t \beta_t} - \frac{\beta'_t}{\beta_t} x$$

□

Since the actual old mortality rates in the HMD and JMD are estimated using the two parameter logistic model, $\tau_{y,t}$ for the old age that we use is expected to satisfy this relationship.

Let us define the LD model satisfying this property, that is, $\tau_{y,t}$ is a linear function of x for each t . Then, we can describe the condition for the model in the continuous form as follows.

$$\tau_{y,t} = f'_t + g'_t x$$

This is the differential form. By integrating both sides with t , we obtain

$$\nu_{y,t} = f_t + g_t x + a_y$$

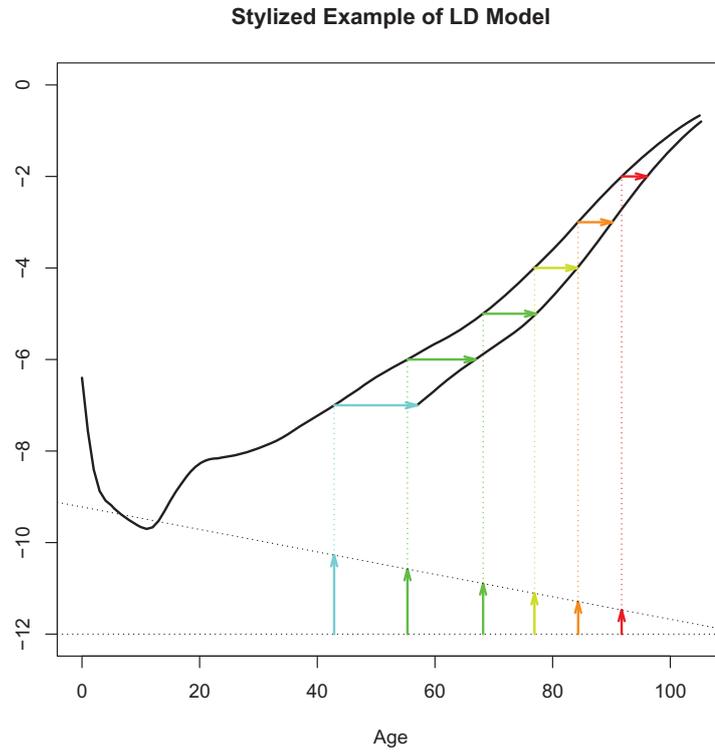
where a_y denotes a standard pattern of inverse log hazard rates.

Fig. 4.4 shows the stylized example of the LD model. The colored horizontal arrows in the upper half show the extent of the shift of the mortality curve during a short period that corresponds to $\tau_{y,t}$. The vertical arrows at the bottom have the same lengths as those on the upper side with the same color whose directions are rotated 90 degrees counter-clockwise. The LD model requires that the extent of the shift is a linear function of age, which means the end point of the arrows form a straight line, as shown by the dotted line.

The parameter g'_t refers to the slope of the dotted line, which indicates that g'_t is negative in this example. Therefore, g_t declines during this short period. If the mortality curve shifts completely parallel, then the slope of the dotted line becomes zero, and thus, $g'_t = 0$ and g_t stay constant. Therefore, we interpret the parameters g'_t and g_t as

- compression $\Leftrightarrow g'_t < 0$ and g_t : decline
- parallel $\Leftrightarrow g'_t = 0$ and g_t : constant

Fig. 4.4: Stylized example of the LD model



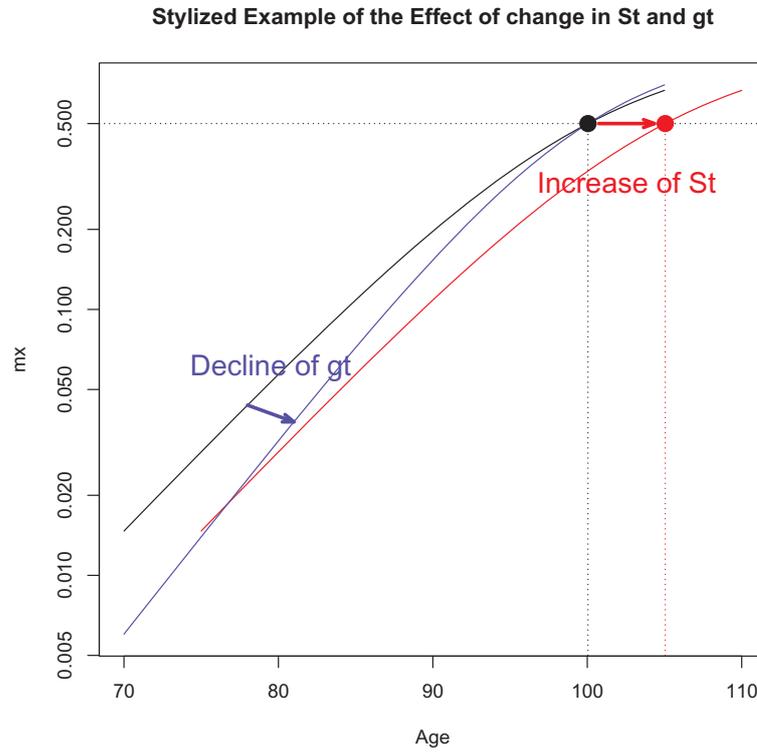
- decompression $\iff g'_t > 0$ and g_t : increase

The absolute value of g_t means the slope of the difference of age increases at time t compared to the baseline mortality that is considered close to the midpoint of the base period t_0 . Therefore, the negative sign of g_t indicates the mortality curve at time t is more compressed compared to the baseline mortality for $t > t_0$ and less compressed for $t < t_0$.

Here, we introduce another variable, S_t , as a location of the mortality curve instead of f'_t or f_t . f_t is the intercept of the line formed by the arrows in Fig. 4.4. It depends on the value of the slope and is not easy to interpret. Here, we define S_t as the age that the mortality rate equals 0.5 at time t . In the Kannisto model,

$$\mu(x, t) = \frac{\alpha(t)e^{\beta(t)x}}{1 + \alpha(t)e^{\beta(t)x}}$$

the age x that satisfies $\mu(x, t) = 0.5$ is $x = -\frac{\log \alpha(t)}{\beta(t)}$, which is the inflection point of the logistic curve and is considered preferable for measuring the location of the mortality curve, excluding the effects of the slope of the mortality. Actually,

Fig. 4.5: Stylized example of the Effect of Change in S_t and g_t 

the shifting logistic model proposed by Bongaarts (2005) measures the extent of the shifting between t_0 and t by $-\frac{\log(\alpha(t)/\alpha(t_0))}{\beta}$, assuming $\beta(t)$ is constant. This is equivalent to the difference of $S(t)$ in our notation and also supports the validity of this parameter as a location of the curve.

In theory, we can convert from S_t to f_t using the value of g_t with the following formula:

$$f_t = (1 - g_t)S_t - S_{t_0} + \int_{t_0}^t g_s S'_s ds$$

This formula is derived from $S_t = S_{t_0} + f_t + g_t f_t - \int_{t_0}^t g_s S'_s ds$, which, in turn, is obtained from the integration of $S'_t = f'_t + g'_t S_t$, which is the definition of the LD model $\tau_{y,t} = f'_t + g'_t x$.

The above formula is approximated by the formula $f_t \approx (1 - g_t)S_t - S_{t_0}$ when the difference between t_0 and t is not large. For Japanese females, $S_t \approx 100$, $S'_t \approx 0.1$, and $|g_t| < 0.25$. Then, if $t - t_0 = 50$, we obtain $\int_{t_0}^t |g_s S'_s| ds < 1.25 \ll S_t$. Therefore, this is a fairly good approximation for ordinary mortality projections and can avoid the process of numerical integrations.

For numerical computation in the discrete form, we can use the following recursive formula that is derived from $S'_t = f'_t + g'_t S_t$

$$f_{t_2} \approx f_{t_1} + (S_{t_2} - S_{t_1}) - (g_{t_2} - g_{t_1}) \frac{S_{t_2} + S_{t_1}}{2}$$

Fig. 4.5 shows a stylized example of the effect of change in S_t and g_t . Assuming that the mortality curve at a base year is shown as the black line, an increase of S_t with g_t fixed changes the curve into that shown as the red line. Therefore, we can recognize the mortality improvement by the increase of S_t as the shifting of the mortality curve. On the other hand, a decline of g_t with S_t fixed changes the curve into that shown as the blue line, which exhibits some compression features of mortality during the improvement.

In the following chapters, we fit these models to actual Japanese old age female mortality for Japanese females and compare them.

4.4 Methods for Parameter Estimations

Next, we discuss the computational methods for parameter estimation and some variants of the LD model.

First, we need to estimate the inverse function of $\lambda_{x,t}$. Section 4.1 noted that we use $m_{x,t}$ in numerical computations, and we estimate $\nu_{y,t}$ as the inverse function of $\log(m_{x,t})$ for $x \geq 25$. To estimate the inverse function, we estimate the age x that corresponds to a value of log mortality y using linear interpolation for

$$y = -10.00, -9.99, -9.98, \dots, -0.02, -0.01$$

for each t when the value of log mortality is available. Then we set the domain of y as the maximum interval that the values of log mortality are available throughout the years from $t_s (= 1970)$ to $t_e (= 2010)$. For Japanese females, the domain is $[y_s, y_e] = [-7.15, -0.01]$.

In addition, we need the differential functions of $\lambda_{x,t}$ and $\nu_{y,t}$. We use the following approximations for the two differential functions in terms of statistical stability.

$$\rho_{x,t} \approx -\frac{\lambda_{x,t+2} - \lambda_{x,t-2}}{4}$$

$$\tau_{y,t} \approx \frac{\nu_{y,t+2} - \nu_{y,t-2}}{4}$$

Therefore, we can compute $\rho_{x,t}$ and $\tau_{y,t}$ only for $t = 1972, \dots, 2008$.

For the two decline models, we need baseline log mortality a_x , whereas we need baseline inverse log mortality a_y for the three shift models. For comparisons of the models in Chapter 5, we use the average from 1970 to 2010 as the baselines,

which represents the same method used in the original LC model. However, it is favorable to use a recent pattern for the purpose of the mortality projections. Therefore, we use the average from 2006 to 2010 in the application to mortality projections in Chapter 6.

Then, the parameters k_t in the PH and HS models are estimated using the least square method. For the LC and HL models, the parameters k_t and b_x or b_y are usually estimated by singular value decomposition, which is essentially equivalent to the least square method. Therefore, we also use the least square method for estimation of the LD model. We can rewrite the normal form of the LD model as $v_{y,t} - a_y = g_t x + f_t$, which means that the difference of the inverse log mortality and the baseline is a linear function of x . Then,

$$\begin{bmatrix} v_{y_0,t_s} - a_{y_0} \\ v_{y_0+1,t_s} - a_{y_0+1} \\ \vdots \\ v_{y_e,t_s} - a_{y_e} \\ v_{y_0,t_s+1} - a_{y_0} \\ v_{y_0+1,t_s+1} - a_{y_0+1} \\ \vdots \\ v_{y_e,t_s+1} - a_{y_e} \\ \vdots \\ v_{y_0,t_e} - a_{y_0} \\ v_{y_0+1,t_e} - a_{y_0+1} \\ \vdots \\ v_{y_e,t_e} - a_{y_e} \end{bmatrix} = \begin{bmatrix} v_{y_0,t_s} & 1 \\ v_{y_0+1,t_s} & 1 \\ \vdots & \vdots \\ v_{y_e,t_s} & 1 \\ & v_{y_0,t_s+1} & 1 \\ & v_{y_0+1,t_s+1} & 1 \\ & \vdots & \vdots \\ & v_{y_e,t_s+1} & 1 \\ & & \ddots \\ & & & v_{y_0,t_e} & 1 \\ & & & v_{y_0+1,t_e} & 1 \\ & & & \vdots & \vdots \\ & & & v_{y_e,t_e} & 1 \end{bmatrix} \begin{bmatrix} g_{t_s} \\ f_{t_s} \\ g_{t_s+1} \\ f_{t_s+1} \\ \vdots \\ g_{t_e} \\ f_{t_e} \end{bmatrix} + \epsilon \quad (4.2)$$

Rewriting this formula as $Y = X\beta + \epsilon$, we can estimate $\hat{\beta}$ as

$$\hat{\beta} = (X'X)^{-1}X'Y$$

by the least square method. Thus, we can estimate the parameters f_t and g_t . We call this the naive LD method. When we compare the performances of the models in Chapter 5, we use the naive LD method.

For the purpose of mortality projection, we use a slightly modified version of the LD method. The first modification is the range of the data. In Chapter 5, we show the LD model's performance is better than that of the LC model for ages over about 75 years old. Following this observation, we propose a blended mortality model that has the LC property in youth and the LD property in older age in Chapter 6. To obtain better performance for the LD part of the model, we restrict the range of the data only for older ages. We compute y_{sm} , which is the minimum value of log mortality at 60 years of age during the period, and y_{em} , which is the

maximum value at 120 years of age. Then, we set the domain of $\nu_{y,t}$ as $[y_{sm}, y_{em}]$ instead of $[y_s, y_e]$. For Japanese females, $[y_{sm}, y_{em}] = [-5.69, -0.06]$.

However, we can only estimate the model that the log mortality rates lie in $[y_{sm}, y_{em}]$ by this method, and we need the estimates outside $[y_{sm}, y_{em}]$. We can obtain them if we extrapolate $\nu_{y,t}$ assuming that the relationship $\nu_{y,t} = f_t + g_t x + a_y$ holds outside of $[y_{sm}, y_{em}]$ in theory. However, this relationship does not hold precisely. Therefore, we repeat the process of parameter estimation by plugging the estimated $\nu_{y,t}$ in the left-hand side of Equation (4.2) into $\nu_{y,t}$ in the right-hand side until it reaches convergence. This is the second modification. We call this the modified LD method.

Chapter 5

Linear Difference Model

In this chapter, we fit the five mortality models to Japanese female old age mortality and examine which model is appropriate for the projection of Japanese mortality.

5.1 Fitting the Mortality Models

5.1.1 Fitting Decline-Type Mortality Models

The Proportional Hazard Model (PH)

First, we fit the PH model to the actual mortality rates. Here, we set a_x as the average log hazard rate for the entire period.

Fig. 5.1 shows the actual log hazard rates ($\lambda_{x,t}$) and the estimated rates with the PH model. We observe that the estimated rates do not exhibit good fit, particularly in the older age groups. Fig. 5.2 shows the difference between the actual and estimated rates. From this graph, we see that the actual values are higher than those of the model for the age range of about 60–80 years in 1970, whereas these values are decreasing over time. However, the opposite movement is observed for ages over 90 years. This is caused by a limitation of the PH model, whereby the rate of mortality improvement is constant with respect to age.

Fig. 5.1: Mortality Rates (Actual and Model, PH)

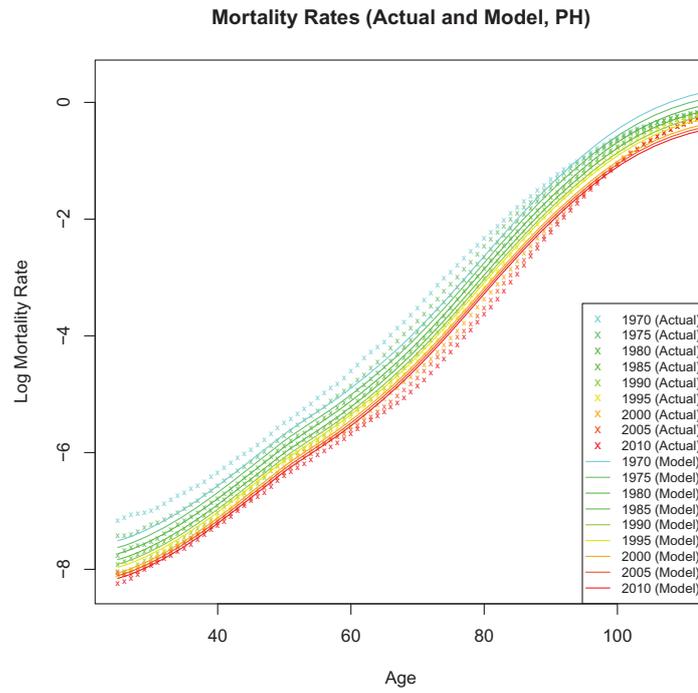
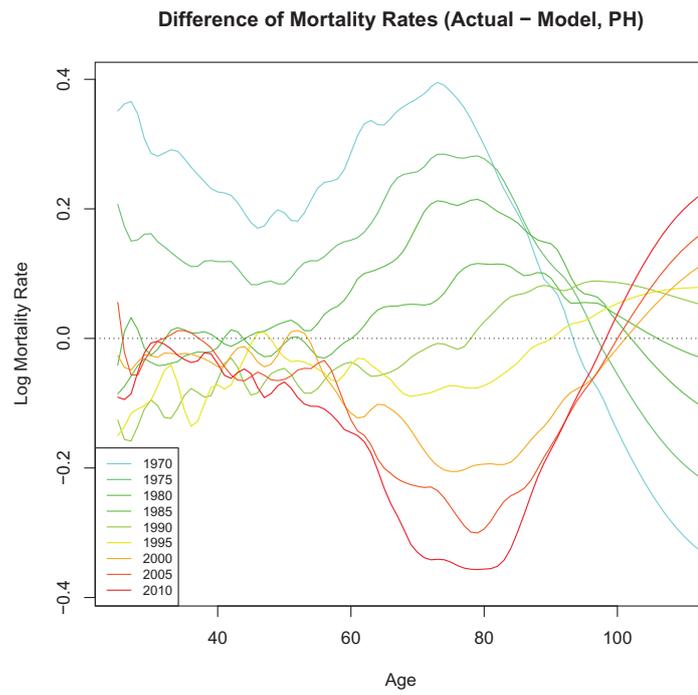


Fig. 5.2: Difference of Mortality Rates (Actual - Model, PH)



The Lee–Carter Model (LC)

Fig. 5.3 shows the actual log hazard rates ($\lambda_{x,t}$) and the estimated rates by the LC model. This figure illustrates that the fit with the actual values is fairly improved by using the LC model because of its flexibility, which allows different mortality improvement rates by age.

However, we can observe from Fig. 5.4 that the difference between the actual and estimated rates exhibits a trend, whereby the actual values are higher in younger age groups and lower in older age groups near the beginning and the end of the entire period; whereas the opposite is true around the middle of the period.

The reason why this trend for the error components is observed is ascribed to the change in the age specific mortality improvement rates over time. Therefore, we examine the $\rho_{x,t}$ functions for these two models next.

Fig. 5.3: Mortality Rates (Actual and Model, LC)

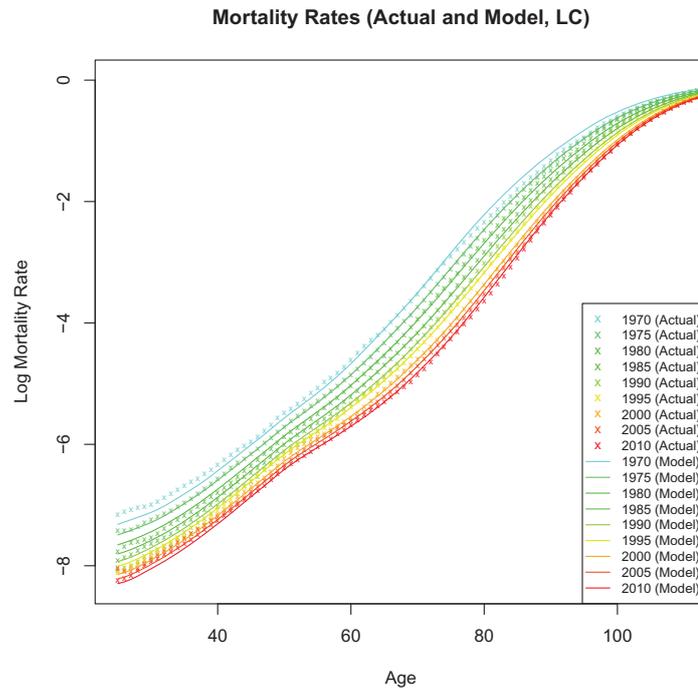
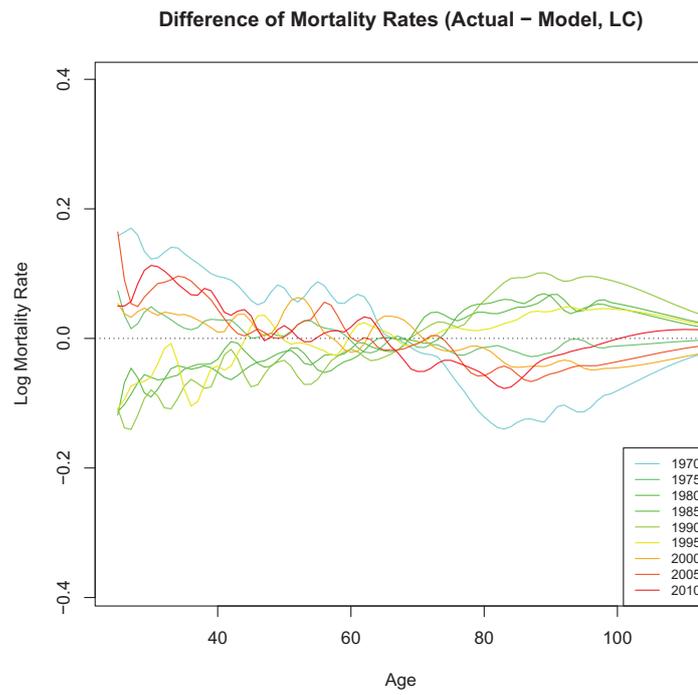


Fig. 5.4: Difference of Mortality Rates (Actual - Model, LC)



Figs. 5.5 and 5.6 show the $\rho_{x,t}$ functions for the actual and estimated values for each of the two models. The blue lines show the $\rho_{x,t}$ by the actual mortality rates. We observe that most of the mortality improvement rates have mountain-shaped curves with peaks. In contrast, the mortality improvement rates under the PH model, expressed by the pink line, are horizontal. This difference in shape is viewed as a cause of the estimates by the PH model being not well fitted, as we observed in Fig. 5.2.

The peak of the mortality improvement rate by the LC model, indicated by the green curves, is like that of the actual value and this improves the fit, as we have seen in Fig. 5.4. However, the age distribution of the rates is fixed in the LC model, whereas it changes dynamically in the actual values.

Thus, the actual age distribution of mortality improvement rates changes over time and is not constant as in the LC model, thus, causing the propensity for error in the LC model, as observed in Fig. 5.4. We could view this result as a limitation when the mortality improvement is considered as a *decline*.

