

Faculté des sciences

Impact of Climate Change on Mortality Risk

A Prospective Approach on Belgian Data

Author: **Owell GUILBEAU COMPÈRE**
Supervisor: **Prof. Donatien HAINAUT**
Reader: **Prof. Pierre ARS**
Academic year 2024–2025
Master [120] en sciences actuarielles



Institute of Statistics, Biostatistics and Actuarial Sciences

Impact of Climate Change on Mortality Risk
A Prospective Approach on Belgian Data

Jury members:

Prof. Donatien HAINAUT *Supervisor*
Prof. Pierre ARS *Reader*

Thesis submitted in partial fulfillment of
the requirements for the Master's degree
in Actuarial Sciences.

by:

Owell GUILBEAU COMPÈRE

Louvain-la-Neuve
June 2025

Preface

This thesis starts from a simple observation: climate change is no longer background noise but a measurable driver of mortality trajectories that actuaries must explicitly incorporate to price credibly. The aim is pragmatic: to quantify, on Belgian data, the effect of temperatures (and their extremes) on mortality by 2050 and to infer the implications for life insurance premiums. The approach remains close to practice: we rely on established stochastic longevity models, extend them with a climate index and a hot-summer trigger, and compare their fit and predictive performance. The work is reproducible, built on public data and R routines, with an emphasis on parsimony, identifiability, and transparency about limitations (cause-of-death measurement uncertainty, sample choices, climate scenario assumptions). Rather than a global attribution exercise, this is an operational contribution for Belgian portfolios: delivering projections and premium deltas usable in pricing, reserving, and risk management.

Acknowledgment

I would like to express my sincere thanks to Prof. Donatien HAINAUT, my supervisor, for insightful guidance, timely feedback, and steadfast support, and to Prof. Pierre ARS, reader, for careful review. I am grateful to the Institute of Statistics, Biostatistics and Actuarial Sciences for an excellent working environment and to the administrative staff for their help. This thesis relies on public data and tools: the Human Mortality Database, the Global Health Data Exchange (IHME/GBD), and the Royal Meteorological Institute's open datasets, together with open-source R software and packages (in particular **StMoMo**). I also thank my classmates for stimulating exchanges and occasional code checks. Any remaining errors are my own. I am thankful to my family and friends for their encouragement throughout.

Note: Generative AI was used for translation purposes.

Contents

Introduction	1
1 Literature Review	3
1.1 Stochastic Modeling of Mortality Rates	3
1.1.1 GAPC Stochastic Mortality Models Family	3
1.1.2 Lee-Carter as an Example of GAPC Model	5
1.1.3 Model Fitting	5
1.2 Boumezoued et al.'s Model	7
1.2.1 Classical Lee-Carter Model with Construction of a Climate Index	8
1.2.2 Harvesting Effect and Age Range Constitution	8
1.2.3 Model Fitting	9
1.3 Milhaud and Gouthon's Model	10
1.3.1 SPO Model with Triggering Event	10
1.3.2 Age Threshold Construction	11
1.3.3 Model Fitting	11
2 Boumezoued et al.'s Approach on Belgian Data	13
2.1 Available Belgian Data	13
2.1.1 Belgian Specific Mortality (Global Health Data, GHD)	13
2.1.2 Belgian Global Mortality (Human Mortality Database, HMD)	14
2.1.3 Belgian Meteorological Data (Royal Meteorological Institute, RMI) . . .	14
2.2 Construction of a Climate Index	15
2.2.1 Climate Variables	15
2.2.2 Period of Calibration	16
2.2.3 Fitting Procedure	17
2.2.4 Selected Variables and Final Climate Index	18
2.3 Model Calibration	20
2.3.1 Estimation of Climate-Related Mortality Across Age Groups	20
2.3.2 Integration of the Harvesting Effect	21
2.3.3 Calibration of the Full Model with Continuous Age	21
2.4 Performance Comparison	23
3 Milhaud and Gouthon's Approach on Belgian Data	27
3.1 Available Belgian Data	27
3.1.1 Weekly Belgian Mortality (Short Term Mortality Fluctuation, STMF) .	27
3.1.2 Belgian Global Mortality (Human Mortality Database, HMD)	28