Fig. 5.5: Comparison of Mortality Improvement Rates (1975–1990)

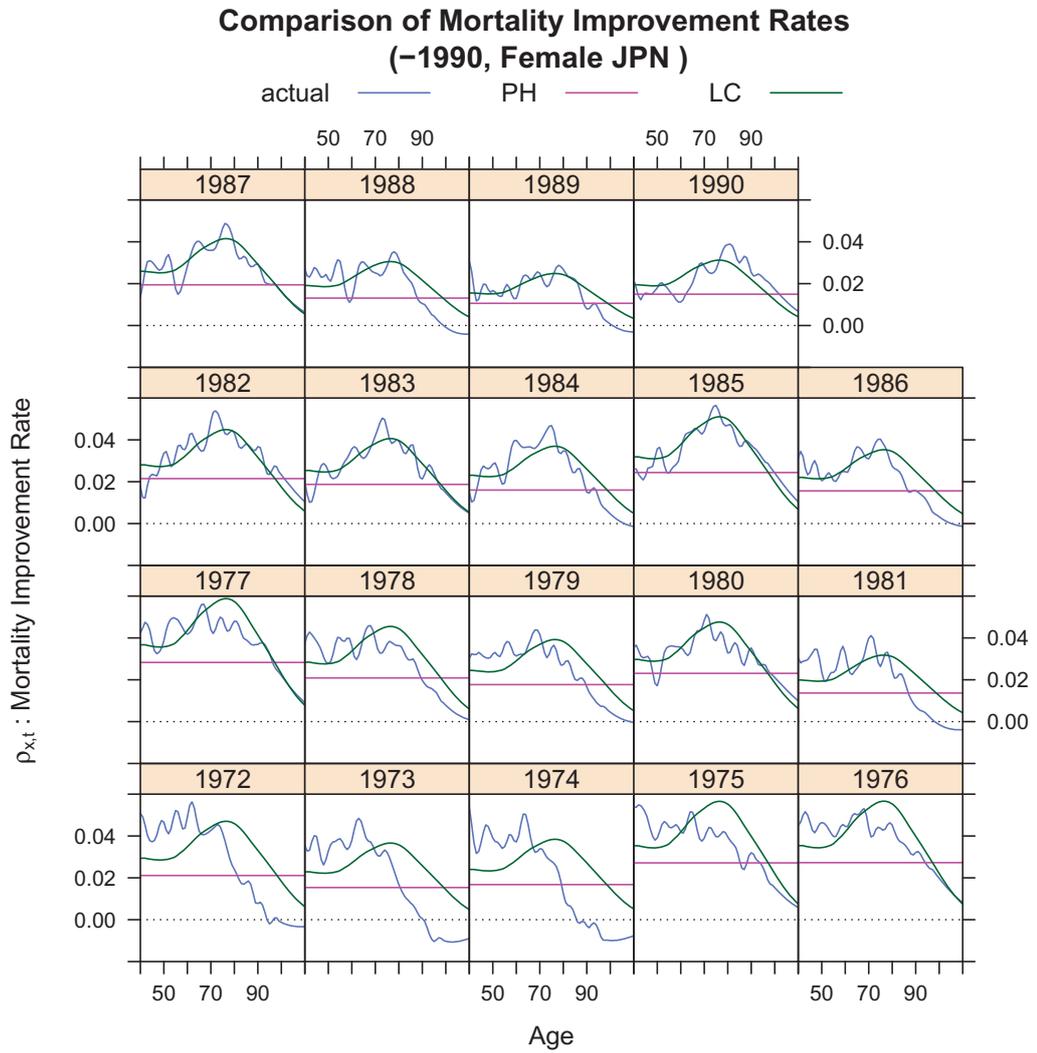
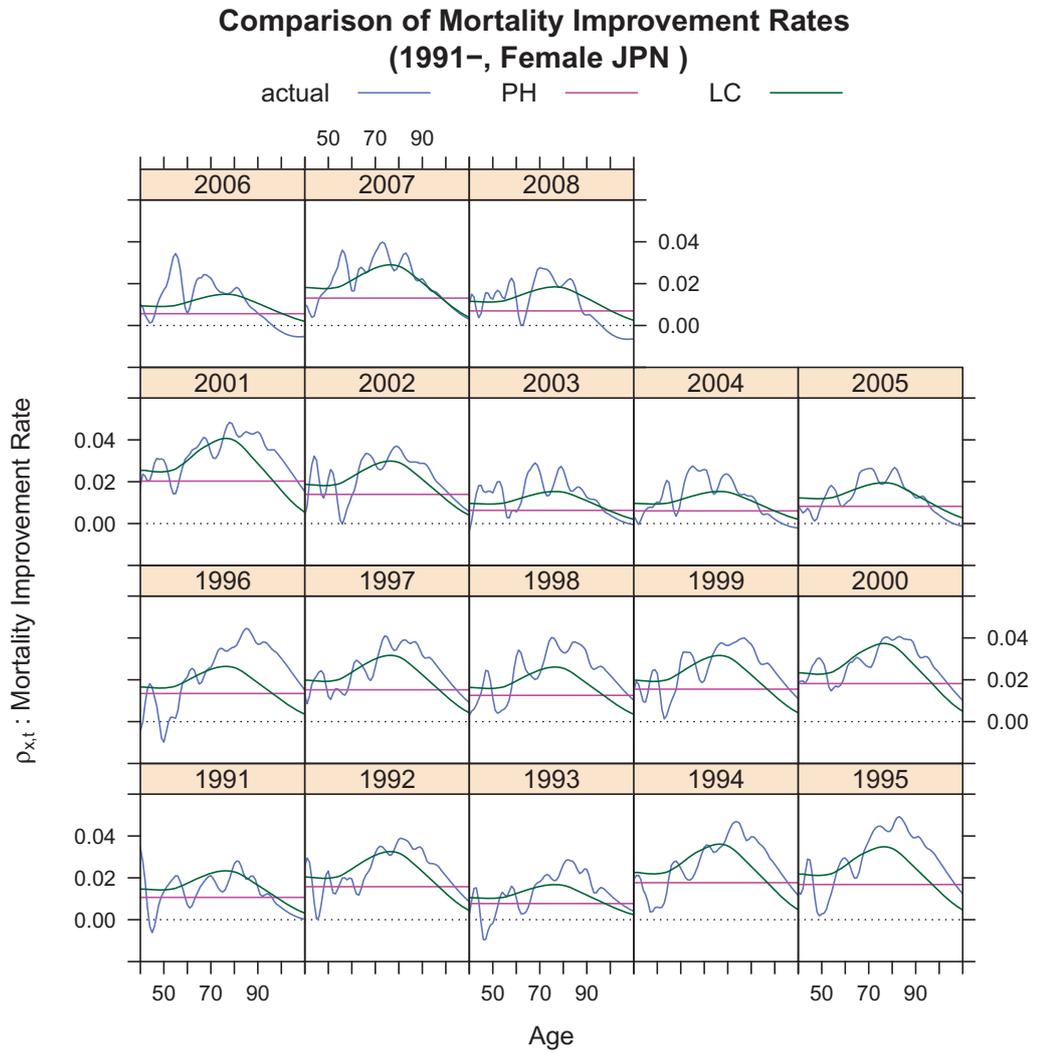


Fig. 5.6: Comparison of Mortality Improvement Rates (1991–2006)



5.1.2 Fitting Shift-Type Mortality Models

The Horizontal Shifting (HS) Model

Next, we fit the *shift* models to actual mortality. First, we consider the HS model.

Parameter estimation for the HS model is completely identical to the PH models, except for adapting these procedures to $\nu_{y,t}$ instead of $\lambda_{x,t}$. Fig. 5.7 shows the actual inverse mortality rates ($\nu_{y,t}$) and the estimated rates by the HS model, and Fig. 5.8 is the difference between the actual and the estimated rates.

We see that the performance of fit by the HS model is much better than that by the PH model, even though both have the same structure. For 1970, which is indicated by the light blue line, the actual values are higher in younger ages and lower in older ages, although the errors are not as high for other years.

Fig. 5.7: Inverse Mortality Rates (Actual and Model, HS)

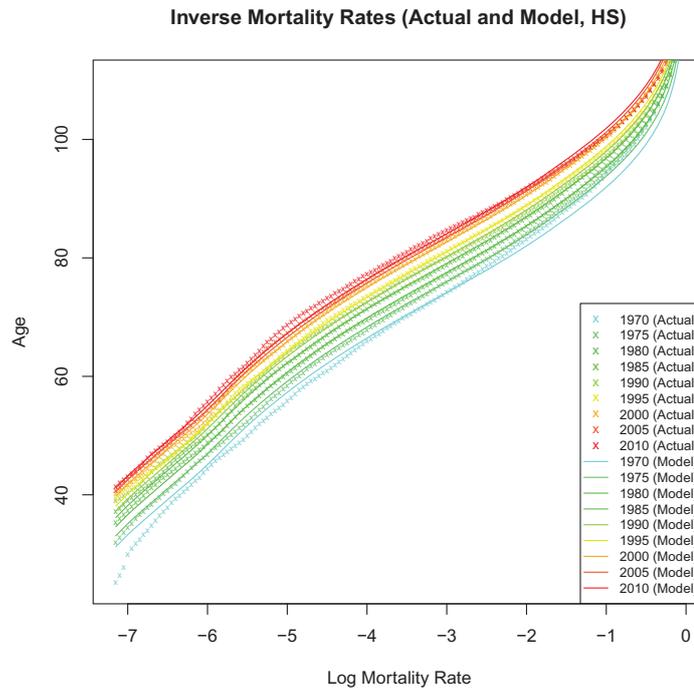
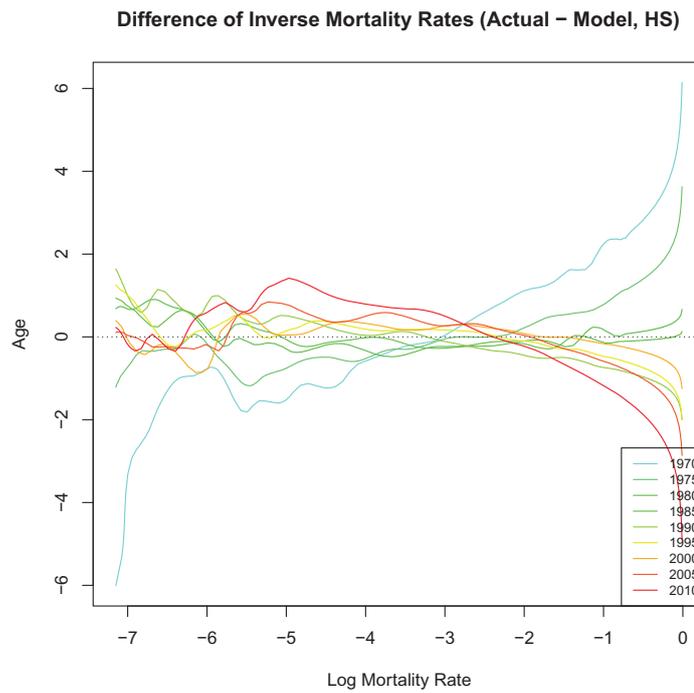


Fig. 5.8: Difference of Inverse Mortality Rates (Actual and Model, HS)



The Horizontal Lee-Carter Model (HL)

Next, we fit the HL model.

Fig. 5.9 shows actual inverse mortality rates ($v_{y,t}$) and the estimated rates under the HS model, and Fig. 5.10 shows the difference between the actual and the estimated rates. We see that the HL model seems to be improved compared to the HS model. However, we also observe that the improvement between the *shift* pair is not as large as the *decline* pair. This means that relaxing the limitation, in which the force of age increase in the HS model is restricted to the constant function, does not cause significant improvement of fit in the HL model. This could be explained by the difference in the shape of $\tau_{y,t}$, the force of age increase.

Fig. 5.9: Inverse Mortality Rates (Actual and Model, HL)

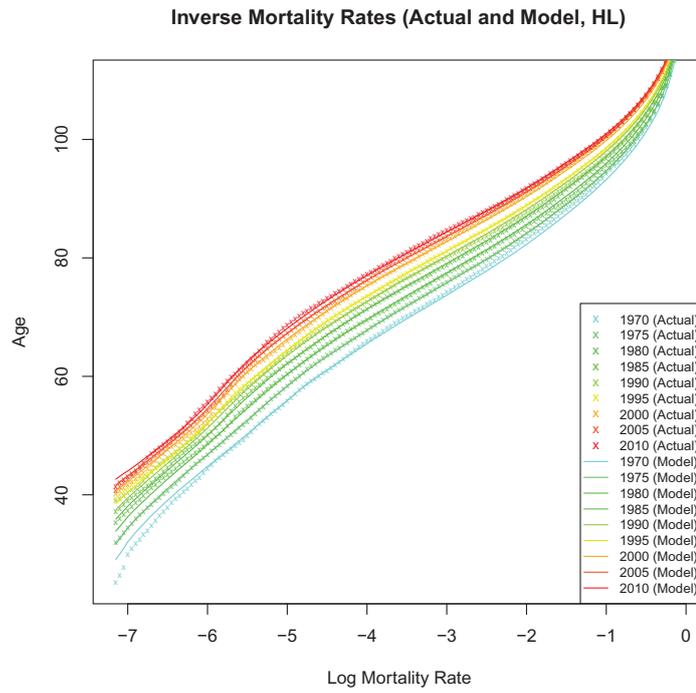
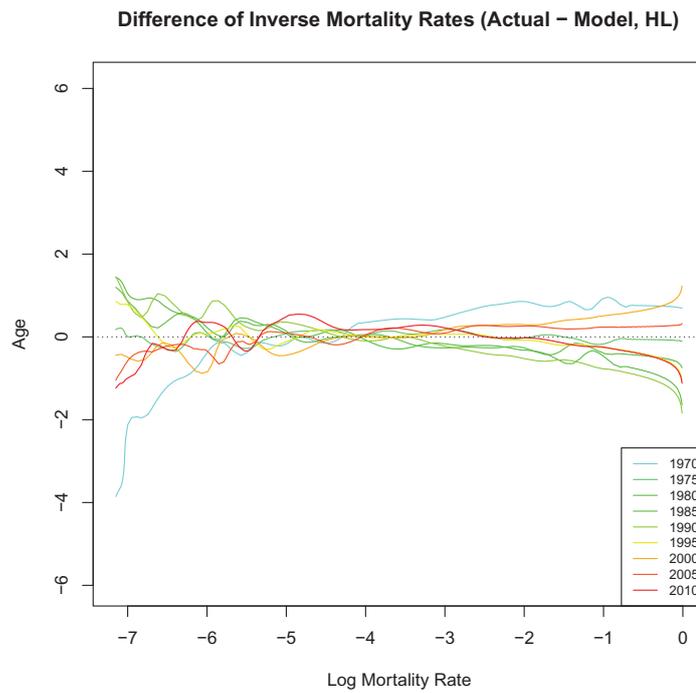


Fig. 5.10: Difference of Inverse Mortality Rates (Actual and Model, HL)

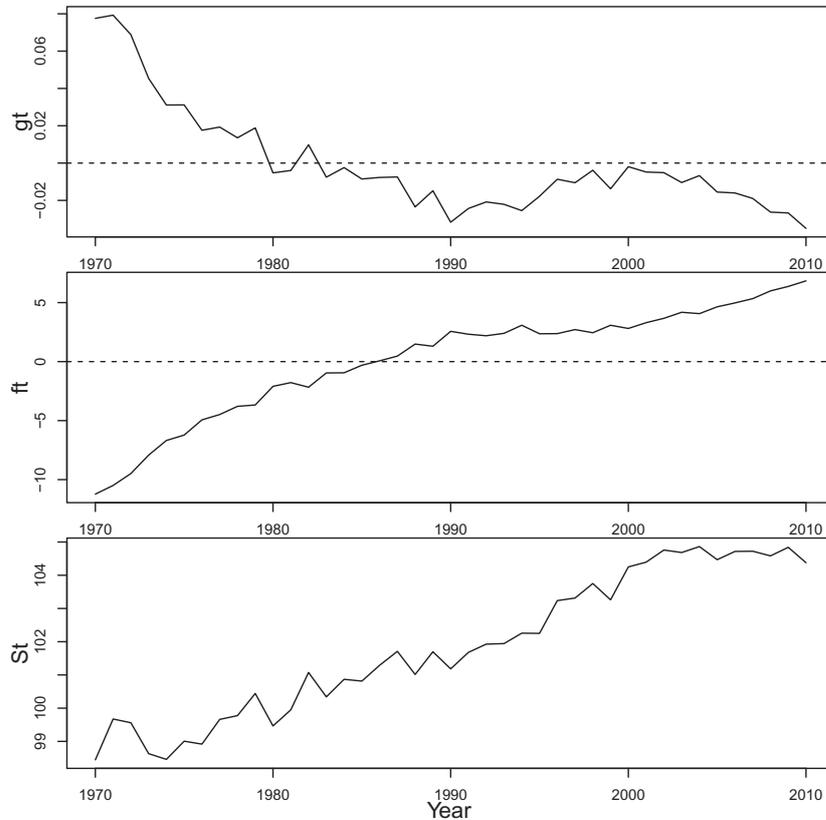


The Linear Difference Model (LD)

Lastly, we fit the LD model. Before observing the results for fit, we note estimated parameters of the LD model in Fig. 5.11.

Fig. 5.11: Estimated Parameters (LD)

Parameters of the LD Model (Female JPN)



The upper box shows the estimated g_t . We observe that g_t decreased between 1970 and 1980, therefore $g'_t < 0$ in this period. We can interpret this as there having been compression of mortality during this period. Between 1980 and 2000, g_t starts near zero, decreases slightly until 1990, and then returns to around 0 in 2000. We can see that g_t remains almost stationary, therefore, g'_t is around zero. This implies that there was little compression of the mortality curve and the shifting is strongly close to parallel in this period. After 2000, g_t decreases slightly.

The middle box shows the trends of f_t . As discussed in Chapter 4, the level of f_t is affected by g_t and this is not easy to interpret. Therefore, we see the parameter S_t in the bottom box instead. We can observe that S_t increases steadily in this period,

which reveals the strong shifting feature of the old mortality in Japan.

Fig. 5.12 is the actual inverse mortality rates and the estimated rates by the LD model, and Fig. 5.13 is the difference between the actual and the estimated rates. From these figures, we observe that the LD model fits well with the actual values. Even though the performance seems to be lower than that of the HL model, the LD model has an advantage that it needs fewer parameters than the HL model. Because the domain of the log mortality $[y_s, y_e] = [-7.15, -0.01]$ for Japanese females, we have 715 levels of the log mortality. Also we have 41 points of time from 1970 to 2010. Therefore, the number of the parameters is 1,471 for the HL, which is the sum of 715 (a_y), 715 (b_y) and 41(k_t), whereas there are 797 parameters for the LD, which is the sum of 715 (a_y), 41 (g_t) and 41(f_t).

Fig. 5.12: Inverse Mortality Rates (Actual and Model, LD)

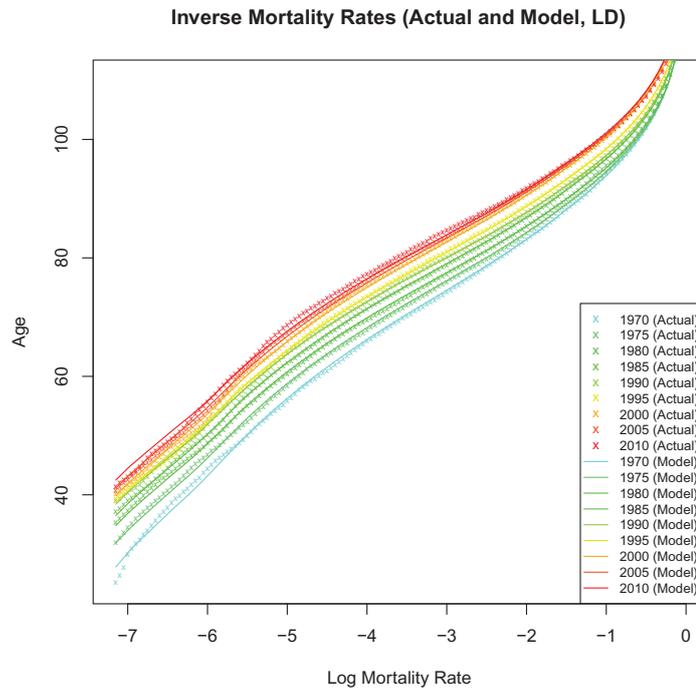
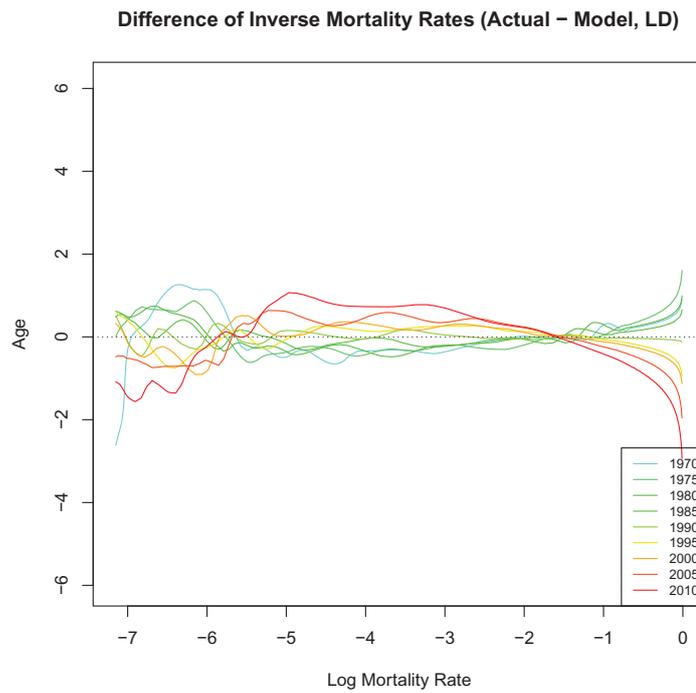


Fig. 5.13: Difference of Inverse Mortality Rates (Actual and Model, LD)



Figs. 5.14 and 5.15 show the relationship between $\tau_{y,t}$ and x (age) for the actual mortality, HS, HL, and LD models. $\tau_{y,t}$ for the LD model are shown by the red lines. We can see that the force of age increase for the HS model (pink lines) is almost horizontal, and that the distribution of $\tau_{y,t}$ for the HL model (green lines) is similar. According to these restrictions, the two models sometimes exhibit big differences from the actual rates, for example, in 1996. Compared to these models, we observe that the LD model performs somewhat better.

In these figures, we observe that the actual mortality rates indicated with the blue lines are well modeled by a linear function of age. They are completely linear in old age according to the proposition shown at the beginning of this subsection. Again, the force of age increase for the HS model (pink lines) is completely horizontal and does not show good fit to the actual rates. $\tau_{y,t}$ for the HL model (green lines) tend to decrease in old ages, whereas that for the actual rates exhibits both increase and decrease. This is caused by the restriction that the distribution is fixed. From these observations, we can conclude that the linear assumption for $\tau_{y,t}$ in the LD model works better.

Fig. 5.14: Comparison of the Force of Age Increase by Age (1975–1990)

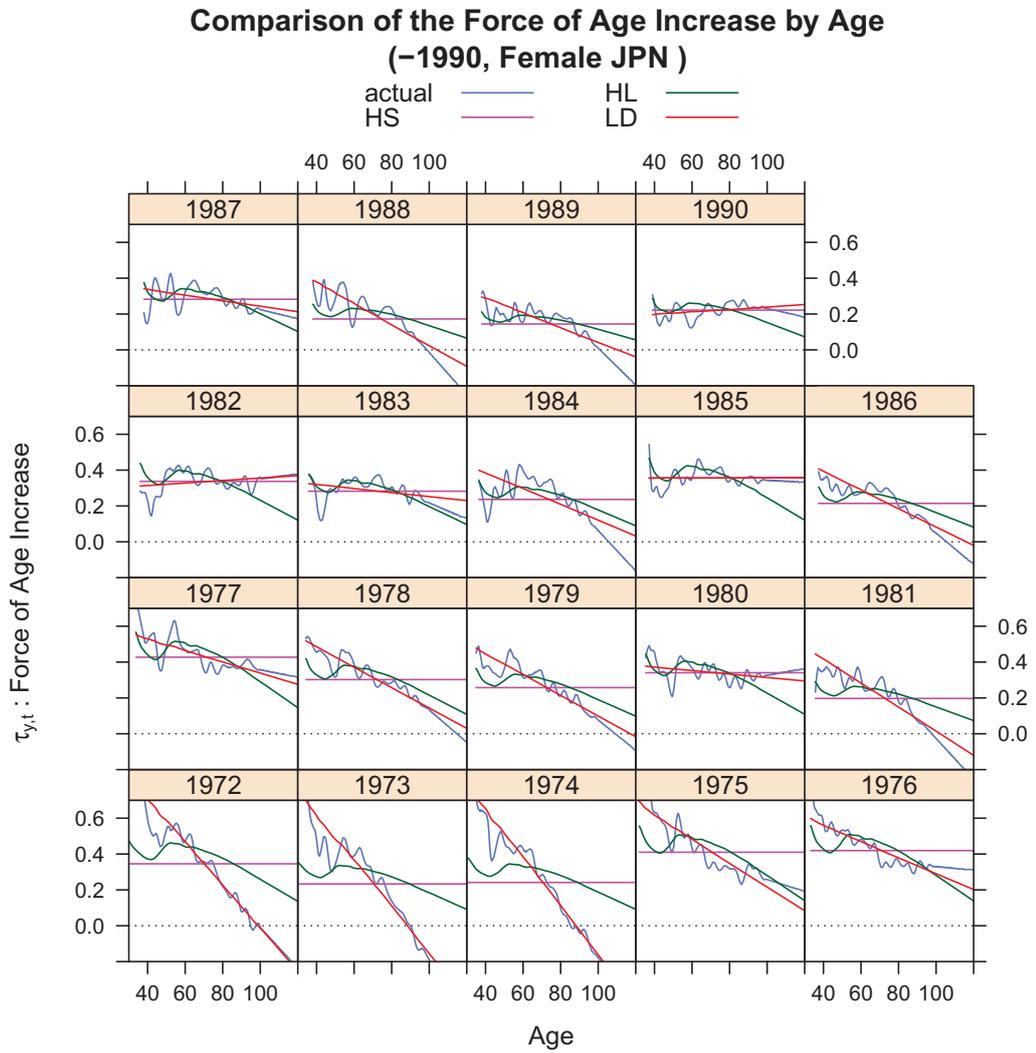
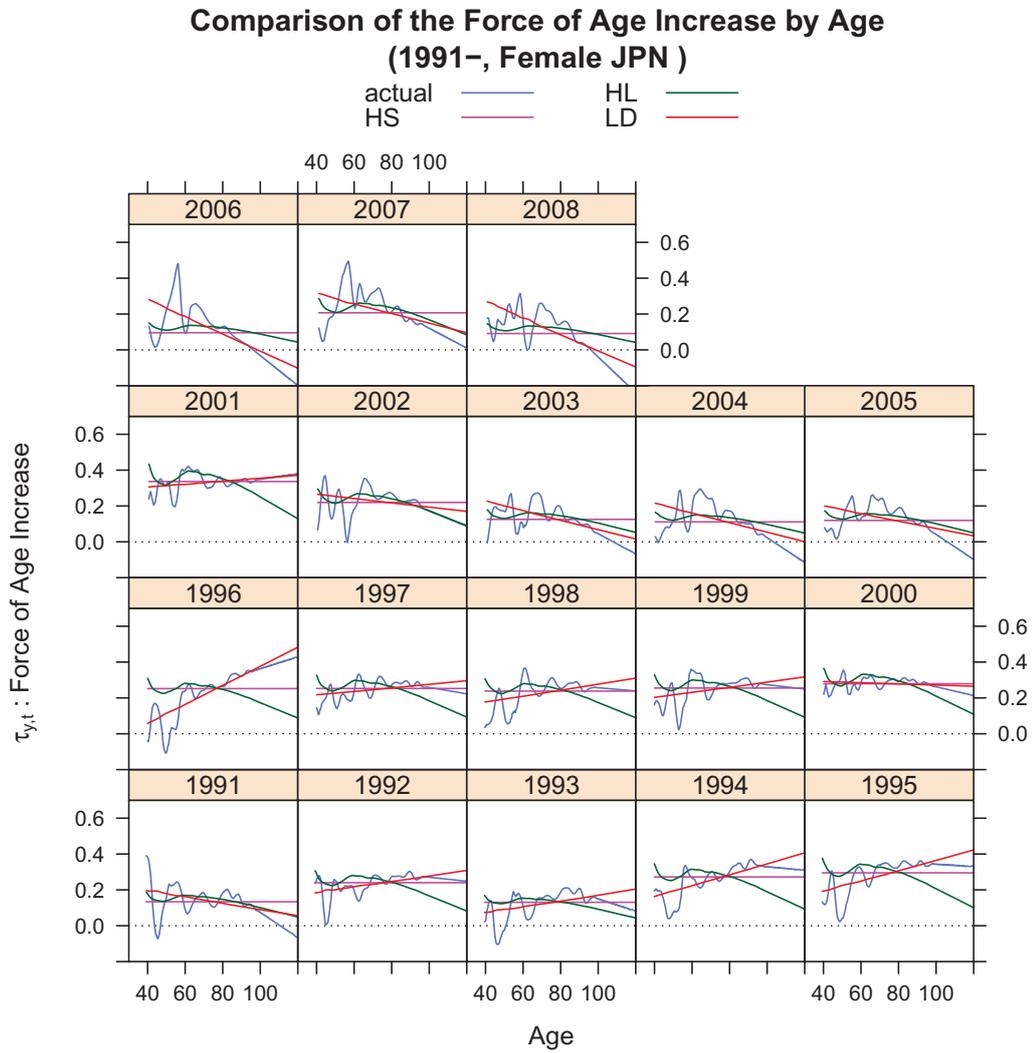


Fig. 5.15: Comparison of the Force of Age Increase by Age (1991–2006)



The definition of the LD model

$$\tau_{y,t} = f'_t + g'_t x$$

is equivalent to the following differential equation:

$$x'(t) = f'(t) + g'(t)x(t)$$

This is a linear ordinary differential equation of order 1 and has a closed-form solution as follows.

$$x(t) = \exp\left(\int_0^t g'(u)du\right) \left(x(0) + \int_0^t f'(u) \exp\left(-\int_0^u g'(v)dv\right) du\right) \quad (5.1)$$

We can verify this formula by differentiating $x(t)$.

$$\begin{aligned} x'(t) &= \frac{d}{dt} \left\{ \exp\left(\int_0^t g'(u)du\right) \left(x(0) + \int_0^t f'(u) \exp\left(-\int_0^u g'(v)dv\right) du\right) \right\} \\ &= \exp\left(\int_0^t g'(u)du\right) g'(t)x(0) \\ &\quad + \exp\left(\int_0^t g'(u)du\right) g'(t) \int_0^t f'(u) \exp\left(-\int_0^u g'(v)dv\right) du \\ &\quad + \exp\left(\int_0^t g'(u)du\right) f'(t) \exp\left(-\int_0^t g'(v)dv\right) \\ &= g'(t) \left\{ \exp\left(\int_0^t g'(u)du\right) \left(x(0) + \int_0^t f'(u) \exp\left(-\int_0^u g'(v)dv\right) du\right) \right\} \\ &\quad + f'(t) \\ &= f'(t) + g'(t)x(t) \end{aligned}$$

Here, we show some examples using this formula and observe the effects of the parameters, in particular $g(t)$. When $f'(t)$ and $g'(t)$ are constant, substituting $f'(t) = f'_0$ and $g'(t) = g'_0$ in Equation 5.1, we obtain

$$x(t) = \exp(g'_0 t)x(0) - \frac{f'_0}{g'_0} \exp(g'_0 t) + \frac{f'_0}{g'_0}$$

when $g'_0 \neq 0$.

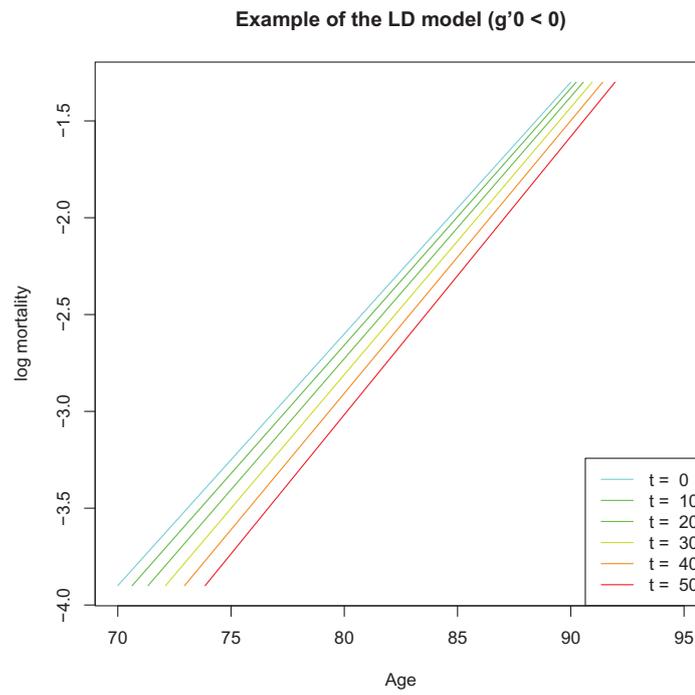
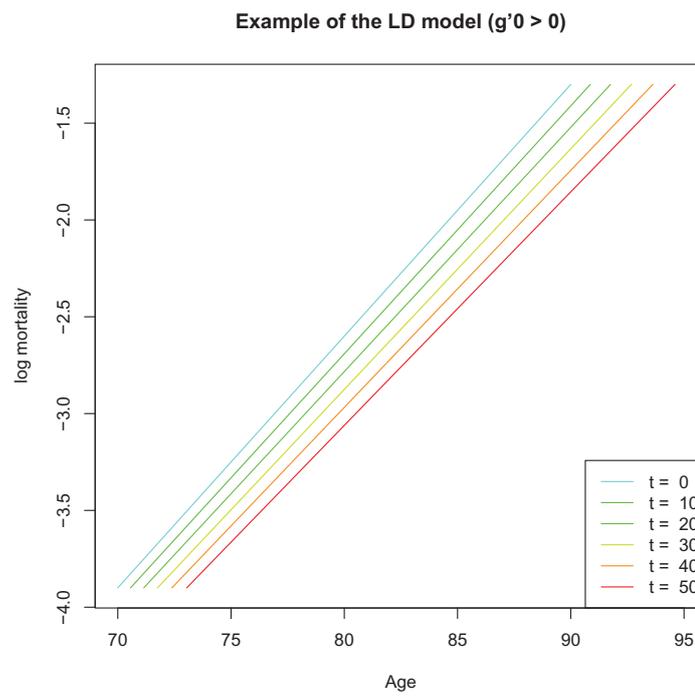
To show the numerical examples, we assume the following Gompertz model as a baseline mortality.

$$\lambda_{x,0} = 0.13x - 13$$

Fig. 5.16 shows the log mortality for $t = 0, 10, \dots, 50$ with $f'_0 = 0.2$ and $g'_0 = -0.002$. In this example, g'_0 is negative and $g(t) = -0.002t$ decreases linearly,

which indicates the mortality curve at time t becomes more compressed compared to the baseline mortality. Fig. 5.16 confirms this.

Fig. 5.17 shows the log mortality when we set $f'_0 = -0.05$ and $g'_0 = 0.0015$. Then, $g(t) = 0.0015t$ increases linearly and the log mortality becomes less compressed over time, as shown in Figs. 5.17.

Fig. 5.16: Example of the LD model ($g'_0 < 0$)Fig. 5.17: Example of the LD model ($g'_0 > 0$)

5.2 Comparison of the Models from a Statistical Viewpoint

In this section, we compare the LC and LD models from a statistical viewpoint to examine whether it is more plausible to understand Japan's recent old age mortality as *decline* or *shift*. Our approach is as follows.

1. The true mortality rates are assumed to be those estimated by models.
2. The number of deaths follows a binomial distribution $B(N_{x,t}, p_{x,t})$, where $N_{x,t}$ is the size of the population and $p_{x,t}$ is the death rate for age x and calendar year t .
3. $N_{x,t}$ is approximated by the closest integer to $E_{x,t}$ (exposure to risk).

Here, we take 0.01% as a critical value to construct the confidence intervals (CI), since $N_{x,t}$ would present too large a value for the Japanese female population.

In Figs. 5.18, 5.19, and 5.20, the blue lines show the difference of log actual mortality rates against the model rates by the LC. The pink and green lines show the difference of upper and lower CIs against the model values, respectively. We observe that actual rates between ages 80 to 100 years tend to move out of the CIs. Actual values are placed over the upper CIs between 1980 and 1995, whereas actual values are placed under the lower CIs before 1975 or after 2000.

Figs. 5.21, 5.22, and 5.23 show the same for the LD model, which differs from the LC model in that the actual values between ages 60 to 80 years tend to move out of the CIs. Actual values are placed over the upper CIs before 1985, whereas actual values are placed under the lower CIs after 1995.

Fig. 5.18: Difference of Log Actual Rate and CI against Model (LC)(1970–1984, critical value = 0.01%)

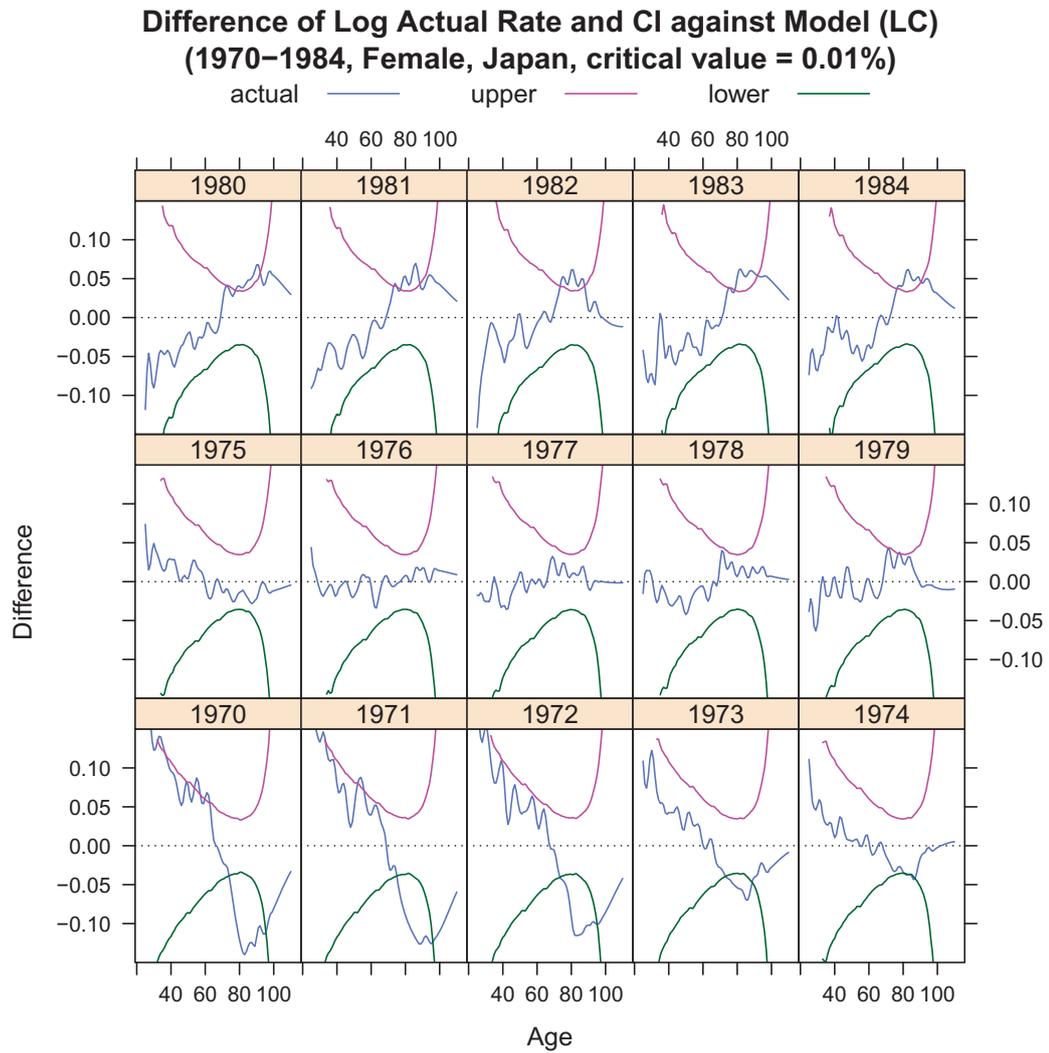


Fig. 5.19: Difference of Log Actual Rate and CI against Model (LC)(1985–1999, critical value = 0.01%)

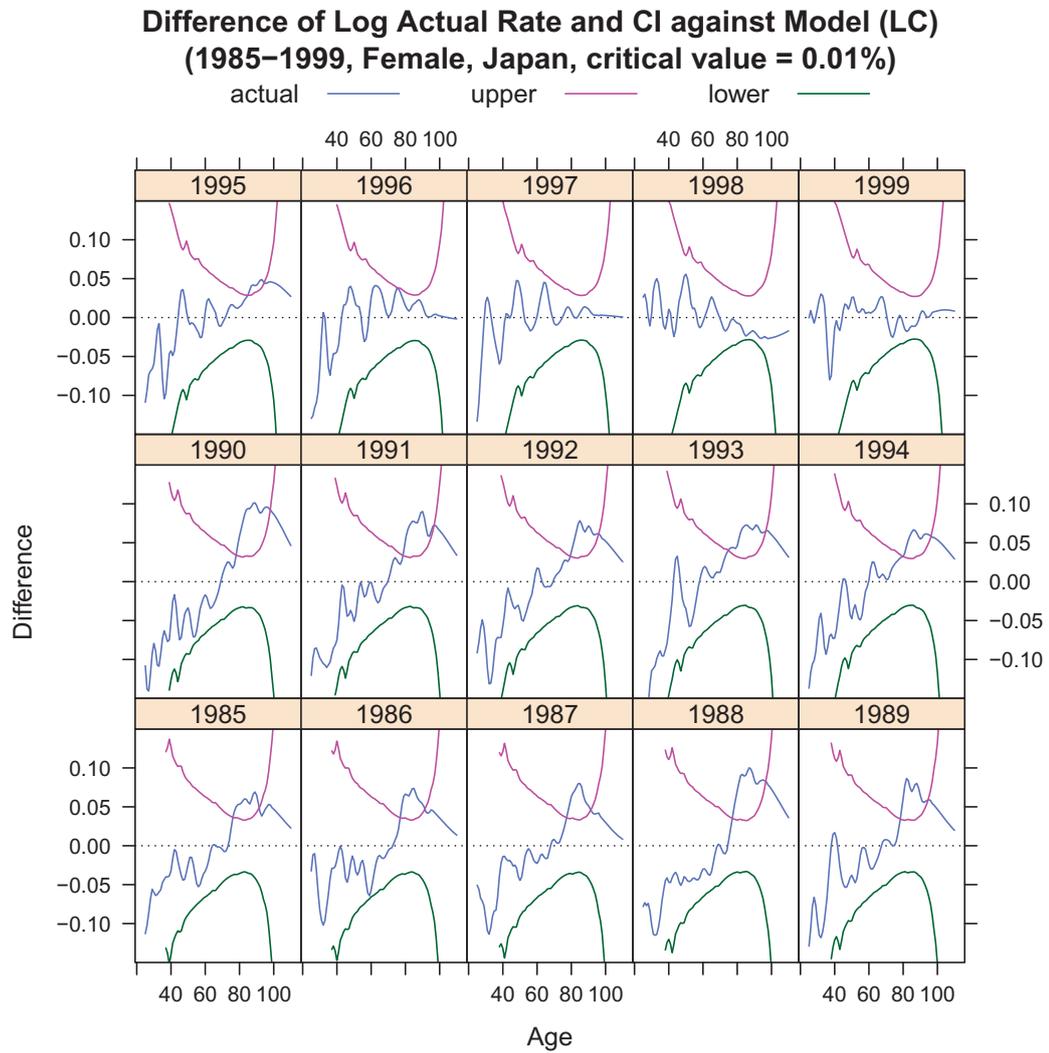


Fig. 5.20: Difference of Log Actual Rate and CI against Model (LC)(2000–2010, critical value = 0.01%)