3.1.3	Belgian Meteorological Data (Royal Meteorological Institute, RMI) . . .	28
3.2	Age Threshold Determination	28
3.2.1	Empirical Evidence	28
3.2.2	Linear Correlations	30
3.3	Model Calibration	31
3.3.1	Fitting Methodology	31
3.3.2	Estimation Results	32
4	Prospective Approach on Belgian Data	37
4.1	Final Prospective Model	37
4.1.1	Proposed Model	37
4.1.2	Performance Comparison	39
4.2	Impact at the 2050 Horizon	40
4.2.1	Classical Lee-Carter Model Projections using Random Walk with Drift .	41
4.2.2	Climate Lee-Carter Model Projections Using RCPs	43
4.2.3	Impact on Longevity and Life Insurance Pricing	47
	Conclusion	51
	Bibliography	56

Introduction

Human activity since the Industrial Era has markedly altered atmospheric conditions, particularly the atmospheric CO₂ concentration. Florides and Christodoulides (2009) place the correlation between CO₂ concentration and temperature at the center of the global warming problem faced by the entire planet. This increase in CO₂ concentration leads to rising temperatures and a higher frequency of natural disasters and exceptional events, as explained in the thesis of Bador (2016). The effects of climate change are felt at multiple levels – global, economic, and health – as shown respectively by Hitz and Smith (2004), Tol (2018), and Hales et al. (2014). One of the most significant consequences is probably an increase in mortality. In its quantitative assessment of climate change risks, the World Health Organization (Hales et al. (2014)) indicates that there could be approximately 250,000 additional deaths per year attributable to climate change between 2030 and 2050. These deaths would be linked in particular to natural disasters, air pollution, disease, undernutrition, and the direct impact of heat on the human body. The most affected populations will, unsurprisingly, be those least prepared: underdeveloped and developing countries. For Western Europe, the majority of climate change-related causes of death are associated with exposure to extreme heat, and thus with rising temperatures. Variations in mortality trends are closely monitored by actuaries to establish sound pricing in life insurance and employee benefits. Indeed, many mortality rate projection methods, such as those presented by Atance et al. (2025), seek to incorporate trends and the impact of environmental factors on mortality. Trends in life expectancy improvement associated with a decline in mortality over the years are well known and are notably captured by longevity models introduced by RD Lee and LR Carter (1992) and Plat (2009). More realistic forward-looking projections and calculations enable insurers to better mitigate pricing risk and thus maintain solvency and/or remain competitive in the market. We may therefore ask whether this trend of improving life expectancy (medical advances, hygiene, etc.) will be affected by the increasing mortality induced by global warming, or, at the very least, how these two opposing forces will interact and what impacts this will have on the pricing of life or pension insurance products. In addition to direct biophysical channels, the interaction also raises issues for capital management, longevity hedging, and product design, since a mis-specification of climate-sensitive mortality trends can propagate into premium adequacy, reserving, and solvency assessments over multi-decade horizons. Belgium’s demographic structure – characterized by population ageing and significant urban concentration – makes this question particularly salient, given the known sensitivity of older age groups to heat and the potential amplification by urban heat island effects.