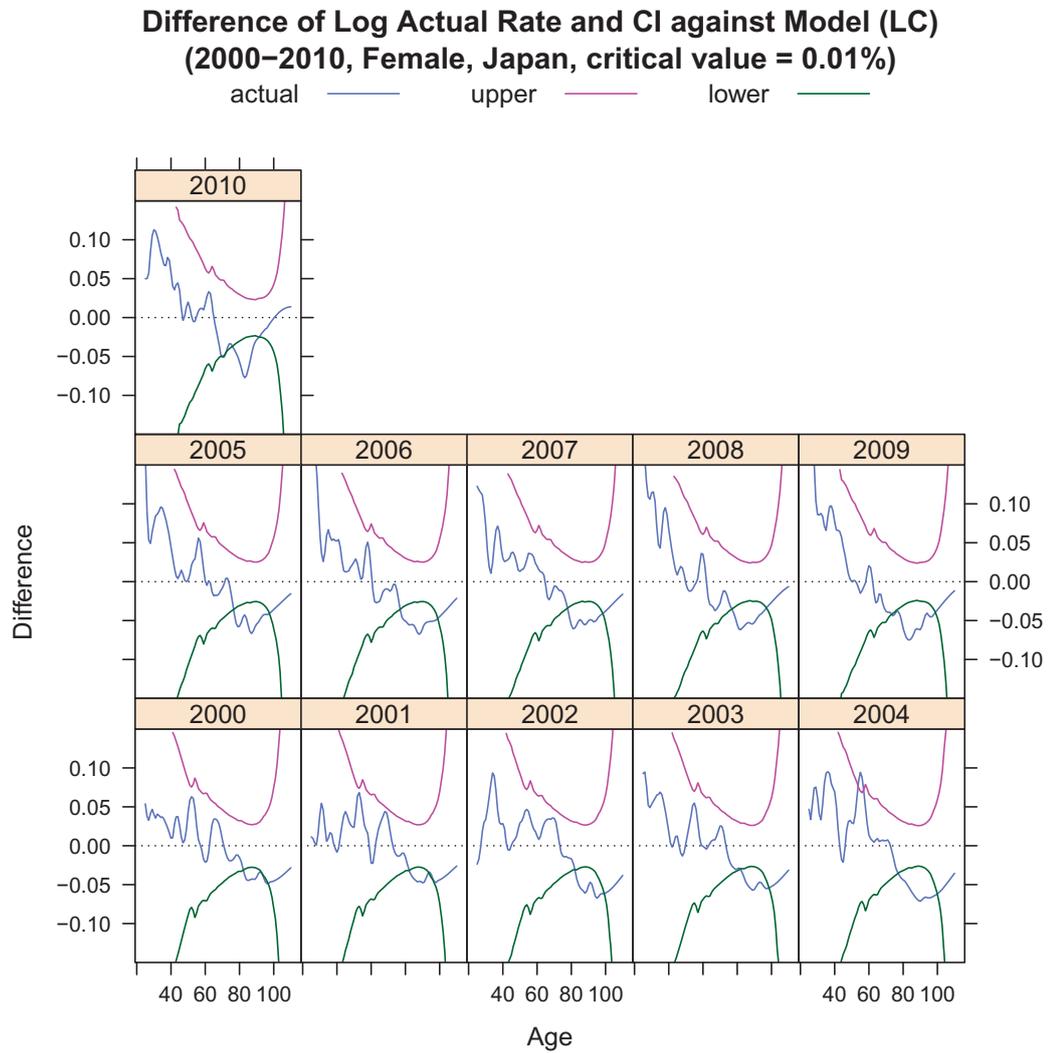


Fig. 5.21: Difference of Log Actual Rate and CI against Model (LD)(1970–1984, critical value = 0.01%)

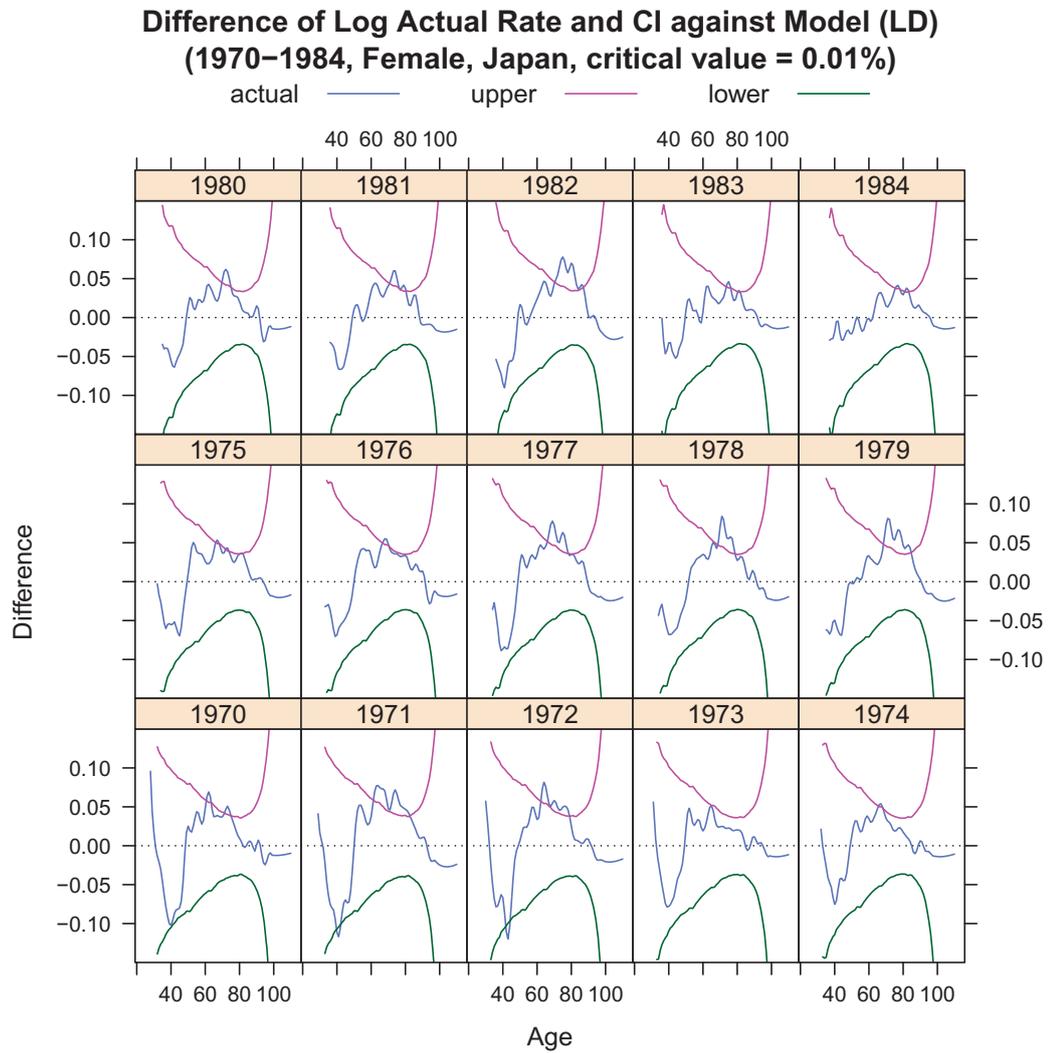


Fig. 5.22: Difference of Log Actual Rate and CI against Model (LD)(1985–1999, critical value = 0.01%)

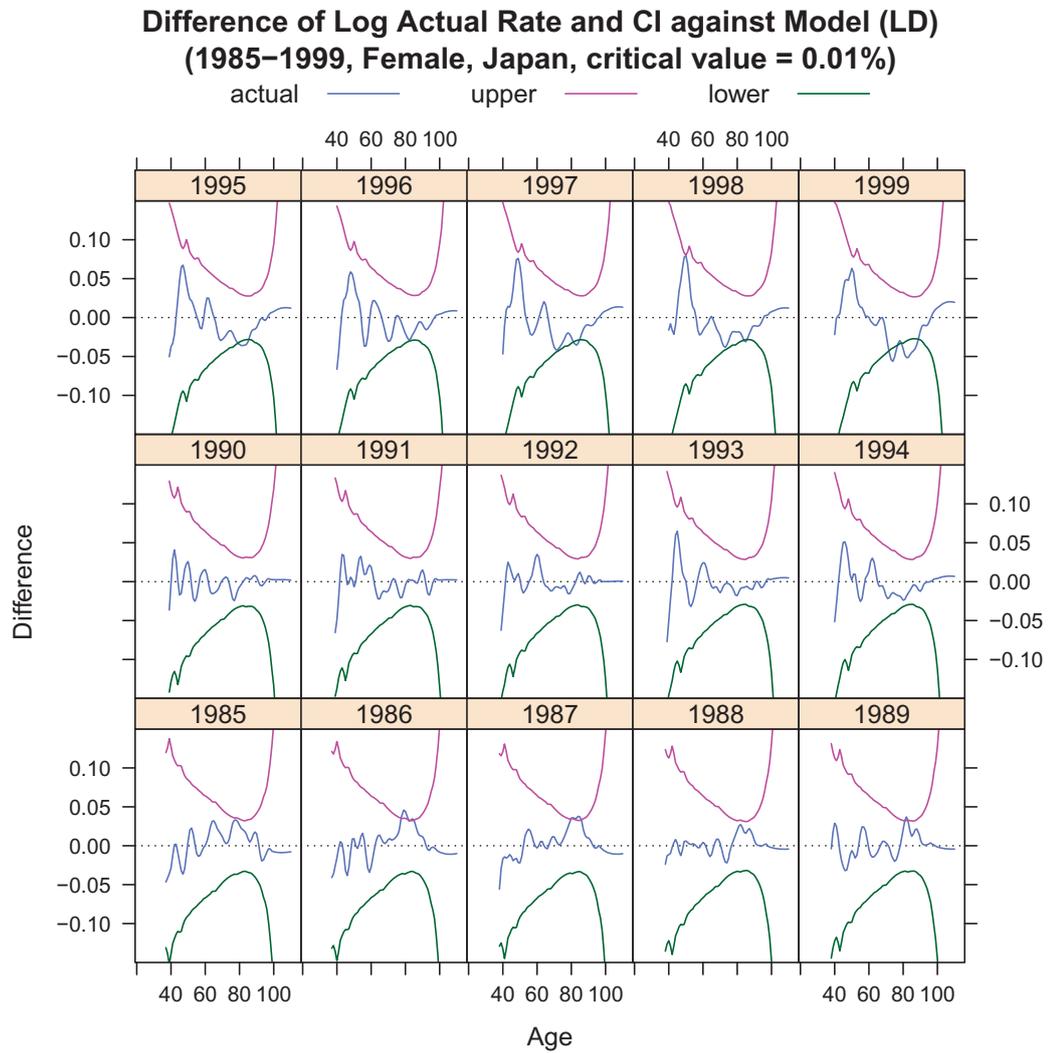


Fig. 5.23: Difference of Log Actual Rate and CI against Model (LD)(2000–2010, critical value = 0.01%)

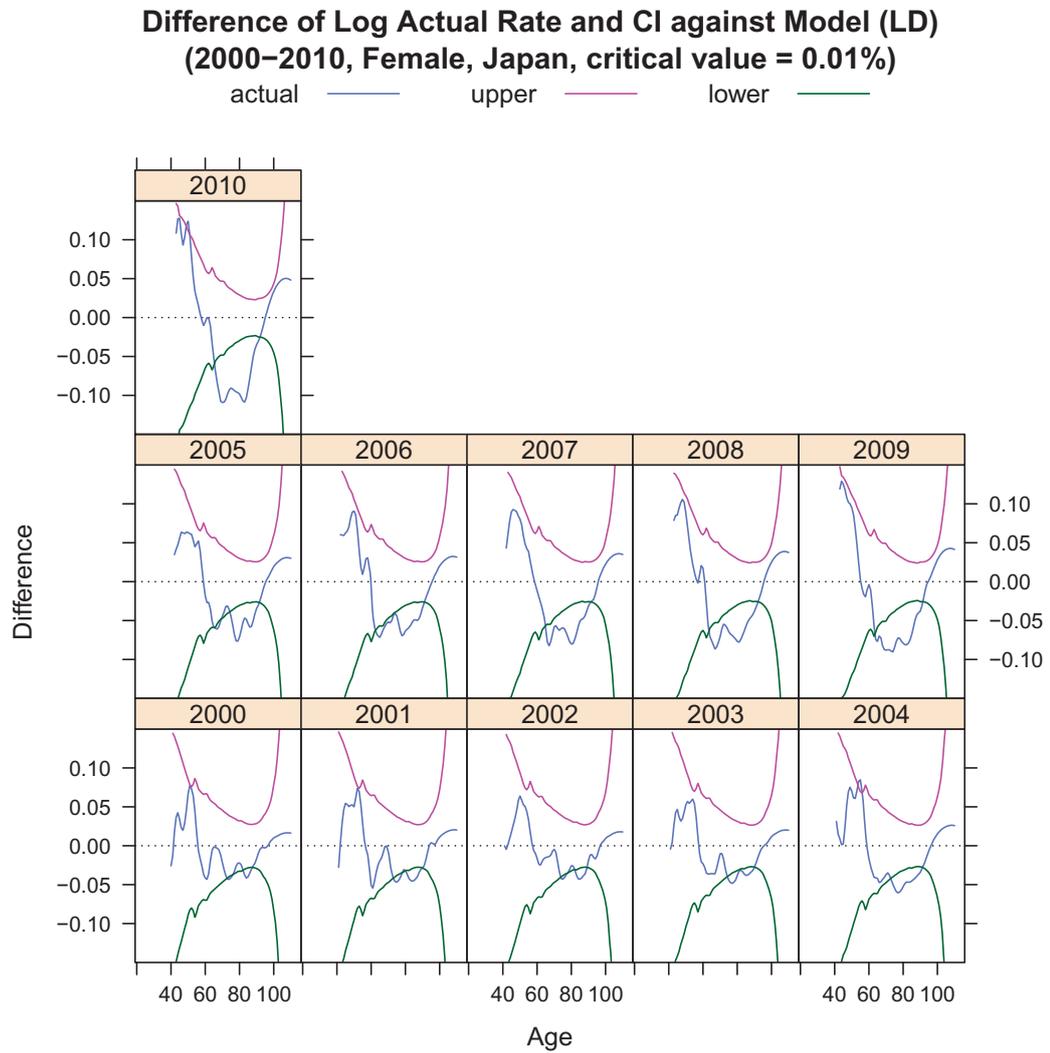
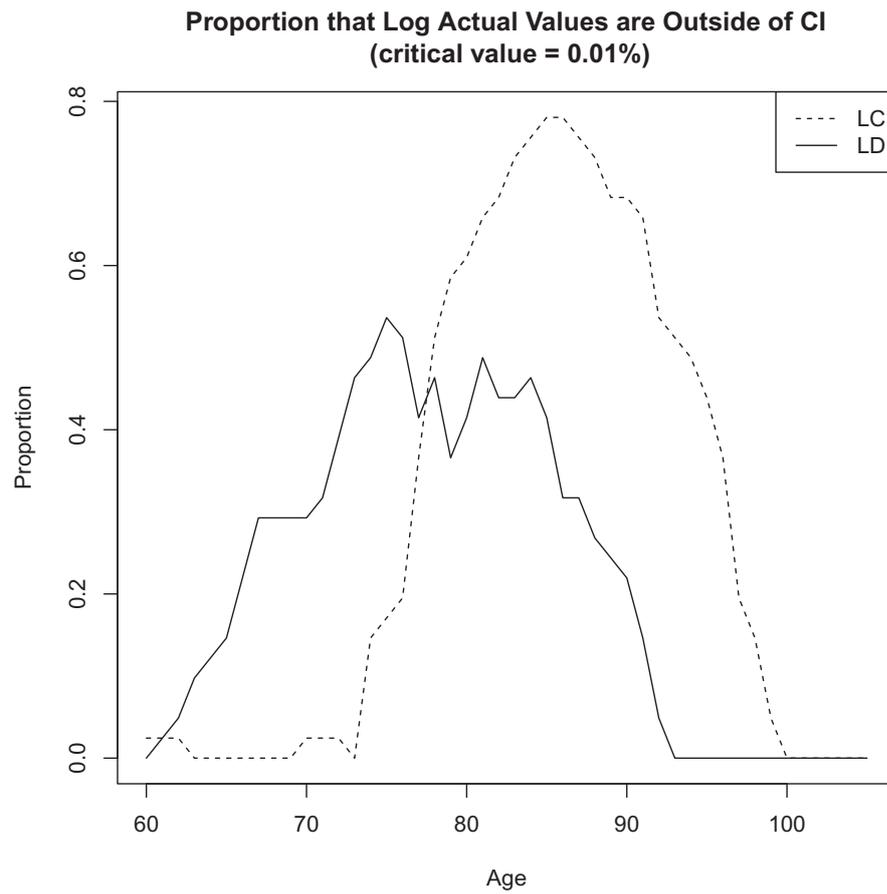


Fig. 5.24 shows the proportion of the log actual mortality rates that are outside of the CIs for each age in the LC and LD models. This indicates that, even though the proportions of LD are higher under about 75 years of age, the performance of LD is considered better than that of LC over 75 years of age. This result suggests that *shift* is more strongly supported as a factor behind the recent old age mortality improvement in Japan than *decline*.

Moreover, this result also acts as a guide for a better construction of a mortality projection model. The *shift*-type models are applicable only for adult mortality and juvenile mortality should be modeled by the *decline*-type models. The results show that old age mortality is better modeled by the LD, and middle age mortality by the LC. Therefore, it would work quite well if we could construct a blended model that has the LC property in youth and the LD property in older age. In Chapter 6, we propose a model for projection based on this observation by applying the idea of tangent vector fields on the log mortality surface.

Fig. 5.24: Proportion of Log Actual Values That Are Outside of CI (critical value = 0.01%)



5.3 Application to Analysis of the Trends of Modal Age at Death

Recently, the modal age at death has received more attention as an indicator of longevity (Horiuchi et al. 2013). Although many studies discuss the modal age at death, there are few articles that examine decomposition analyses of the change of the modal age in terms of the shifting and/or the compression of the mortality curve. In this section, we propose a new decomposition method for the modal age at death using the LD model, and provide decomposition analyses with the method.

5.3.1 Decomposition of the Change in Modal Age using the LD Model

First, we describe the methods for estimating M_t , the modal age at death. It is often difficult to estimate M_t from the raw d_x functions in the life tables because of the fluctuations. Therefore, smoothing methods and/or parametric modelings are usually used in estimation. Canudas-Romo (2008) uses the approximation by quadratic function originally proposed by Kannisto. Horiuchi et al. (2013) use a nonparametric smoothing method based on P-splines. Here, we use the minimum- R_3 moving averages with nine terms by Greville (1981) for smoothing the m_x functions that are used in the official life tables for Japan, and estimate M_t using quadratic approximations used in Canudas-Romo (2008).

In the LD model, we derive the following decomposition of the trends of M_t .

$$\frac{d}{dt}M_t = f'_t + g'_t \left(M_t - \frac{1}{\mu_{x,t} - \frac{\frac{\partial^2}{\partial x^2} \lambda_{x,t}}{\frac{\partial}{\partial x} \lambda_{x,t}}} \right) = S'_t + g'_t(M_t - S_t) - g'_t D_t$$

where $D_t = \frac{1}{\mu_{x,t} - \frac{\frac{\partial^2}{\partial x^2} \lambda_{x,t}}{\frac{\partial}{\partial x} \lambda_{x,t}}}$.

This formula is interpreted as follows. S'_t represents the amount of shifting, $g'_t(M_t - S_t)$ is the effect of compression at the modal age, and $-g'_t D_t$ represents the gap of the modal age at $t + dt$ and the age at t that the value of $\lambda_{x,t}$ of the modal age at t takes. Moreover, the formula $\frac{d}{dt}M_t = f'_t + g'_t(M_t - D_t)$ could be understood that the change of the modal age at death is equal to the force of age increase for the age $M_t - D_t$.

To derive this decomposition, we first notice a relationship that holds on M_t .

Proposition 2. When $x = M_t$, then

$$\frac{\partial}{\partial t} \mu_{x,t} = \frac{\partial}{\partial t} \frac{\partial}{\partial x} \lambda_{x,t}$$

Proof.

$$\begin{aligned} \frac{\partial^2}{\partial x^2} l_{x,t} &= -\frac{\partial}{\partial x} (\mu_{x,t} l_{x,t}) \\ &= -\frac{\partial \mu_{x,t}}{\partial x} l_{x,t} - \mu_{x,t} \frac{\partial l_{x,t}}{\partial x} \\ &= -l_{x,t} \left\{ \frac{\partial \mu_{x,t}}{\partial x} - \mu_{x,t}^2 \right\} \end{aligned}$$

If $x = M_t$, then $\frac{\partial^2}{\partial x^2} l_{x,t} = 0$. Therefore,

$$\begin{aligned} \frac{\partial \mu_{x,t}}{\partial x} &= \mu_{x,t}^2 \\ \Leftrightarrow \frac{\partial}{\partial x} \log \mu_{x,t} &= \mu_{x,t} \\ \Rightarrow \frac{\partial}{\partial t} \mu_{x,t} &= \frac{\partial}{\partial t} \frac{\partial}{\partial x} \lambda_{x,t} \end{aligned}$$

□

In the following discussion, we consider the expression of M_t as a linear combination of the tangent vectors on S whose directions, defined as either x or y , are fixed. Then, we use the abovementioned relationship to describe the location of M_t .

Before we provide the formula for M_t , we show the following relationships that hold in the LD model.

Proposition 3. When x is fixed,

$$\begin{aligned} \frac{\partial}{\partial t} \mu_{x,t} &= -\frac{\partial}{\partial x} \mu_{x,t} (f'_t + g'_t x) \\ \frac{\partial}{\partial t} \left(\frac{\partial}{\partial x} \lambda_{x,t} \right) &= -\frac{\partial^2}{\partial x^2} \lambda_{x,t} (f'_t + g'_t x) - \frac{\partial}{\partial x} \lambda_{x,t} g'_t \end{aligned}$$

Proof. The log mortality surface S is defined by the equation $\lambda_{x,t} - y = 0$, and the tangent space on (x_0, t_0, y_0) is $Y - y_0 = \frac{\partial \lambda_{x,t}}{\partial x} (X - x_0) + \frac{\partial \lambda_{x,t}}{\partial t} (T - t_0)$. $(\tau_{y,t}, 1, 0)$ is a tangent vector on S , therefore, we have

$$\begin{aligned} 0 &= \frac{\partial \lambda_{x,t}}{\partial x} \tau_{y,t} + \frac{\partial \lambda_{x,t}}{\partial t} \\ \Leftrightarrow \frac{1}{\mu_{x,t}} \frac{\partial \mu_{x,t}}{\partial t} &= -\frac{1}{\mu_{x,t}} \frac{\partial \mu_{x,t}}{\partial x} (f'_t + g'_t x) \end{aligned}$$

This proves the first equation.

Then, using this formula,

$$\begin{aligned}\frac{\partial}{\partial t} \left(\frac{\partial}{\partial x} \lambda_{x,t} \right) &= \frac{\partial}{\partial x} \left(\frac{\partial}{\partial t} \lambda_{x,t} \right) \\ &= -\frac{\partial}{\partial x} \left\{ \frac{\partial}{\partial x} \lambda_{x,t} (f'_t + g'_t x) \right\} \\ &= -\frac{\partial^2}{\partial x^2} \lambda_{x,t} (f'_t + g'_t x) - \frac{\partial}{\partial x} \lambda_{x,t} g'_t\end{aligned}$$

This completes the proof of the second equation. \square

Next we observe a similar relationship when y is fixed. Obviously, $\frac{\partial}{\partial t} \mu_{x,t} = 0$ when $y = \lambda_{x,t}$ is fixed. We consider the directional derivative along $\tau_{y,t}$ of the slope $\frac{\partial}{\partial x} \lambda_{x,t}$.