Indeed, the question of the impact of climate change on mortality risk in the insurance sector is central to many concerns, as it is a topical issue and the effects of climate change are expected to become increasingly pronounced given the various projections synthesized by

climatologists in different RCP (Representative Concentration Pathways) scenarios, some of which are dramatic, as explained by Jubb et al. (2013). Numerous studies, such as the work of Guibert et al. (2024), attempt to incorporate the effects of climate change into existing models that focus on past trends or events in order to produce stochastic mortality projections. Other authors, such as Boumezoued et al. (2022) and Milhaud and Gouthon (2024), have also sought to incorporate these climate-related changes into classical longevity models known in the literature, using French data, and have obtained rather encouraging results. This leads us to attempt a similar study on data from the Belgian population in order to quantify the impact that global warming will have on Belgian mortality by 2050 and the implications for insurers covering policyholders living in Belgium. In doing so, we will clarify the mapping from RCP narratives to stochastic paths of climate covariates used in the mortality models. We will also distinguish, where the data allow, between heat-related and cold-related contributions to excess mortality so as to avoid attributing all climatic effects to a single tail of the temperature distribution.

Our study will seek to determine whether global warming has a non-negligible impact on mortality in Belgium and, if so, to quantify it. We will also examine whether the models introduced by various authors for populations other than Belgium are suitable for Belgium and to what extent. We further aim to quantify this impact of climate change in terms of variations in the premiums of a basic life insurance product, in order to assess the real implications for insurance companies. Most of the results presented stem from analyses and linkages between mortality data and climate data available online for Belgium. The fitting of existing models and their extensions is performed numerically using the open-source programming language R, which is widely used as a reference in actuarial science and statistics and has extensive online documentation.

To best address the problem of assessing the impact of climate change on mortality risk, we begin by reviewing the foundations of stochastic mortality modeling. We provide the specifications and general principles of well-known longevity models before analyzing two extensions proposed by Boumezoued et al. (2022) and Milhaud and Gouthon (2024) to capture the effects of climate change. We then delve more deeply into the first extension of Boumezoued et al. (2022), following the authors' methodology to fit the proposed model to available Belgian data before comparing the performance of this model with the classical RD Lee and LR Carter (1992) model. We subsequently follow the same steps for the model proposed by Milhaud and Gouthon (2024). Finally, we propose an extension of the models studied in the previous sections in order to retain the best characteristics of each. This extension will serve to produce projections for constructing prospective mortality tables and to use these results to measure the impact of climate change on longevity. In parallel, we will translate model outputs into directly interpretable insurance metrics (changes in life expectancy, premium deltas) and discuss their operational use in pricing, reserving, and risk monitoring for Belgian portfolios.

Chapter 1

Literature Review

In this thesis, we compare two existing prospective modeling approaches for mortality that incorporate climate change data, in order to capture the impact of climate variations on mortality. The work of Boumezoued et al. (2022) and Milhaud and Gouthon (2024) suggests relying on well-known stochastic models, which we adapt by adding terms that isolate these causes. We will therefore first review the existing stochastic mortality rate models in the literature before delving deeper into the methodologies of the models proposed by Boumezoued et al. (2022) and Milhaud and Gouthon (2024).

1.1 Stochastic Modeling of Mortality Rates

1.1.1 GAPC Stochastic Mortality Models Family

Before delving into the specifications of the models under study, we first recall the foundations of stochastic mortality rates modeling. One of the most widely used models for mortality projections is the model introduced by RD Lee and LR Carter (1992). This model allows for the decomposition of a matrix of observed mortality rates – with ages as rows and time as columns – into a bilinear combination of age-related and period-related parameters. Subsequently, other authors have extended this model by incorporating, for example, quadratic terms and cohort effects. We focus on the work of Villegas et al. (2018) that groups these various extensions into a family of Generalized Age-Period-Cohort (GAPC) stochastic mortality models. We now introduce the required notation and the formal expression of this family.