Proposition 4. *When y is fixed,*

$$\begin{aligned}\frac{\partial}{\partial t} \mu_{x,t} &= 0 \\ D_{\tau_{y,t}} \left(\frac{\partial}{\partial x} \lambda_{x,t} \right) &= -\frac{\partial}{\partial x} \lambda_{x,t} g'_t\end{aligned}$$

Proof. The slope at $t = t_0$ is expressed as $\frac{\Delta y}{\Delta x}$ when $\Delta x \rightarrow 0$. Then, at $t = t_0 + h$, for a small h , the slope is expressed as

$$\frac{\Delta y'}{\Delta x'} = \frac{\Delta y}{(1 + g'_{t_0})h\Delta x}$$

since $\Delta x' = \{x_0 + \Delta x + (f'_{t_0} + g'_{t_0}(x_0 + \Delta x))h\} - \{x_0 + (f'_{t_0} + g'_{t_0}(x_0))h\}$.

Then,

$$\begin{aligned}D_{\tau_{y,t}} \left(\frac{\partial}{\partial x} \lambda_{x,t} \right) &= \lim_{h \rightarrow 0} \frac{1}{h} \left\{ \frac{\Delta y}{(1 + g'_{t_0})h\Delta x} - \frac{\Delta y}{\Delta x} \right\} \\ &= \lim_{h \rightarrow 0} \frac{1}{h} \frac{\Delta y}{\Delta x} \frac{-hg'_{t_0}}{1 + hg'_{t_0}} \\ &= -\frac{\partial}{\partial x} \lambda(x_0, t_0) g'_{t_0}\end{aligned}$$

\square

We are now ready to derive the first formula. Assume that \mathbf{M} , the tangent vector on S along $x = M_t$, is expressed by a linear combination of \mathbf{A} and \mathbf{B} as

$\mathbf{M} = (1 - k)\mathbf{A} + k\mathbf{B}$, where \mathbf{A} and \mathbf{B} are the tangent vectors on S when x and y , respectively, are fixed. Then, $\frac{d}{dt}M_t = k(f'_t + g'_t x)$.

Using Proposition 3,

$$\begin{aligned} & - (1 - k) \left\{ \frac{\partial^2}{\partial x^2} \lambda_{x,t} (f'_t + g'_t x) - \frac{\partial}{\partial x} \lambda_{x,t} g'_t \right\} - k \frac{\partial}{\partial x} \lambda_{x,t} g'_t = - (1 - k) \frac{\partial}{\partial x} \mu_{x,t} (f'_t + g'_t x) \\ \Leftrightarrow k(f'_t + g'_t x) &= - \frac{g'_t}{\mu_{x,t} - \frac{\frac{\partial^2}{\partial x^2} \lambda_{x,t}}{\frac{\partial}{\partial x} \lambda_{x,t}}} + (f'_t + g'_t x) \end{aligned}$$

Substituting $x = M_t$,

$$\frac{d}{dt}M_t = f'_t + g'_t \left(M_t - \frac{1}{\mu_{x,t} - \frac{\frac{\partial^2}{\partial x^2} \lambda_{x,t}}{\frac{\partial}{\partial x} \lambda_{x,t}}} \right) = S'_t + g'_t(M_t - S_t) - g'_t D_t$$

where $D_t = \frac{1}{\mu_{x,t} - \frac{\frac{\partial^2}{\partial x^2} \lambda_{x,t}}{\frac{\partial}{\partial x} \lambda_{x,t}}}$.

This completes the proof of the decomposition.

5.3.2 Results of Decomposition

Fig. 5.25 shows the trends of M_t for the actual mortality and the LD model. We observe that both trends are similar, thus, we analyze the trends with the mortality rates by the LD model.

Fig. 5.26 shows the results of the decomposition of the change of M_t for the LD model every 10 years. We observe that the increase of M_t is caused mainly by the shift from 1980 to 2000, whereas compression plays a larger part before 1970 and after 2000.

Fig. 5.25: Trends of the Modal Age at Death (Actual and LD, Females, Japan)

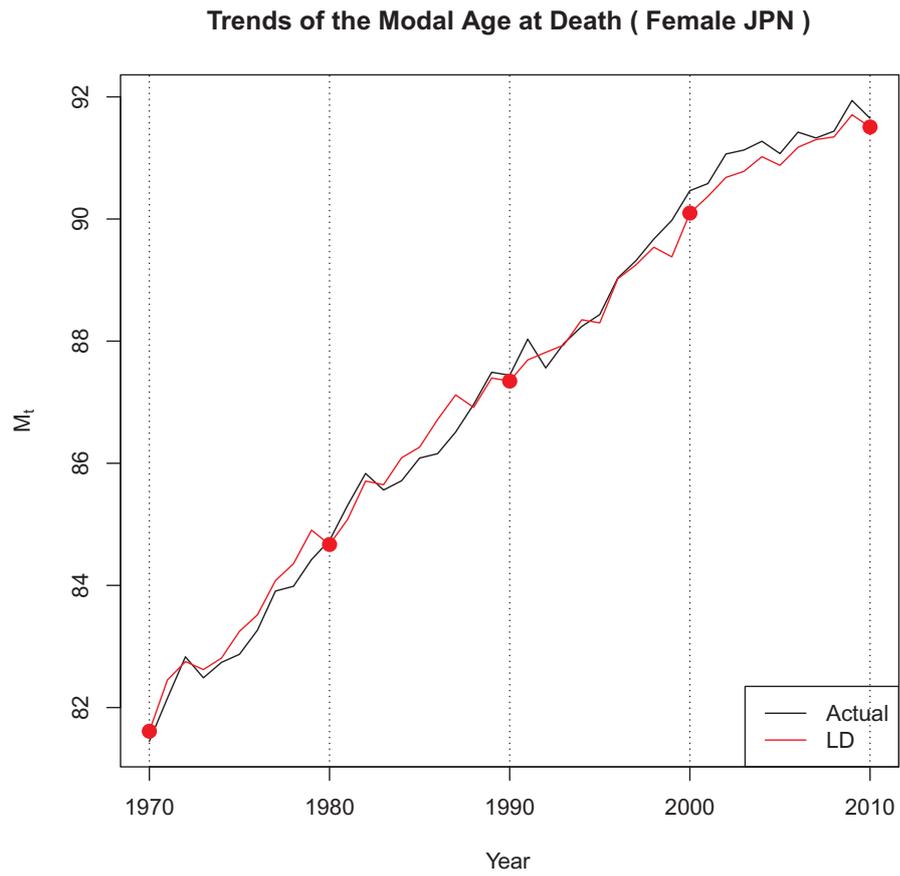
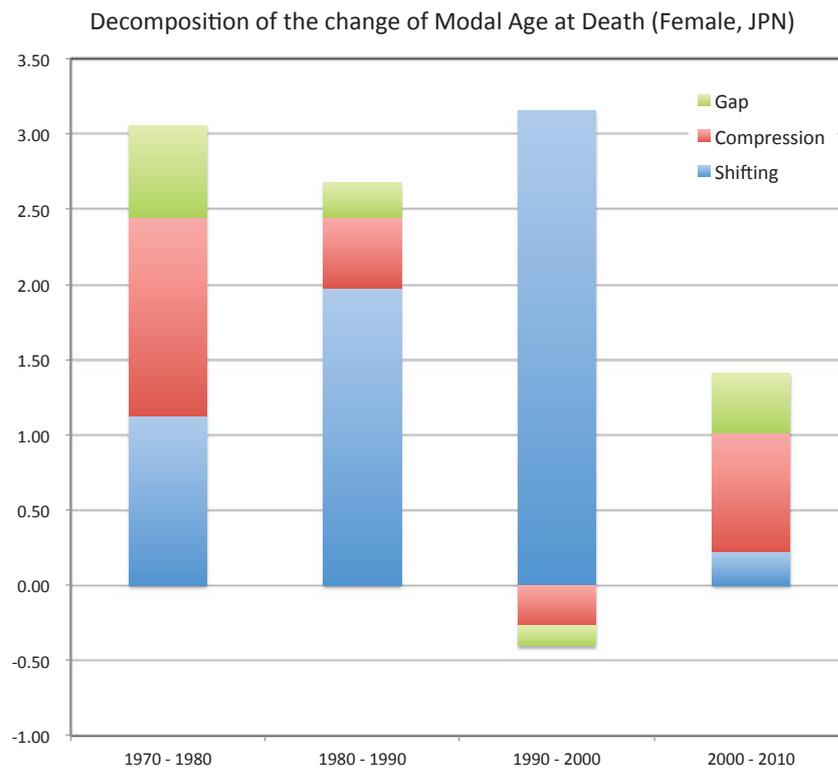


Fig. 5.26: Decomposition of the change of the Modal Age at Death (Females Japan)



Chapter 6

Tangent Vector Field Approach to Mortality Projection

6.1 Tangent Vector Field Approach to Mortality Projection

6.1.1 Building an Entire Age Model Using Tangent Vector Fields

In Chapter 4, we propose the LD model for adult mortality for Japan. However, we need an entire age model for mortality projection. In this section, we propose a method for solving this problem by applying tangent vector fields on the log mortality surface.

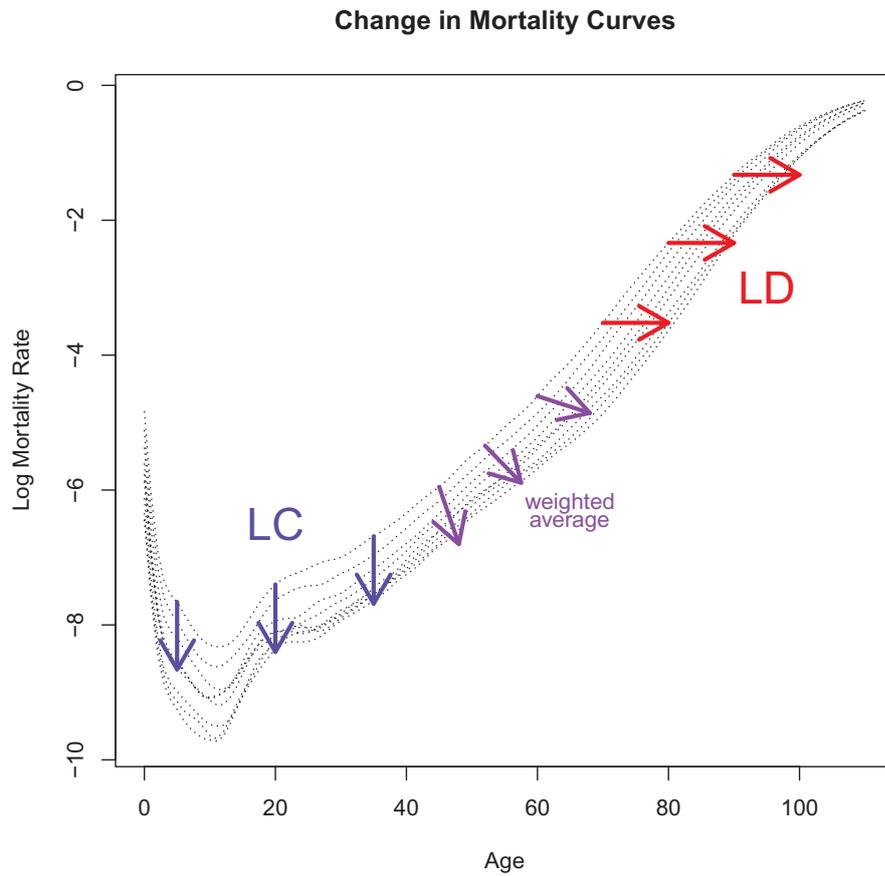
We begin with a stylized example of the change in mortality curves shown in Fig. 6.1. Now, we use the LD model for adult mortality, whose direction of mortality improvements is expressed by the age increases shown by the red arrows. On the other hand, the mortality improvements in juvenile mortality are well modeled by the *decline*-type models, such as the LC model, whose mortality improvements are shown by the blue arrows. The arrows show the directions in which the points on the log mortality curves are heading. Mathematically, these arrows are formulated using tangent vector fields on the log mortality surface.

In Chapter 4, we define the two differential functions, $\rho_{x,t}$ and $\tau_{y,t}$. Here, the following vectors that make use of these functions are tangent vectors on S , as shown in Fig. 6.2.

$$\begin{aligned}\rho(x_0, t_0, y_0) &= (0, 1, -\rho_{x_0, t_0}) \\ \tau(x_0, t_0, y_0) &= (\tau_{y_0, t_0}, 1, 0)\end{aligned}$$

Each tangent vector defines a tangent vector field on S . If $\rho_{x,t}$ and/or $\tau_{y,t}$ satisfy the conditions for the LC model and/or the LD model, respectively, we can say the upper vector field corresponds to the LC model, and the lower vector field to the LD model.

Fig. 6.1: Change in the Mortality Curves



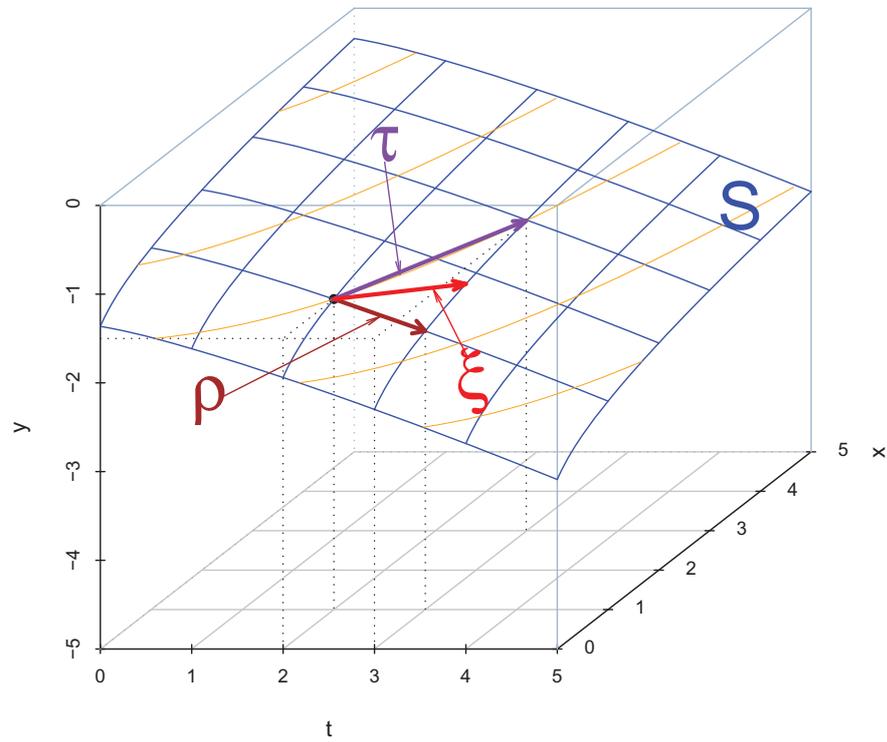
Then, from these two vector fields, we can construct one vector field as follows:

- For each point on the adult mortality area, we pick the vector from the LD model.
- For each point on the juvenile mortality area, we pick the vector from the LC model.
- For each point between the two areas, we take a weighted average of the two vectors.

More precisely, we use a weight function $w(x, t)$ that takes 0 in young age and 1 in old age, and define a new tangent vector field ξ :

$$\xi = (1 - w(x, t))\rho(x, t, y) + w(x, t)\tau(x, t, y)$$

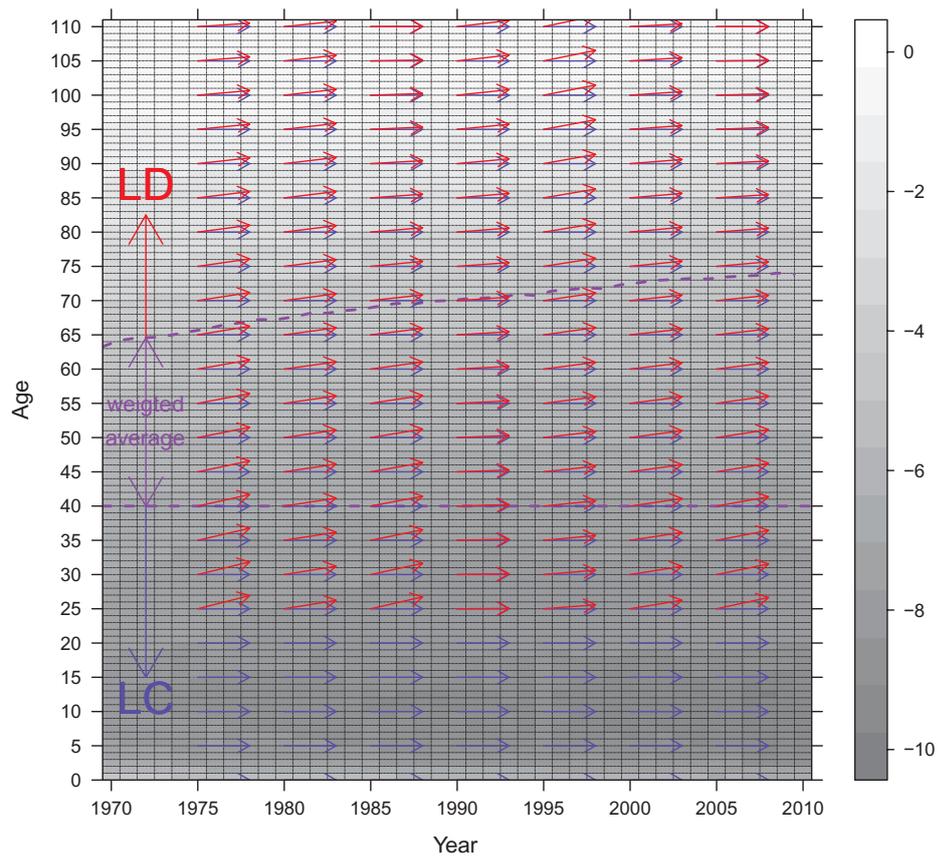
Fig. 6.2: Tangent Vectors on the Log Mortality Surface



This vector field induces a blended mortality model that has the LC property in youth and the LD property in older age. We call it the Tangent Vector Fields (TVF) model. Fig. 6.3 shows an example of the construction of the TVF model.

Fig. 6.3: Example of Construction of a Tangent Vector Field

Construction of a Tangent Vector Field



6.1.2 Application to Japanese Mortality Projection

We can project mortality rates with the TVF model by projecting the parameters for the LC and LD models. Actually, the TVF model is used in the 2012 official population projection in Japan (NIPSSR 2012). Moreover, Ishii and Lanzieri (2013) perform some experimental mortality projections for the EU countries using the TVF model and compare them with those by the LC model. Here, we simplify the method used in the Japanese official projection and perform a mortality projection as an example of application of the LD model¹.

Note that we use the base period from 2006 to 2010 and the modified LD method for mortality projection, as discussed in Section 4.4. Therefore, the value of the parameters is different from those in Chapter 5.

To project the mortality by the TVF model, we need the following parameters.

- a baseline log mortality a_x (estimated in the LC model)
- a set of Lee–Carter b_x sensitivity constants
- estimated and projected values of k_t for the LC model
- estimated and projected values of g_t for the LD model
- estimated and projected values of S_t for the LD model (and the converted values of f_t)
- a weight function $w(x, t)$

In our example, we project k_t for the LC model by a non-linear fitting, and g_t and S_t for the LD model by linear regressions. Therefore, we need some more parameters for projections of k_t , S_t , and g_t . However, any other methods that can produce their projected values can be chosen. In the following discussion, we show how the parameters are estimated and projected in our example.

First, we estimate the parameters a_x and b_x by the usual LC method using a singular value decomposition. Table 6.1 shows the estimated a_x and b_x for $x = 0, 5, \dots, 110$. The results for all ages x ($x \leq 110$) are shown in appendix (Table A.1).

Recall that when we work on the discrete form, age x actually represents the age interval $[x, x + 1)$. In the following discussion, we treat non-integer values for x that should be understood not as the exact ages but as the age intervals.

At the same time, we obtain the estimated parameters k_t from the LC procedure. We need future values of k_t for mortality projection. Usually, a linear extrapolation is used for projecting k_t in the LC method. However, the trajectory of k_t

¹Note that the results of the projection here are different from the 2012 official projection because, for example, we simplify the method and use other procedures for parameter estimation.

Table 6.1: Estimated a_x and b_x

x	a_x	b_x
0	-6.40128	0.04189
5	-9.18976	0.03974
10	-9.66274	0.03446
15	-9.08868	0.02449
20	-8.27315	0.01682
25	-8.11789	0.01803
30	-7.93639	0.01947
35	-7.63442	0.02000
40	-7.23336	0.01945
45	-6.81780	0.02022
50	-6.38395	0.02074
55	-6.02434	0.02220
60	-5.65698	0.02491
65	-5.29455	0.03069
70	-4.81176	0.03394
75	-4.23470	0.03584
80	-3.60353	0.03636
85	-2.90363	0.03244
90	-2.22321	0.02672
95	-1.61599	0.02077
100	-1.08484	0.01479
105	-0.66740	0.00919
110	-0.37838	0.00499

for Japan is different from a linear function because of the rapid improvement of mortality after World War II. In the 2012 official population projection, k_t is projected with extrapolation using non-linear fitting. The function is the average of the exponential and log function:

$$\frac{1}{2} [\{A_1 \exp(B_1(t - t_0 + 1)) + C_1\} + \{A_2 \log(B_2 + (t - t_0 + 1)) + C_2\}]$$

where $t_0 = 1970$.

Extrapolation using non-linear fitting exhibits a good performance in fitting k_t for Japan and has been used since the 2002 projection. Here, we use the same non-linear fitting method. The estimated function for non-linear fitting is

$$\frac{1}{2} [\{64.941595 \exp(-0.024326(t - 1970 + 1)) - 24.887423\} \\ + \{-40.352946 \log(22.568412 + (t - 1970 + 1)) + 166.417569\}]$$

The actual and projected k_t is shown in Fig. 6.4 and Tables 6.2 and 6.3. In Fig. 6.4, the black and red lines show the estimated and projected parameters respectively. We observe that fitting of k_t is quite good and confirm the efficiency of the non-linear fitting used in the official projections.

Fig. 6.4: Projection of k_t

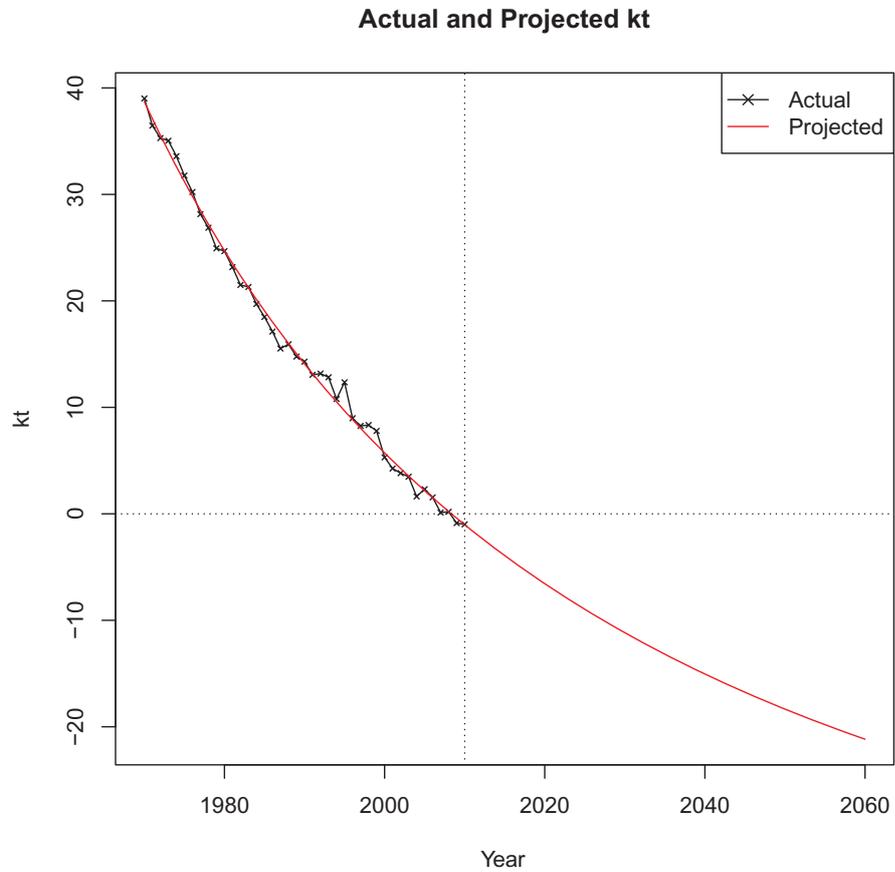


Table 6.2: Estimated k_t

Year	k_t
1970	39.00959
1971	36.46058
1972	35.28939
1973	35.04089
1974	33.58641
1975	31.76990
1976	30.22631
1977	28.14799
1978	26.86690
1979	24.93465
1980	24.66965
1981	23.17092
1982	21.49438
1983	21.30911
1984	19.70636
1985	18.49024
1986	17.11708
1987	15.52292
1988	15.92624
1989	14.76668
1990	14.28867
1991	13.05575
1992	13.17009
1993	12.82659
1994	10.78384
1995	12.35414
1996	8.97750
1997	8.25865
1998	8.33862
1999	7.78368
2000	5.30455
2001	4.25513
2002	3.81737
2003	3.48299
2004	1.63163
2005	2.31065
2006	1.54488
2007	0.12343
2008	0.19286
2009	-0.87076
2010	-0.99041

Table 6.3: Projected k_t

Year	k_t	Year	k_t
1970	38.69973	2015	-3.93249
1971	37.09971	2016	-4.47949
1972	35.55145	2017	-5.01619
1973	34.05193	2018	-5.54286
1974	32.59846	2019	-6.05976
1975	31.18856	2020	-6.56714
1976	29.82001	2021	-7.06522
1977	28.49076	2022	-7.55426
1978	27.19892	2023	-8.03447
1979	25.94278	2024	-8.50607
1980	24.72074	2025	-8.96927
1981	23.53132	2026	-9.42429
1982	22.37314	2027	-9.87132
1983	21.24492	2028	-10.31055
1984	20.14546	2029	-10.74217
1985	19.07364	2030	-11.16637
1986	18.02841	2031	-11.58332
1987	17.00879	2032	-11.99320
1988	16.01382	2033	-12.39617
1989	15.04264	2034	-12.79240
1990	14.09441	2035	-13.18205
1991	13.16834	2036	-13.56526
1992	12.26369	2037	-13.94219
1993	11.37973	2038	-14.31299
1994	10.51580	2039	-14.67780
1995	9.67125	2040	-15.03675
1996	8.84546	2041	-15.38998
1997	8.03785	2042	-15.73762
1998	7.24786	2043	-16.07979
1999	6.47494	2044	-16.41663
2000	5.71859	2045	-16.74824
2001	4.97831	2046	-17.07476
2002	4.25363	2047	-17.39628
2003	3.54409	2048	-17.71293
2004	2.84927	2049	-18.02482
2005	2.16874	2050	-18.33204
2006	1.50209	2051	-18.63470
2007	0.84894	2052	-18.93290
2008	0.20892	2053	-19.22674
2009	-0.41834	2054	-19.51631
2010	-1.03318	2055	-19.80170
2011	-1.63595	2056	-20.08302
2012	-2.22695	2057	-20.36034
2013	-2.80651	2058	-20.63374
2014	-3.37492	2059	-20.90333
		2060	-21.16916

Next, we discuss the parameters for the LD model. Applying the modified LD method described in Section 4.4, we can obtain the estimated parameters g_t (and f_t) for 1970 – 2010. We can also estimate the parameter S_t for age x for t that satisfies $m_{x,t} = 0.5$. Table 6.4 shows the estimated values for g_t and S_t .

For the projection of g_t and S_t , we first regress them on the projected k_t from 1970 to 2010, and then extrapolate them using the value of k_t after 2010. The formulae are:

$$g_t = 0.0035978 + 0.0031131k_t$$

$$S_t = 104.59967 - 0.17440k_t$$

The projected values for g_t and S_t are shown in Table 6.5 and Fig. 6.5. Even though the fitness of S_t and g_t is less effective than that of k_t , we see that they succeed in capturing the long-term trends and the projected trajectories are considered to be plausible.

Fig. 6.5: Projection of g_t and S_t

Projections of g_t and S_t (Female JPN)

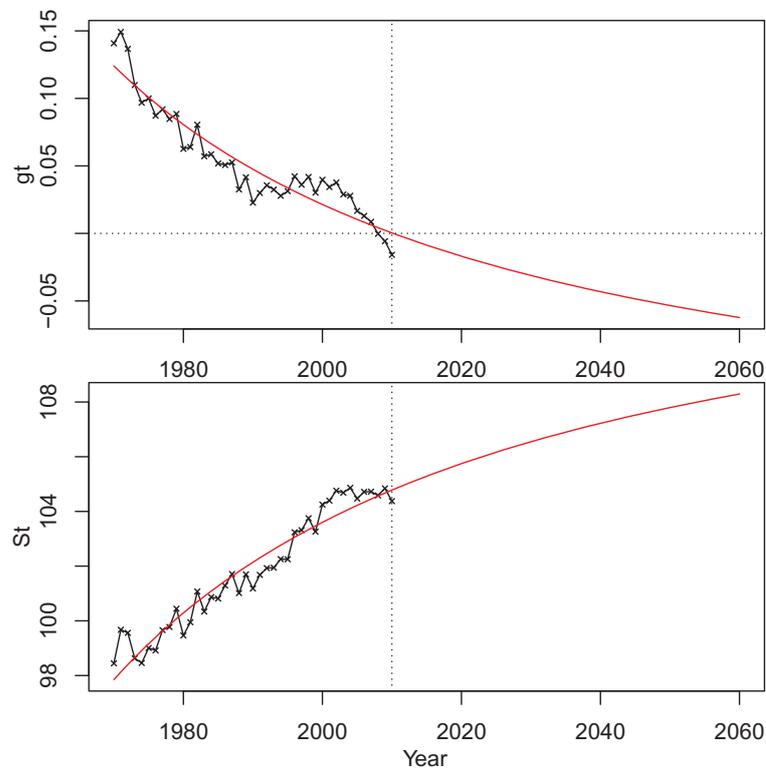


Table 6.4: Estimated g_t and S_t

Year	g_t	S_t
1970	0.14081	98.44532
1971	0.14922	99.67369
1972	0.13666	99.56049
1973	0.10989	98.62987
1974	0.09683	98.45862
1975	0.09999	99.00625
1976	0.08716	98.91745
1977	0.09192	99.66118
1978	0.08472	99.77548
1979	0.08848	100.44246
1980	0.06264	99.46466
1981	0.06397	99.94836
1982	0.08051	101.07333
1983	0.05710	100.33948
1984	0.05862	100.86782
1985	0.05179	100.81415
1986	0.05066	101.29046
1987	0.05258	101.71017
1988	0.03255	101.01421
1989	0.04154	101.69801
1990	0.02278	101.18247
1991	0.02994	101.67975
1992	0.03562	101.92643
1993	0.03249	101.94270
1994	0.02784	102.25687
1995	0.03130	102.25062
1996	0.04224	103.23662
1997	0.03609	103.31385
1998	0.04180	103.75121
1999	0.03021	103.25987
2000	0.03972	104.25056
2001	0.03434	104.39306
2002	0.03771	104.75828
2003	0.02886	104.68350
2004	0.02787	104.86259
2005	0.01668	104.46800
2006	0.01296	104.71687
2007	0.00860	104.72304
2008	-0.00026	104.58248
2009	-0.00578	104.84067
2010	-0.01585	104.37655

Table 6.5: Projected g_t and S_t

Year	g_t	S_t	Year	g_t	S_t
1970	0.12407	97.85043	2015	-0.00864	105.28550
1971	0.11909	98.12948	2016	-0.01035	105.38090
1972	0.11427	98.39949	2017	-0.01202	105.47450
1973	0.10960	98.66101	2018	-0.01366	105.56635
1974	0.10508	98.91450	2019	-0.01527	105.65650
1975	0.10069	99.16038	2020	-0.01685	105.74498
1976	0.09643	99.39906	2021	-0.01840	105.83185
1977	0.09229	99.63088	2022	-0.01992	105.91714
1978	0.08827	99.85618	2023	-0.02141	106.00088
1979	0.08436	100.07525	2024	-0.02288	106.08313
1980	0.08056	100.28837	2025	-0.02432	106.16391
1981	0.07685	100.49581	2026	-0.02574	106.24327
1982	0.07325	100.69779	2027	-0.02713	106.32123
1983	0.06973	100.89456	2028	-0.02850	106.39783
1984	0.06631	101.08630	2029	-0.02984	106.47311
1985	0.06298	101.27323	2030	-0.03116	106.54709
1986	0.05972	101.45551	2031	-0.03246	106.61981
1987	0.05655	101.63334	2032	-0.03374	106.69129
1988	0.05345	101.80686	2033	-0.03499	106.76157
1989	0.05043	101.97623	2034	-0.03623	106.83067
1990	0.04747	102.14160	2035	-0.03744	106.89862
1991	0.04459	102.30311	2036	-0.03863	106.96546
1992	0.04178	102.46088	2037	-0.03981	107.03119
1993	0.03902	102.61504	2038	-0.04096	107.09586
1994	0.03633	102.76571	2039	-0.04210	107.15948
1995	0.03371	102.91300	2040	-0.04321	107.22208
1996	0.03113	103.05702	2041	-0.04431	107.28369
1997	0.02862	103.19787	2042	-0.04539	107.34431
1998	0.02616	103.33564	2043	-0.04646	107.40399
1999	0.02375	103.47044	2044	-0.04751	107.46273
2000	0.02140	103.60235	2045	-0.04854	107.52057
2001	0.01910	103.73145	2046	-0.04956	107.57751
2002	0.01684	103.85784	2047	-0.05056	107.63359
2003	0.01463	103.98158	2048	-0.05154	107.68881
2004	0.01247	104.10276	2049	-0.05251	107.74320
2005	0.01035	104.22144	2050	-0.05347	107.79678
2006	0.00827	104.33771	2051	-0.05441	107.84957
2007	0.00624	104.45162	2052	-0.05534	107.90157
2008	0.00425	104.56324	2053	-0.05626	107.95282
2009	0.00230	104.67263	2054	-0.05716	108.00332
2010	0.00038	104.77986	2055	-0.05805	108.05309
2011	-0.00150	104.88498	2056	-0.05892	108.10215
2012	-0.00333	104.98805	2057	-0.05979	108.15052
2013	-0.00514	105.08913	2058	-0.06064	108.19820
2014	-0.00691	105.18826	2059	-0.06148	108.24522
			2060	-0.06230	108.29158

Next, we calculate f_t . For the computation of f_t , we use the following recursive formula.

$$f_{t_2} \approx f_{t_1} + (S_{t_2} - S_{t_1}) - (g_{t_2} - g_{t_1}) \frac{S_{t_2} + S_{t_1}}{2}$$

To project f_t , we begin with the baseline mortality. At the baseline, we have $f_t = g_t = 0$. The value of S_t is 104.64792, which is obtained by a linear interpolation from the a_x shown in Table A.1. Then, the parameter f_t at $t = 2009$ that is considered close to the baseline is

$$\begin{aligned} f_{2009} &= 0 + (104.67263 - 104.64792) - (0.0022955 - 0) \frac{104.67263 + 104.64792}{2} \\ &= -0.21554 \end{aligned}$$

Then, we can compute the values for the next (or previous) years step by step. For example, f_{2010} is

$$\begin{aligned} f_{2010} &= -0.21554 + (104.77986 - 104.67263) \\ &\quad - (0.00038143 - 0.0022955) \frac{104.77986 + 104.67263}{2} = 0.09214 \end{aligned}$$

Table 6.6 shows the projected values of f_t .

Table 6.6: Projected f_t

Year	f_t	Year	f_t
1970	-19.36897	2015	1.54578
1971	-18.60184	2016	1.82055
1972	-17.85820	2017	2.09030
1973	-17.13674	2018	2.35516
1974	-16.43626	2019	2.61525
1975	-15.75569	2020	2.87069
1976	-15.09404	2021	3.12159
1977	-14.45042	2022	3.36806
1978	-13.82399	2023	3.61021
1979	-13.21401	2024	3.84814
1980	-12.61976	2025	4.08195
1981	-12.04060	2026	4.31175
1982	-11.47591	2027	4.53761
1983	-10.92513	2028	4.75965
1984	-10.38772	2029	4.97794
1985	-9.86320	2030	5.19257
1986	-9.35108	2031	5.40363
1987	-8.85094	2032	5.61121
1988	-8.36235	2033	5.81537
1989	-7.88492	2034	6.01621
1990	-7.41828	2035	6.21379
1991	-6.96207	2036	6.40819
1992	-6.51597	2037	6.59948
1993	-6.07964	2038	6.78773
1994	-5.65278	2039	6.97301
1995	-5.23511	2040	7.15539
1996	-4.82635	2041	7.33494
1997	-4.42622	2042	7.51170
1998	-4.03448	2043	7.68575
1999	-3.65088	2044	7.85715
2000	-3.27519	2045	8.02596
2001	-2.90718	2046	8.19222
2002	-2.54663	2047	8.35600
2003	-2.19335	2048	8.51735
2004	-1.84712	2049	8.67633
2005	-1.50777	2050	8.83298
2006	-1.17509	2051	8.98735
2007	-0.84891	2052	9.13950
2008	-0.52907	2053	9.28947
2009	-0.21554	2054	9.43731
2010	0.09214	2055	9.58306
2011	0.39398	2056	9.72677
2012	0.69012	2057	9.86849
2013	0.98070	2058	10.00824
2014	1.26588	2059	10.14608
		2060	10.28204

Next, we consider the weight function $w(x, t)$. This function is used to carry out weighted averages of the vectors that correspond to the LC and LD models, and then, $w(x, t)$ should take 0 in young age and 1 in old age. In young age, it is preferable that the upper bound of the area in which $w(x, t)$ takes 0 is constant with respect to age because the LC model is used in this area. However, it is more appropriate that the lower bound of the area in which $w(x, t)$ takes 1 in old age corresponds to the level of the mortality rate where the LD model is used. Therefore, we move the lower bound $x_1(t)$ corresponding to the log mortality rate for the age of 70 years at year 2008. Thus, $x_1(t)$ is set as $x_1(2008) = 70$ at $t = 2008$ and is varied satisfying the condition that $\lambda_{x_1(t),t} = \lambda_{70,2008}$.

For the projection of $x_1(t)$, we use the following equation that holds in the discrete form.

Proposition 5. *If $x(t_1)$ and $x(t_2)$ are the points on the log mortality surface of the LD model and are on the same flow that has the same values of log mortality, then*

$$x(t_2) = \frac{1 - g_{t_1}}{1 - g_{t_2}} x(t_1) + \frac{f_{t_2} - f_{t_1}}{1 - g_{t_2}}$$

Proof. In the LD model, $v_{y,t} = a_y + g_t x + f_t$. Therefore, $x(t_1)$ and $x(t_2)$ are expressed as follows using $x(t_0)$, which is the value of x for the baseline mortality curve on the flow:

$$\begin{aligned} x(t_1) &= x(t_0) + g_{t_1} x(t_1) + f_{t_1} \\ x(t_2) &= x(t_0) + g_{t_2} x(t_2) + f_{t_2} \end{aligned}$$

Subtracting the upper equation from the lower,

$$(1 - g_{t_2})x(t_2) = (1 - g_{t_1})x(t_1) + f_{t_2} - f_{t_1}$$

This proves the proposition. □

Then, substituting $t_1 = 2008$, $t_1 = t$, and $x(t_1) = 70$,

$$x_1(t) = \frac{1 - g_{2008}}{1 - g_t} \cdot 70 + \frac{f_t - f_{2008}}{1 - g_t}$$

Table 6.7 shows the projected values of $x_1(t)$.

Using the projected $x_1(t)$, we define the weight function $w(x, t)$ as the following piecewise linear function.

$$w(x, t) = \begin{cases} 0 & (x < x_0(t) = 40) \\ \frac{x - x_0(t)}{x_1(t) - x_0(t)} & (x_0(t) \leq x < x_1(t)) \\ 1 & (x \geq x_1(t)) \end{cases}$$

Fig. 6.6 shows the projected $w(x, t)$ for $t = 1970, 1980, \dots, 2060$ and 2008.

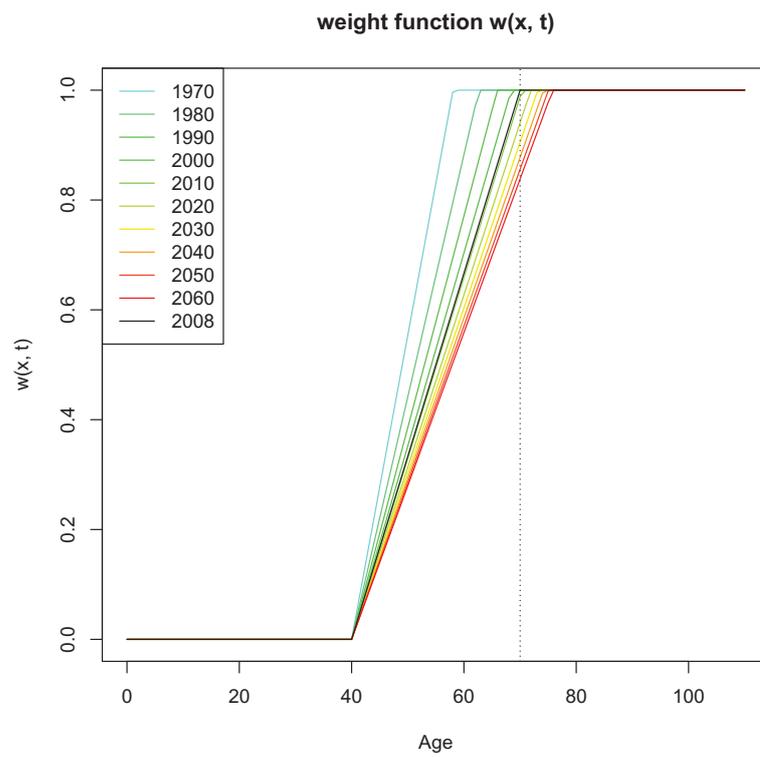
Fig. 6.6: Projection of $w(x, t)$ 

Table 6.7: Projected $x_1(t)$

Year	$x_1(t)$	Year	$x_1(t)$
1970	58.06732	2015	71.16233
1971	58.60982	2016	71.31434
1972	59.13046	2017	71.46315
1973	59.63073	2018	71.60885
1974	60.11196	2019	71.75153
1975	60.57536	2020	71.89129
1976	61.02200	2021	72.02820
1977	61.45287	2022	72.16234
1978	61.86888	2023	72.29380
1979	62.27083	2024	72.42264
1980	62.65949	2025	72.54895
1981	63.03554	2026	72.67279
1982	63.39963	2027	72.79423
1983	63.75233	2028	72.91333
1984	64.09420	2029	73.03016
1985	64.42575	2030	73.14478
1986	64.74744	2031	73.25725
1987	65.05972	2032	73.36763
1988	65.36301	2033	73.47596
1989	65.65768	2034	73.58231
1990	65.94410	2035	73.68673
1991	66.22261	2036	73.78926
1992	66.49354	2037	73.88996
1993	66.75718	2038	73.98887
1994	67.01381	2039	74.08603
1995	67.26372	2040	74.18150
1996	67.50714	2041	74.27531
1997	67.74433	2042	74.36751
1998	67.97552	2043	74.45814
1999	68.20091	2044	74.54723
2000	68.42073	2045	74.63482
2001	68.63515	2046	74.72095
2002	68.84438	2047	74.80566
2003	69.04859	2048	74.88897
2004	69.24794	2049	74.97093
2005	69.44261	2050	75.05157
2006	69.63274	2051	75.13092
2007	69.81849	2052	75.20900
2008	70.00000	2053	75.28585
2009	70.17725	2054	75.36150
2010	70.35067	2055	75.43597
2011	70.52024	2056	75.50930
2012	70.68608	2057	75.58150
2013	70.84830	2058	75.65261
2014	71.00702	2059	75.72266
		2060	75.79165

Now we have all the parameters to perform projection of the mortality rates by the TVF model. Unlike the LC model, the TVF model does not have explicit formula for mortality rates for a year, t . We carry out the projection from the mortality rates in year t to those for $t + 1$ recursively, starting from the baseline mortality, just like the population projection procedure by the cohort component method. Here, we begin with the case of the LC model in the vector approach used in the TVF model.