First, we denote by D_{xt} the random variable representing the number of deaths in a population at age $x = x_1, \dots, x_k$, for individuals whose last birthday occurred during calendar year $t = t_1, \dots, t_n$. We also denote by d_{xt} and E_{xt}^c the observed number of deaths and the central exposure to risk at age x during year t , respectively. Regarding mortality rates, we denote the force of mortality and the central mortality rate by μ_{xt} and m_{xt} , respectively. Throughout this thesis, we assume that the force of mortality is piecewise constant, which implies that the force of mortality, μ_{xt} , and the central mortality rate, m_{xt} , coincide and thus we have $\mu_{xt} \approx d_{xt}/E_{xt}^c$. As explained by Villegas et al. (2018), a GAPC stochastic mortality model consists of four components:

1. The *random component*: the number of deaths, D_{xt} , which in this thesis is assumed to

follow a Poisson distribution as follows:

$$D_{xt} \sim \text{Poisson}(E_{xt}^c \mu_{xt}) \quad (1.1.1)$$

with $\mathbb{E}(D_{xt}/E_{xt}^c) = \mu_{xt}$.

2. The *systematic component*: the effects of age x , calendar year t , and year of birth (cohort) $c = t - x$ are captured by a *predictor*, η_{xt} , given by:

$$\eta_{xt} = \alpha_x + \sum_{i=1}^N \beta_x^{(i)} \kappa_t^{(i)} + \beta_x^{(0)} \gamma_{t-x} \quad (1.1.2)$$

with

- α_x as a *static age function* capturing the general trend of mortality by age,
- $N \geq 0$ as an integer indicating the number of age-period terms used to describe mortality trends. Each time index $\kappa_t^{(i)}$, for $i = 1, \dots, N$, contributes to capturing the overall mortality dynamics, while the corresponding $\beta_x^{(i)}$ modulates its effect across ages,
- γ_{t-x} as a parameter accounting for the cohort effect and $\beta_x^{(0)}$ modulating its effect across ages.

Note that we assume period indexes, $\kappa_t^{(i)}$, $i = 1, \dots, N$, and cohort index, γ_{t-x} , to be stochastic processes in order to perform stochastic projection of GAPC models. The reader is referred to the work of Hainaut (2024c) for further information regarding stochastic processes.

3. The *link function*, which connects the random and systematic components, such that:

$$g\left(\mathbb{E}\left(\frac{D_{xt}}{E_{xt}^c}\right)\right) = \eta_{xt}. \quad (1.1.3)$$

We will pair the Poisson distribution with the log link function commonly referred to as the canonical link.

4. The *set of parameter constraints* that ensures the model is identifiable. These constraints are applied through a *constraint function*, v , which maps an arbitrary vector of parameters

$$\theta := \left(\alpha_x, \beta_x^{(1)}, \dots, \beta_x^{(N)}, \kappa_t^{(1)}, \dots, \kappa_t^{(N)}, \beta_x^{(0)}, \gamma_{t-x}\right) \quad (1.1.4)$$

into a vector of transformed parameters

$$v(\theta) = \tilde{\theta} = \left(\tilde{\alpha}_x, \tilde{\beta}_x^{(1)}, \dots, \tilde{\beta}_x^{(N)}, \tilde{\kappa}_t^{(1)}, \dots, \tilde{\kappa}_t^{(N)}, \tilde{\beta}_x^{(0)}, \tilde{\gamma}_{t-x}\right) \quad (1.1.5)$$

which satisfies the model constraints such that θ and $\tilde{\theta}$ result in the same η_{xt} .

Most stochastic mortality models found in the literature belong to this family. It is therefore useful to recall their structure and components, in order to subsequently develop new models within this framework that incorporate parameters accounting for climatic variations.

1.1.2 Lee-Carter as an Example of GAPC Model

Following the work of Villegas et al. (2018), the GAPC family can be illustrated by the well-known model introduced by RD Lee and LR Carter (1992). This parsimonious model serves as the foundation for the climate-related extensions proposed by Boumezoued et al. (2022). It is therefore useful to present its formulation here. Since we are working with the force of mortality, μ_{xt} , we consider the Lee-Carter model under a Poisson framework.