For the LC model, we can directly project $\lambda_{x,t}$ from a_x , b_x , and k_t , shown in Tables 6.1 and 6.3. For example, $\lambda_{x,t}$ for $t = 2009$ are obtained as follows.

$$\begin{bmatrix} \lambda_{0,2009} \\ \lambda_{5,2009} \\ \lambda_{10,2009} \\ \lambda_{15,2009} \\ \lambda_{20,2009} \\ \lambda_{25,2009} \\ \lambda_{30,2009} \\ \lambda_{35,2009} \\ \lambda_{40,2009} \\ \lambda_{45,2009} \\ \lambda_{50,2009} \\ \lambda_{55,2009} \\ \lambda_{60,2009} \\ \lambda_{65,2009} \\ \lambda_{70,2009} \\ \lambda_{75,2009} \\ \lambda_{80,2009} \\ \lambda_{85,2009} \\ \lambda_{90,2009} \\ \lambda_{95,2009} \\ \lambda_{100,2009} \\ \lambda_{105,2009} \\ \lambda_{110,2009} \end{bmatrix} = \begin{bmatrix} -6.40128 \\ -9.18976 \\ -9.66274 \\ -9.08868 \\ -8.27315 \\ -8.11789 \\ -7.93639 \\ -7.63442 \\ -7.23336 \\ -6.81780 \\ -6.38395 \\ -6.02434 \\ -5.65698 \\ -5.29455 \\ -4.81176 \\ -4.23470 \\ -3.60353 \\ -2.90363 \\ -2.22321 \\ -1.61599 \\ -1.08484 \\ -0.66740 \\ -0.37838 \end{bmatrix} - 0.41834 \begin{bmatrix} 0.04189 \\ 0.03974 \\ 0.03446 \\ 0.02449 \\ 0.01682 \\ 0.01803 \\ 0.01947 \\ 0.02000 \\ 0.01945 \\ 0.02022 \\ 0.02074 \\ 0.02220 \\ 0.02491 \\ 0.03069 \\ 0.03394 \\ 0.03584 \\ 0.03636 \\ 0.03244 \\ 0.02672 \\ 0.02077 \\ 0.01479 \\ 0.00919 \\ 0.00499 \end{bmatrix} = \begin{bmatrix} -6.41880 \\ -9.20639 \\ -9.67716 \\ -9.09892 \\ -8.28018 \\ -8.12543 \\ -7.94453 \\ -7.64278 \\ -7.24150 \\ -6.82626 \\ -6.39262 \\ -6.03363 \\ -5.66740 \\ -5.30739 \\ -4.82596 \\ -4.24970 \\ -3.61874 \\ -2.91721 \\ -2.23439 \\ -1.62468 \\ -1.09103 \\ -0.67125 \\ -0.38047 \end{bmatrix}$$

We can obtain $\lambda_{x,2010}$ by changing $k_{2009} = -0.41834$ to $k_{2010} = -1.03318$ in the above formula. Then, the difference between $\lambda_{x,2010}$ and $\lambda_{x,2009}$ is equal to $(k_{2010} - k_{2009})b_x$. Therefore, we can obtain $\lambda_{x,2010}$ by adding $-\rho_{x,2009} = (k_{2010} - k_{2009})b_x$ to $\lambda_{x,2009}$, which corresponds to the vector field approach.

This procedure is expressed in the following diagram. The left-hand box shows the coordinates $(x, y) = (x, \lambda_{x,2009})$ by the LC model for $t = 2009$. The center box is the vector ρ that indicates the change in each point on the mortality curve for the LC model. The right box showing the coordinates $(x, y) = (x, \lambda_{x,2010})$ is obtained by adding the center box to the left-hand one.

t = 2009		Change		t = 2010	
x	y	x	y	x	y
0	-6.41880	0	-0.02575	0	-6.44456
5	-9.20639	0	-0.02444	5	-9.23082
10	-9.67716	0	-0.02119	10	-9.69835
15	-9.09892	0	-0.01506	15	-9.11398
20	-8.28018	0	-0.01034	20	-8.29052
25	-8.12543	0	-0.01109	25	-8.13652
30	-7.94453	0	-0.01197	30	-7.95650
35	-7.64278	0	-0.01230	35	-7.65508
40	-7.24150	0	-0.01196	40	-7.25346
45	-6.82626	0	-0.01243	45	-6.83869
50	-6.39262	0	-0.01275	50	-6.40537
55	-6.03363	0	-0.01365	55	-6.04728
60	-5.66740	0	-0.01532	60	-5.68272
65	-5.30739	0	-0.01887	65	-5.32626
70	-4.82596	0	-0.02087	70	-4.84683
75	-4.24970	0	-0.02203	75	-4.27173
80	-3.61874	0	-0.02236	80	-3.64110
85	-2.91721	0	-0.01995	85	-2.93715
90	-2.23439	0	-0.01643	90	-2.25082
95	-1.62468	0	-0.01277	95	-1.63746
100	-1.09103	0	-0.00910	100	-1.10012
105	-0.67125	0	-0.00565	105	-0.67690
110	-0.38047	0	-0.00307	110	-0.38354

Next, we consider the LD model. From Proposition 5, we can derive the difference $\tau_{y,t} = x(t+1) - x(t)$ as follows.

$$x(t+1) - x(t) = \frac{g_{t+1} - g_t}{1 - g_{t+1}} x(t) + \frac{f_{t+1} - f_t}{1 - g_{t+1}}$$

For $t = 2009$, we have

$$x(2010) - x(2009) = -0.0019148x(t) + 0.30780 = -0.0019148 \begin{bmatrix} 0 \\ 5 \\ 10 \\ 15 \\ 20 \\ 25 \\ 30 \\ 35 \\ 40 \\ 45 \\ 50 \\ 55 \\ 60 \\ 65 \\ 70 \\ 75 \\ 80 \\ 85 \\ 90 \\ 95 \\ 100 \\ 105 \\ 110 \end{bmatrix} + 0.30780 = \begin{bmatrix} 0.30780 \\ 0.29822 \\ 0.28865 \\ 0.27908 \\ 0.26950 \\ 0.25993 \\ 0.25035 \\ 0.24078 \\ 0.23121 \\ 0.22163 \\ 0.21206 \\ 0.20248 \\ 0.19291 \\ 0.18334 \\ 0.17376 \\ 0.16419 \\ 0.15461 \\ 0.14504 \\ 0.13547 \\ 0.12589 \\ 0.11632 \\ 0.10675 \\ 0.09717 \end{bmatrix}$$

Now we consider the projection procedure for the LD model with a similar diagram as that for the LC.

The first box from the left is $(x, y) = (x, \lambda_{x,2009})$ by the LD model. The second box expresses the vector τ of change for the LD model. Note that $\tau_{y,t} = x(t+1) - x(t)$ is plugged in the "x" (left) column here, which corresponds to the direction of the flow for the LD model. We obtain the third box by adding the first and second ones. It shows the relationship between x and y . However, this is not a normal representation because the values of x are not integers. Therefore, we "standardize" it by linear interpolations and obtain a normal representation, shown in the fourth box.

t = 2009		Change		t = 2010		t = 2010	
x	y	x	y	x	y	x	y
0	-6.41880	0.30780	0	0.30780	-6.41880	0	
5	-9.20639	0.29822	0	5.29822	-9.20639	5	-9.17441
10	-9.67716	0.28865	0	10.28865	-9.67716	10	-9.65753
15	-9.09892	0.27908	0	15.27908	-9.09892	15	-9.16408
20	-8.28018	0.26950	0	20.26950	-8.28018	20	-8.31101
25	-8.12543	0.25993	0	25.25993	-8.12543	25	-8.13085
30	-7.94453	0.25035	0	30.25035	-7.94453	30	-7.95642
35	-7.64278	0.24078	0	35.24078	-7.64278	35	-7.66127
40	-7.24150	0.23121	0	40.23121	-7.24150	40	-7.25807
45	-6.82327	0.22163	0	45.22163	-6.82327	45	-6.84214
50	-6.38706	0.21206	0	50.21206	-6.38706	50	-6.40529
55	-6.02569	0.20248	0	55.20248	-6.02569	55	-6.04032
60	-5.65690	0.19291	0	60.19291	-5.65690	60	-5.67083
65	-5.29181	0.18334	0	65.18334	-5.29181	65	-5.30685
70	-4.80574	0.17376	0	70.17376	-4.80574	70	-4.82438
75	-4.22945	0.16419	0	75.16419	-4.22945	75	-4.24906
80	-3.59925	0.15461	0	80.15461	-3.59925	80	-3.61961
85	-2.90073	0.14504	0	85.14504	-2.90073	85	-2.92162
90	-2.22207	0.13547	0	90.13547	-2.22207	90	-2.23970
95	-1.61629	0.12589	0	95.12589	-1.61629	95	-1.63079
100	-1.08622	0.11632	0	100.11632	-1.08622	100	-1.09765
105	-0.66927	0.10675	0	105.10675	-0.66927	105	-0.67709
110	-0.38016	0.09717	0	110.09717	-0.38016	110	-0.38483

In this case, we cannot obtain the value for age 0 by linear interpolation. However, it causes no problem because we do not use the projected mortality rates by the LD model for the juvenile area.

Finally, we describe the TVF procedure that combines the two models. In the TVF model, we make the vector of change ζ by the weighted average of the vectors ρ and τ . Using $w(x, t)$, which we have already prepared, ζ is constructed as in the following diagram.

ρ		$1 - w(x, t)$	τ		$w(x, t)$	ξ	
x	y		x	y		x	y
0	-0.02575	1.00000	0.30780	0	0.00000	0.00000	-0.02575
0	-0.02444	1.00000	0.29822	0	0.00000	0.00000	-0.02444
0	-0.02119	1.00000	0.28865	0	0.00000	0.00000	-0.02119
0	-0.01506	1.00000	0.27908	0	0.00000	0.00000	-0.01506
0	-0.01034	1.00000	0.26950	0	0.00000	0.00000	-0.01034
0	-0.01109	1.00000	0.25993	0	0.00000	0.00000	-0.01109
0	-0.01197	1.00000	0.25035	0	0.00000	0.00000	-0.01197
0	-0.01230	1.00000	0.24078	0	0.00000	0.00000	-0.01230
0	-0.01196	1.00000	0.23121	0	0.00000	0.00000	-0.01196
0	-0.01243	0.83526	0.22163	0	0.16474	0.03651	-0.01038
0	-0.01275	0.67052	0.21206	0	0.32948	0.06987	-0.00855
0	-0.01365	0.50578	0.20248	0	0.49422	0.10007	-0.00690
0	-0.01532	0.34104	0.19291	0	0.65896	0.12712	-0.00522
0	-0.01887	0.17630	0.18334	0	0.82370	0.15102	-0.00333
0	-0.02087	0.01155	0.17376	0	0.98845	0.17175	-0.00024
0	-0.02203	0.00000	0.16419	0	1.00000	0.16419	0.00000
0	-0.02236	0.00000	0.15461	0	1.00000	0.15461	0.00000
0	-0.01995	0.00000	0.14504	0	1.00000	0.14504	0.00000
0	-0.01643	0.00000	0.13547	0	1.00000	0.13547	0.00000
0	-0.01277	0.00000	0.12589	0	1.00000	0.12589	0.00000
0	-0.00910	0.00000	0.11632	0	1.00000	0.11632	0.00000
0	-0.00565	0.00000	0.10675	0	1.00000	0.10675	0.00000
0	-0.00307	0.00000	0.09717	0	1.00000	0.09717	0.00000

Then, we can perform the projection like the procedure for the LD model, except we use ξ for the vector for change, as shown in the following diagram.

t = 2009		Change		t = 2010		t = 2010	
x	y	x	y	x	y	x	y
0	-6.41880	0.00000	-0.02575	0.00000	-6.44456	0	-6.44456
5	-9.20639	0.00000	-0.02444	5.00000	-9.23082	5	-9.23082
10	-9.67716	0.00000	-0.02119	10.00000	-9.69835	10	-9.69835
15	-9.09892	0.00000	-0.01506	15.00000	-9.11398	15	-9.11398
20	-8.28018	0.00000	-0.01034	20.00000	-8.29052	20	-8.29052
25	-8.12543	0.00000	-0.01109	25.00000	-8.13652	25	-8.13652
30	-7.94453	0.00000	-0.01197	30.00000	-7.95650	30	-7.95650
35	-7.64278	0.00000	-0.01230	35.00000	-7.65508	35	-7.65508
40	-7.24150	0.00000	-0.01196	40.00000	-7.25346	40	-7.25346
45	-6.82327	0.03651	-0.01038	45.03651	-6.83366	45	-6.83675
50	-6.38706	0.06987	-0.00855	50.06987	-6.39560	50	-6.40159
55	-6.02569	0.10007	-0.00690	55.10007	-6.03260	55	-6.03981
60	-5.65690	0.12712	-0.00522	60.12712	-5.66212	60	-5.67128
65	-5.29181	0.15102	-0.00333	65.15102	-5.29514	65	-5.30752
70	-4.80574	0.17175	-0.00024	70.17175	-4.80598	70	-4.82441
75	-4.22945	0.16419	0.00000	75.16419	-4.22945	75	-4.24906
80	-3.59925	0.15461	0.00000	80.15461	-3.59925	80	-3.61961
85	-2.90073	0.14504	0.00000	85.14504	-2.90073	85	-2.92162
90	-2.22207	0.13547	0.00000	90.13547	-2.22207	90	-2.23970
95	-1.61629	0.12589	0.00000	95.12589	-1.61629	95	-1.63079
100	-1.08622	0.11632	0.00000	100.11632	-1.08622	100	-1.09765
105	-0.66927	0.10675	0.00000	105.10675	-0.66927	105	-0.67709
110	-0.38016	0.09717	0.00000	110.09717	-0.38016	110	-0.38483

The projected log mortality rates for $t = 2009$ and 2010 by the LC and TVF models for all ages x ($x \leq 110$) are shown in the Appendix (Tables A.2 and A.3).

Fig. 6.7 shows the projected e_0 for the LC and TVF models. We see that the two models yield almost the same trajectories for future life expectancy. However, the projected age patterns differ.

We compare the relative mortality rates for the actual and estimated rates by the LC and TVF models, compared with the recent level (the average from 2006 to 2010). Fig. 6.8 shows the actual relative mortality rates. With the LC model, actual mortality is estimated and projected as in Fig. 6.9. We observe that the LC model expresses mortality improvement in a vertical direction, which is a characteristic of decline-type models.

On the other hand, Fig. 6.10, an estimation and projection by the TVF model, shows that the TVF model succeeds in expressing the shifting of mortality improvement in the direction of older ages.

Fig. 6.11 compares the projected age pattern of m_x for the LC and TVF models. The projected m_x curves by the LC model exhibit an unnatural pattern because their slope diminishes once around the age of 60 years and then becomes much steeper above the age of 80 years. The curve of the TVF model is more plausible.

Fig. 6.7: Projected Life Expectancy

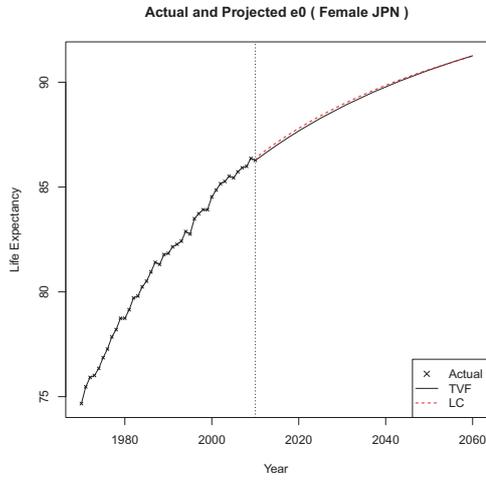


Fig. 6.8: Relative Mortality Rates (Actual)

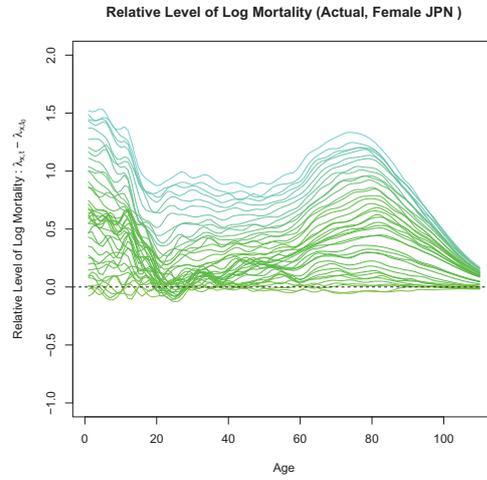


Fig. 6.9: Relative Mortality Rates (LC)

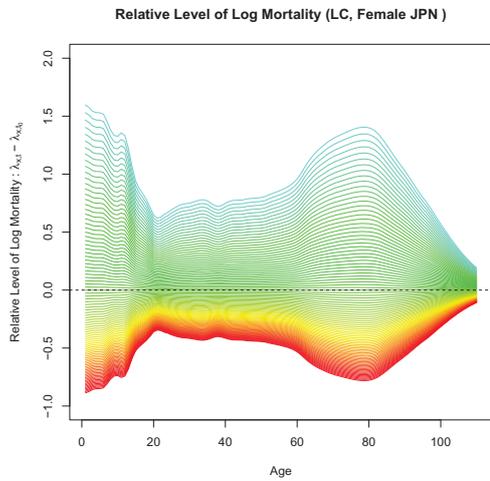


Fig. 6.10: Relative Mortality Rates (TVF)

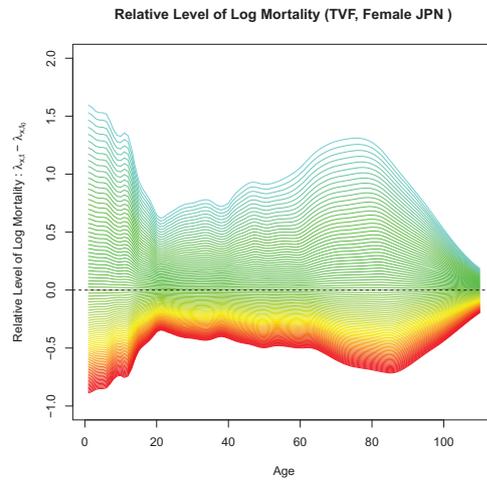
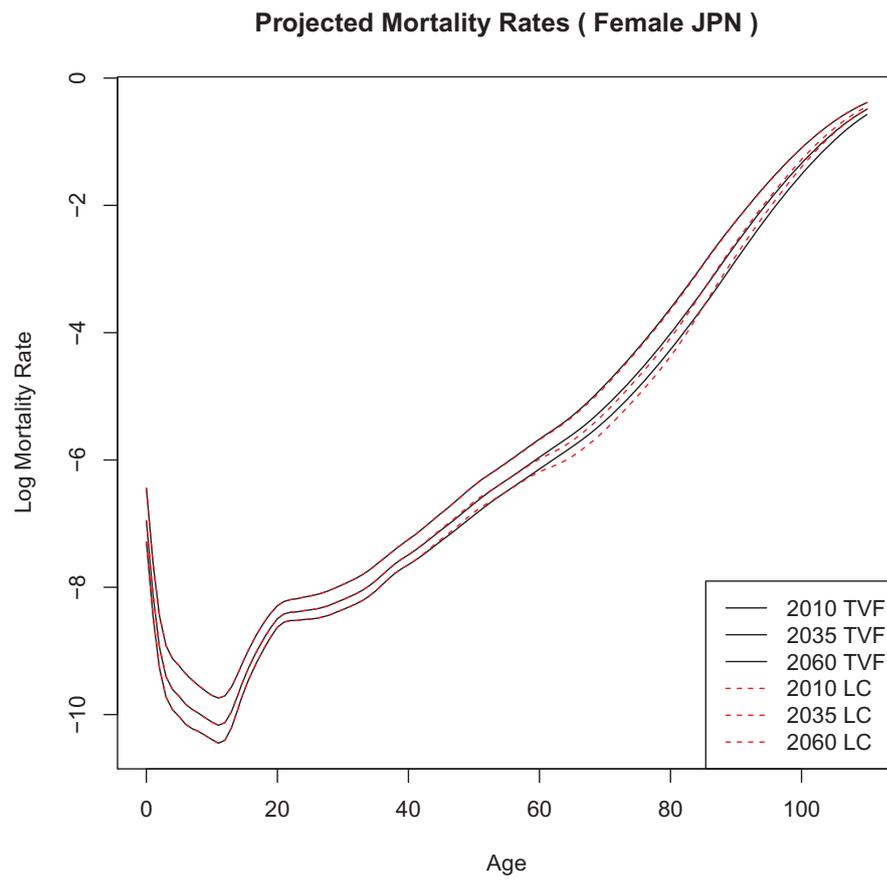


Fig. 6.11: Comparison of m_x curves (LC and TVF)

6.1.3 Discussion of Mortality Projection

Li et al. (2013) propose an extension of the LC model, the Lee–Carter method extended with rotation (LC-ER). In this subsection, we review their method and discuss a comparison with the TVF model.

The first step of their model is the modification of b_x in the LC model. They smooth b_x at ages 15–65 years to equal the average value for this age range, and then reduce the values at ages 0–14 years to this average. Next, they proportionally adjust the values of b_x at ages 70 years and older to make $b_{70} = b_{65}$, and then, proportionally adjust the values of b_x at all ages to make b_x sum to 1. They call this modified b_x the ultimate values, $b_u(x)$.

The second step of their model is to set $B(x, t)$, which means the extended LC age pattern of mortality decline at time t starting with $b_o(x)$ by the LC model and smoothly converging to $b_u(x)$. They define a linear weight function, $w(t)$, and a smooth weight function, $w_s(t)$, as:

$$w(t) = \frac{e_o(t) - 80}{e_o^u - 80}$$

$$w_s(t) = \left\{ 0.5 \left[1 + \sin \left[\frac{\pi}{2} (2 \cdot w(t) - 1) \right] \right] \right\}^p$$

where $e_o(t)$ is the life expectancy at time t estimated by the LC method, and e_o^u is the level of the life expectancy at which the rotation finishes, which they choose to be 102 years. p adjusts the speed of rotation and $p = 0.5$ is taken as the default.

Then, they set $B(x, t)$ as:

$$B(x, t) = \begin{cases} b_o(x), & e_o(t) < 80 \\ (1 - w_s(t))b_o(x) + w_s(t)b_u(x), & 80 \leq e_o(t) < e_o^u \\ b_u(x), & e_o^u \leq e_o(t) \end{cases}$$

Two points should be noticed as unique characteristics of their model. First, it needs only the parameters that are estimated in the usual LC method. Therefore, it is easy to apply their rotation if one already has the estimated parameters by the LC method. Second, their model does not depend on the actual change of age pattern of mortality decline. They mention that the rotation is subtle and difficult to handle in mortality models. This would be one of the reasons why they do not use the actual pattern in the model.

However, for Japanese case, we see that actual mortality already exhibits a strong shifting feature. Therefore, we have sufficient data to model the shift of mortality curves. Moreover, their model is based essentially on the LC framework, which is a decline-type model, even though the rotation feature is captured more clearly using the shift-type model. It is usually considered difficult to blend a shift-type and a decline-type model, so the choice might be made to stay in the

decline-type framework. However, we show it is possible to construct a hybrid model with a tangent vector approach. Thus, it is more efficient to use the shift-type model directly in this situation. Therefore, the TVF model is considered better for Japanese mortality projection.

Chapter 7

Summary and Conclusion

In this dissertation, we proposed the Tangent Vector Field (TVF) model, which is an extended Linear Difference (LD) model, for Japanese mortality projection, and showed its applications.

Following the mathematical formulations for *decline*-type and *shift*-type models, we compared five models; the Proportional Hazard (PH) model, the Lee-Carter (LC) model, the Horizontal Shifting (HS) model, the Horizontal Lee-Carter (HL) model, and the Linear Difference (LD) model.

We first examined the *decline*-type models. For the PH model, the estimated rates did not exhibit good fit, particularly in the older age groups. On the other hand, we saw that the fit with the actual values was fairly improved by using the LC model because of its flexibility, which allows different mortality improvement rates by age. From the observation of $\rho_{x,t}$, we saw that most of the actual mortality improvement rates have mountain-shaped curves with peaks. In contrast, the mortality improvement rates under the PH model are horizontal. This difference in shape was viewed as a cause for the estimates by the PH model not being well fitted. The peak of the mortality improvement rate by the LC model was like that of the actual value and this improves the fit. However, the age distribution of the rates was fixed in the LC model, whereas it changed dynamically in the actual values. Thus, the actual age distribution of mortality improvement rates changed over time and was not constant, as in the LC model, causing a propensity for error in the LC model. We could see this result as a limitation when the mortality improvement is considered as a *decline*.

Next, we observed *shift*-type models. We saw that the performance of fit by the HS model is much better than by the PH model, even though both have the same structure. We saw that the HL model seemed to be an improvement on the HS model. However, we also observed that the improvement between the *shift* pair was not as large as the *decline* pair. We observed that the LD model fit well with the actual values. Even though the performance seemed to be lower than the HL model, the LD model had an advantage in that it needed fewer parameters than

the HL model. From the observation of $\tau_{y,t}$, we saw that the force of age increase for the HS model was almost horizontal, and that the distribution of $\tau_{y,t}$ for the HL model was similar. According to these restrictions, the two models sometimes exhibited big differences from the actual rates. Compared to these models, we observed that the LD model performed better.

Following these considerations, we compared the LC and LD models from a statistical viewpoint. From the observation of the proportion where the log actual mortality rates were outside of the CIs for each age in the LC and LD models, we saw that even though the proportions of LD were higher under about 75 years of age, the performance of LD was considered better than that of LC over 75 years of age. This result suggested that *shift* was more strongly supported as a factor behind the recent old age mortality improvement in Japan compared to *decline*. This result also implied that a better construction for mortality projection model would be a blended model that had the LC property in youth and the LD property in older age. We also showed that the LD model is useful for not only projection but also analyses of mortality. We proposed a new decomposition method for the modal age at death using the LD model, and gave decomposition analyses with the method. Because the LD model was developed originally for mortality projections, the number of parameters was reduced for parsimony. This might be a restriction in terms of a flexible expression for various types of mortality situation. However, this feature brought another possibility to derive simple analytical formulas. The decomposition that we proposed was easy to apply when the mortality curves were modeled by the LD model, and had a clear interpretation composed by shifting, compression, and other parts. From the results of the decomposition, we observed a strong parallel shifting feature from 1980 to 2000 that also increased M_t by shifting components. On the other hand, the compression components played a larger part for the increase of M_t before 1970 and after 2000. The analytical decomposition of trends in the modal age at death would be considered useful for understanding old age mortality.

Lastly, we proposed the TVF model applying the idea of the tangent vector fields on the log mortality surface. We showed a fully specified example of the projection procedure of the TVF model with all the constants and coefficients applied for Japanese mortality projection. Then, we compared the results of the mortality projection by the TVF model with those by the LC model. From the observation of the relative mortality rates, we saw that the LC model expressed mortality improvement only in a vertical direction, whereas the TVF model succeeded in expressing the shifting of mortality improvement in the direction of older ages that were observed in the actual mortality. We also compared the projected m_x curves. m_x by the LC model exhibited an unnatural pattern, since the slope of the curve diminished once around the age of 60 years and then became much steeper above the age of 80 years. The curve of the TVF model was more plausible. As a whole, we revealed that the TVF model had many advantages for Japanese mortality pro-

jection compared with the LC model.

We believe we can show that the TVF model proposed in this dissertation is not only quite useful for Japanese mortality projection but has various applicability. At this point in time, there may be few countries with such strong shifting features for old age mortality as Japan. However, some countries are likely to experience the same mortality situation as Japan in the future through the extension of life expectancy. The TVF model will be a useful tool for projections in such situations.

References

- Beard, R. E. (1971) "Some aspects of theories of mortality, cause of death analysis, forecasting and stochastic processes", in *Biological Aspects of Demography*, London: Taylor & Francis Ltd.
- Bongaarts, J. (2005) "Long-range Trends in Adult Mortality: Models and Projection Methods", *Demography*, Vol. 42, No. 1, pp. 23–49.
- Box, G. E., G. M. Jenkins, and G. C. Reinsel (2013) *Time series analysis: forecasting and control*: John Wiley & Sons.
- Brass, W. (1971) "On the Scale of Mortality", in W. Brass ed. *Biological Aspects of Demography*: Taylor and Francis Ltd, pp. 69–110.
- Canudas-Romo, V. (2008) "The modal age at death and the shifting mortality hypothesis", *Demographic Research*, Vol. 19, No. 30, pp. 1179–1204.
- Coale, A. J. and P. Demeny (1983) *Regional Model Life Tables and Stable Populations, 2nd Edition*, New York: Academic Press.
- Fries, J. F. (1980) "Aging, Natural Death, and the Compression of Morbidity", *New England Journal of Medicine*, Vol. 303, pp. 130–135.
- Gompertz, B. (1825) "On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies", *Philosophical Transactions of the Royal Society of London*, Vol. 115, pp. 513–583.
- Greville, T. (1981) "Moving-weighted-average smoothing extended to the extremities of the data. II. Methods", *Scandinavian Actuarial Journal*, Vol. 1981, No. 2, pp. 65–81.
- Heligman, L. and J. H. Pollard (1980) "The age pattern of mortality", *Journal of the Institute of Actuaries*, Vol. 107, pp. 49–80.
- Horiuchi, S. and J. Wilmoth (1995) "Aging of mortality decline", Paper Presented at the Annual Meeting of the Population Association of America, San Francisco, California.
- Horiuchi, S., N. Ouellette, S. L. K. Cheung, and J.-M. Robine (2013) "Modal Age at Death: Lifespan Indicator in the Era of Longevity Extension", Paper presented at the XXVII IUSSP International Population Conference.

- Human Mortality Database. University of California, Berkeley (USA) and Max Planck Institute for Demographic Research (Germany). Available at www.mortality.org or www.humanmortality.de.
- Igawa, T. (2013) "Analysis of the Residual Structure of the Lee–Carter Model: The Case of Japanese Mortality", *Asia-Pacific Journal of Risk and Insurance*, Vol. 7, No. 2, pp. 53–80.
- Ishii, F. and G. Lanzieri (2013) "Interpreting and Projecting Mortality Trends for European Countries by Using the LD Model", Paper presented at the XXVII IUSSP International Population Conference.
- Ishii, F. (2006) "Trends of Japanese Life Expectancy and Mortality Projection Models (in Japanese)", *Journal of Population Problems*, Vol. 62, No. 3, pp. 21–30.
- Japanese Mortality Database. National Institute of Population and Social Security Research, Available at <http://www.ipss.go.jp/p-toukei/JMD/index-en.html>.
- Kogure, A. and T. Hasegawa (2005) "Statistical Modelling of the Projected Life Tables: the Lee-Carter method and its extensions (in Japanese)", *Policy and Governance Working Paper Series*, Vol. 71.
- Komatsu, R. (2002) "A Construction of Future Life Table in Japan Using a Relational Model (in Japanese)", *Journal of Population Problems*, Vol. 58, No. 3, pp. 3–14.
- Lee, R. and L. Carter (1992) "Modeling and Forecasting U.S. Mortality", *Journal of the American Statistical Association*, Vol. 87, No. 419, pp. 659–675.
- Lee, R. and T. Miller (2001) "Evaluating the Performance of the Lee-Carter Method for Forecasting Mortality", *Demography*, Vol. 38, No. 4, pp. 537–549.
- Li, N., R. Lee, and P. Gerland (2013) "Extending the Lee-Carter Method to Model the Rotation of Age Patterns of Mortality Decline for Long-Term Projections", *Demography*, Vol. 50, No. 6, pp. 2037–2051.
- Makeham, W. M. (1860) "On the Law of Mortality and the Construction of Annuity Tables", *The Assurance Magazine, and Journal of the Institute of Actuaries*, Vol. 8, No. 6, pp. 301–310.
- Ministry of Health, Labor and Welfare "Life Tables".
- Nanjo, Z. and K. Yoshinaga (2003) "Forecasting Japan's Life Tables with Special Reference to the Lee-Carter Method (in Japanese)", *The 55th Annual Meeting of the Population Association of Japan, Abstract Booklet*, p. 57.

- NIPSSR (2002) *Population Projections for Japan: 2001-2050 (With Long-Range Population Projections: 2051-2100)*.
- NIPSSR (2012) *Population Projections for Japan (January 2012) - 2011 to 2060 -*.
- Ogawa, N., M. Kondo, M. Tamura, R. Matsukura, T. Saito, A. Mason, S. Tuljapurkar, and N. Li (2002) *Zinko, Keizai, Shakaihoshō Moderu niyoru Chokitenbo - Zintekishihon nimotozuku Approach- (in Japanese)*: Nihon University Population Research Institute.
- Oikawa, K. (2006) "Study on Future Mortality Rate Estimation (in Japanese)", *Kaiho, the Institute of Actuaries in Japan*, Vol. 59, No. 2, pp. 1–28.
- Olshansky, S. and A. Ault (1986) "The Fourth Stage of the Epidemiologic Transition: The Age of Delayed Degenerative Diseases", *The Milbank Quarterly*, Vol. 64, No. 3, pp. 355–391.
- Omran, A. (1971) "The Epidemiologic Transition: A Theory of the Epidemiology of Population Change", *The Milbank Memorial Fund Quarterly*, Vol. 49, No. 4, pp. 509–538.
- Ozeki, M. (2005) "Application of Mortality Models to Japan", Presented at The Living to 100 and Beyond Symposium.
- Perks, W. (1932) "On some experiments on the graduation of mortality statistics", *Journal of the Institute of Actuaries*, Vol. 63, pp. 12–40.
- Rogers, A. and J. S. Little (1994) "Parameterizing Age Patterns of Demographic Rates with the Multiexponential Model Schedule", *Mathematical Population Studies*, Vol. 4, No. 3, pp. 175–195.
- Siler, W. (1979) "A Competing-Risk Model for Animal Mortality", *Ecology*, Vol. 60, No. 4, pp. 750–757.
- Thatcher, A. R., V. Kannisto, and J. W. Vaupel (1998) *The Force of Mortality at Ages 80 to 120*: Odense University Press.
- Tuljapurkar, S., N. Li, and C. Boe (2000) "A Universal Pattern of Mortality Decline in the G7 Countries", *Nature*, Vol. 405, pp. 789–782.
- United Nations (1956) *Methods for Population Projections by Sex and Age, Series A, Population Studies No.25*.
- (1982) *Model Life Tables for Developing Countries, Population Studies No.77*.
- (2006) *World Population Prospects: The 2004 Revision, Volume III: Analytical Report*: Department of Economic and Social Affairs, Population Division.

- Wilmoth, J. R. (1995) "Are Mortality Projections Always More Pessimistic When Disaggregated by Cause of Death?", *Mathematical Population Studies*, Vol. 5, No. 4, pp. 293–319.
- (1996) "Mortality Projections for Japan", in G. Caselli and A. D. Lopez eds. *Health and Mortality among Elderly Populations*: Oxford Univ. Press, pp. 266–287.
- (1997) "In Search of Limits", in K. W. Wachter and C. E. Finch eds. *Between Zeus and the Salmon*, Washington, D.C.: National Academy Press, pp. 38–64.
- (2011) "Increase of Human Longevity: Past, Present, and Future", *The Japanese Journal of Population*, Vol. 9, No. 1, pp. 155–161.

Appendix A

Supplement Tables

Table A.1: Estimated a_x and b_x

x	a_x	b_x	x	a_x	b_x
0	-6.40128	0.04189	55	-6.02434	0.02220
1	-7.56570	0.04148	56	-5.95121	0.02256
2	-8.40707	0.04093	57	-5.87799	0.02299
3	-8.87631	0.04008	58	-5.80321	0.02345
4	-9.08272	0.03977	59	-5.72895	0.02406
5	-9.18976	0.03974	60	-5.65698	0.02491
6	-9.31893	0.03945	61	-5.58885	0.02609
7	-9.41992	0.03788	62	-5.52176	0.02741
8	-9.50769	0.03590	63	-5.45129	0.02865
9	-9.59478	0.03467	64	-5.37567	0.02974
10	-9.66274	0.03446	65	-5.29455	0.03069
11	-9.70433	0.03523	66	-5.20694	0.03148
12	-9.66996	0.03467	67	-5.11474	0.03217
13	-9.53375	0.03176	68	-5.01886	0.03284
14	-9.32034	0.02773	69	-4.91827	0.03346
15	-9.08868	0.02449	70	-4.81176	0.03394
16	-8.88028	0.02280	71	-4.70201	0.03439
17	-8.69083	0.02161	72	-4.58860	0.03479
18	-8.52585	0.02031	73	-4.47221	0.03514
19	-8.38658	0.01858	74	-4.35407	0.03549
20	-8.27315	0.01682	75	-4.23470	0.03584
21	-8.20709	0.01613	76	-4.11360	0.03614
22	-8.17738	0.01632	77	-3.98995	0.03635
23	-8.16238	0.01698	78	-3.86393	0.03649
24	-8.13896	0.01742	79	-3.73518	0.03651
25	-8.11789	0.01803	80	-3.60353	0.03636
26	-8.09379	0.01849	81	-3.46956	0.03604
27	-8.06580	0.01895	82	-3.33195	0.03545
28	-8.02747	0.01921	83	-3.19100	0.03461
29	-7.98383	0.01937	84	-3.04772	0.03360
30	-7.93639	0.01947	85	-2.90363	0.03244
31	-7.89004	0.01973	86	-2.76162	0.03126
32	-7.83823	0.01997	87	-2.62259	0.03007
33	-7.77986	0.02021	88	-2.48642	0.02891
34	-7.71094	0.02023	89	-2.35342	0.02781
35	-7.63442	0.02000	90	-2.22321	0.02672
36	-7.54937	0.01943	91	-2.09564	0.02555
37	-7.46423	0.01895	92	-1.97088	0.02429
38	-7.38079	0.01873	93	-1.84934	0.02303
39	-7.30510	0.01897	94	-1.73122	0.02186
40	-7.23336	0.01945	95	-1.61599	0.02077
41	-7.15649	0.01984	96	-1.50350	0.01971
42	-7.07419	0.02008	97	-1.39322	0.01854
43	-6.98839	0.02015	98	-1.28600	0.01728
44	-6.90223	0.02016	99	-1.18309	0.01602
45	-6.81780	0.02022	100	-1.08484	0.01479
46	-6.73427	0.02036	101	-0.99128	0.01359
47	-6.64950	0.02056	102	-0.90262	0.01243
48	-6.55913	0.02062	103	-0.81902	0.01130
49	-6.46950	0.02067	104	-0.74060	0.01021
50	-6.38395	0.02074	105	-0.66740	0.00919
51	-6.30493	0.02093	106	-0.59946	0.00822
52	-6.23445	0.02131	107	-0.53670	0.00731
53	-6.16550	0.02163	108	-0.47905	0.00647
54	-6.09599	0.02191	109	-0.42634	0.00570
			110	-0.37838	0.00499

Table A.2: Projected Log Mortality Rates by the LC model

Age	2009	2010	Age	2009	2010
0	-6.41880	-6.44456	55	-6.03363	-6.04728
1	-7.58306	-7.60856	56	-5.96065	-5.97452
2	-8.42419	-8.44936	57	-5.88761	-5.90175
3	-8.89308	-8.91772	58	-5.81302	-5.82744
4	-9.09936	-9.12382	59	-5.73901	-5.75380
5	-9.20639	-9.23082	60	-5.66740	-5.68272
6	-9.33543	-9.35969	61	-5.59976	-5.61580
7	-9.43577	-9.45906	62	-5.53323	-5.55008
8	-9.52271	-9.54478	63	-5.46328	-5.48090
9	-9.60928	-9.63060	64	-5.38811	-5.40639
10	-9.67716	-9.69835	65	-5.30739	-5.32626
11	-9.71907	-9.74073	66	-5.22011	-5.23946
12	-9.68446	-9.70577	67	-5.12819	-5.14797
13	-9.54703	-9.56656	68	-5.03260	-5.05280
14	-9.33194	-9.34899	69	-4.93227	-4.95284
15	-9.09892	-9.11398	70	-4.82596	-4.84683
16	-8.88982	-8.90384	71	-4.71639	-4.73754
17	-8.69987	-8.71315	72	-4.60315	-4.62454
18	-8.53435	-8.54683	73	-4.48691	-4.50851
19	-8.39436	-8.40578	74	-4.36892	-4.39074
20	-8.28018	-8.29052	75	-4.24970	-4.27173
21	-8.21383	-8.22375	76	-4.12872	-4.15094
22	-8.18421	-8.19424	77	-4.00516	-4.02751
23	-8.16949	-8.17993	78	-3.87920	-3.90163
24	-8.14625	-8.15696	79	-3.75045	-3.77290
25	-8.12543	-8.13652	80	-3.61874	-3.64110
26	-8.10152	-8.11289	81	-3.48464	-3.50680
27	-8.07372	-8.08537	82	-3.34678	-3.36858
28	-8.03551	-8.04733	83	-3.20548	-3.22676
29	-7.99193	-8.00384	84	-3.06177	-3.08243
30	-7.94453	-7.95650	85	-2.91721	-2.93715
31	-7.89829	-7.91042	86	-2.77469	-2.79391
32	-7.84658	-7.85886	87	-2.63517	-2.65366
33	-7.78832	-7.80075	88	-2.49851	-2.51629
34	-7.71941	-7.73185	89	-2.36505	-2.38215
35	-7.64278	-7.65508	90	-2.23439	-2.25082
36	-7.55750	-7.56945	91	-2.10633	-2.12204
37	-7.47216	-7.48381	92	-1.98104	-1.99597
38	-7.38863	-7.40015	93	-1.85898	-1.87314
39	-7.31304	-7.32470	94	-1.74036	-1.75381
40	-7.24150	-7.25346	95	-1.62468	-1.63746
41	-7.16479	-7.17699	96	-1.51175	-1.52386
42	-7.08259	-7.09494	97	-1.40097	-1.41237
43	-6.99682	-7.00921	98	-1.29323	-1.30385
44	-6.91067	-6.92307	99	-1.18980	-1.19964
45	-6.82626	-6.83869	100	-1.09103	-1.10012
46	-6.74279	-6.75531	101	-0.99697	-1.00533
47	-6.65810	-6.67073	102	-0.90782	-0.91546
48	-6.56776	-6.58043	103	-0.82375	-0.83069
49	-6.47814	-6.49085	104	-0.74487	-0.75115
50	-6.39262	-6.40537	105	-0.67125	-0.67690
51	-6.31369	-6.32656	106	-0.60289	-0.60795
52	-6.24336	-6.25646	107	-0.53976	-0.54426
53	-6.17455	-6.18785	108	-0.48175	-0.48573
54	-6.10515	-6.11862	109	-0.42872	-0.43222
			110	-0.38047	-0.38354

Table A.3: Projected Log Mortality Rates by the TVF model

Age	2009	2010	Age	2009	2010
0	-6.41880	-6.44456	55	-6.02569	-6.03981
1	-7.58306	-7.60856	56	-5.95219	-5.96652
2	-8.42419	-8.44936	57	-5.87855	-5.89295
3	-8.89308	-8.91772	58	-5.80344	-5.81806
4	-9.09936	-9.12382	59	-5.72896	-5.74357
5	-9.20639	-9.23082	60	-5.65690	-5.67128
6	-9.33543	-9.35969	61	-5.58856	-5.60252
7	-9.43577	-9.45906	62	-5.52101	-5.53490
8	-9.52271	-9.54478	63	-5.44995	-5.46430
9	-9.60928	-9.63060	64	-5.37367	-5.38868
10	-9.67716	-9.69835	65	-5.29181	-5.30752
11	-9.71907	-9.74073	66	-5.20354	-5.22006
12	-9.68446	-9.70577	67	-5.11072	-5.12776
13	-9.54703	-9.56656	68	-5.01416	-5.03158
14	-9.33194	-9.34899	69	-4.91282	-4.93078
15	-9.09892	-9.11398	70	-4.80574	-4.82441
16	-8.88982	-8.90384	71	-4.69605	-4.71494
17	-8.69987	-8.71315	72	-4.58275	-4.60204
18	-8.53435	-8.54683	73	-4.46654	-4.48610
19	-8.39436	-8.40578	74	-4.34862	-4.36825
20	-8.28018	-8.29052	75	-4.22945	-4.24906
21	-8.21383	-8.22375	76	-4.10852	-4.12819
22	-8.18421	-8.19424	77	-3.98507	-4.00490
23	-8.16949	-8.17993	78	-3.85923	-3.87921
24	-8.14625	-8.15696	79	-3.73068	-3.75084
25	-8.12543	-8.13652	80	-3.59925	-3.61961
26	-8.10152	-8.11289	81	-3.46549	-3.48595
27	-8.07372	-8.08537	82	-3.32810	-3.34886
28	-8.03551	-8.04733	83	-3.18742	-3.20840
29	-7.99193	-8.00384	84	-3.04445	-3.06550
30	-7.94453	-7.95650	85	-2.90073	-2.92162
31	-7.89829	-7.91042	86	-2.75910	-2.77941
32	-7.84658	-7.85886	87	-2.62044	-2.64005
33	-7.78832	-7.80075	88	-2.48462	-2.50357
34	-7.71941	-7.73185	89	-2.35196	-2.37022
35	-7.64278	-7.65508	90	-2.22207	-2.23970
36	-7.55750	-7.56945	91	-2.09481	-2.11184
37	-7.47216	-7.48381	92	-1.97035	-1.98676
38	-7.38863	-7.40015	93	-1.84910	-1.86486
39	-7.31304	-7.32470	94	-1.73125	-1.74634
40	-7.24150	-7.25346	95	-1.61629	-1.63079
41	-7.16418	-7.17655	96	-1.50405	-1.51799
42	-7.08134	-7.09411	97	-1.39400	-1.40746
43	-6.99496	-7.00805	98	-1.28701	-1.29989
44	-6.90824	-6.92155	99	-1.18430	-1.19647
45	-6.82327	-6.83675	100	-1.08622	-1.09765
46	-6.73921	-6.75289	101	-0.99281	-1.00351
47	-6.65380	-6.66780	102	-0.90427	-0.91425
48	-6.56292	-6.57741	103	-0.82077	-0.83002
49	-6.47288	-6.48752	104	-0.74241	-0.75094
50	-6.38706	-6.40159	105	-0.66927	-0.67709
51	-6.30792	-6.32214	106	-0.60134	-0.60848
52	-6.23712	-6.25086	107	-0.53859	-0.54506
53	-6.16778	-6.18150	108	-0.48091	-0.48675
54	-6.09779	-6.11164	109	-0.42817	-0.43340
			110	-0.38016	-0.38483