As previously explained, we implement the model such that the number of deaths (1.1.1) follows a Poisson distribution, and we therefore use the logarithmic link function (1.1.3) with respect to the force of mortality, μ_{xt} . We now need to define the structure of the predictor (1.1.2). The model proposed by RD Lee and LR Carter (1992) assumes a static age function, α_x , a single ($N = 1$) age-period term to capture the mortality trend over time, and no cohort effect. We thus have:

$$\eta_{xt} = \alpha_x + \beta_x^{(1)} \kappa_t^{(1)} \quad (1.1.6)$$

with the period index modeled using an ARIMA process in order to project mortality data. The work of Villegas et al. (2018) suggests using a random walk with drift as a reasonable fit, given by:

$$\kappa_t^{(1)} = \Delta + \kappa_{t-1}^{(1)} + \xi_t, \quad \xi_t \sim N(0, \sigma_\kappa^2) \quad i.i.d., \quad (1.1.7)$$

where Δ is the drift parameter and ξ_t is a Gaussian white noise process with variance σ_κ^2 .

The final step in model construction is to identify a set of parameter constraints, applied through a constraint function that transforms (1.1.4) into (1.1.5), ensuring the identifiability of the model. As proposed by Villegas et al. (2018), for arbitrary real constants c_1 and $c_2 \neq 0$, the parameters of equation (1.1.6) can be transformed such that:

$$\left(\alpha_x, \beta_x^{(1)}, \kappa_t^{(1)} \right) \rightarrow \left(\alpha_x + c_1 \beta_x^{(1)}, \frac{1}{c_2} \beta_x^{(1)}, c_2 (\kappa_t^{(1)} - c_1) \right), \quad (1.1.8)$$

leaving η_{xt} unchanged. RD Lee and LR Carter (1992) propose the following set of parameter constraints:

$$\sum_x \beta_x^{(1)} = 1, \quad \sum_t \kappa_t^{(1)} = 0, \quad (1.1.9)$$

which can be enforced by setting

$$c_1 = \frac{1}{n} \sum_t \kappa_t^{(1)}, \quad c_2 = \sum_x \beta_x^{(1)} \quad (1.1.10)$$

in transformation (1.1.8).

1.1.3 Model Fitting

The **StMoMo** package available in R provides a convenient way to fit a Lee-Carter model or some other members of the GAPC family to mortality data. We refer the reader to Sections 4 and 5 of Villegas et al. (2018) for further details on the R implementations of these models.

Nevertheless, we can outline the general parameter estimation procedure for a GAPC stochastic mortality model by adapting the development proposed by Hainaut (2024a). In a GAPC model, the force of mortality is driven by

$$\log(\mu_{xt}) = \alpha_x + \sum_{i=1}^N \beta_x^{(i)} \kappa_t^{(i)} + \beta_x^{(0)} \gamma_{t-x}, \quad (1.1.11)$$

under the assumption (1.1.1) regarding the number of deaths in the population. In order to obtain a unique set of parameters, we use the Singular Value Decomposition (SVD) estimator (Akritas and Malaschonok, 2004) for α_x ,

$$\hat{\alpha}_x = \frac{1}{n} \sum_t \log \hat{\mu}_{xt}, \quad (1.1.12)$$

and e.g. the constraints

$$\sum_x \beta_x^{(0)} = 1, \quad \sum_x \beta_x^{(i)} = 1, \quad \sum_t \kappa_t^{(i)} = 0 \quad \text{and} \quad \sum_{c=t_1-x_k}^{t_n-x_1} \gamma_c = 0, \quad i = 1, \dots, N, \quad (1.1.13)$$

where $\hat{\mu}_{xt} = d_{xt}/E_{xt}^c$. Let $\hat{d}_{xt}$ denote the estimator of d_{xt} ,

$$\begin{aligned} \hat{d}_{xt} &= \mathbb{E}(D_{xt}) \\ &= E_{xt}^c \exp \left(\hat{\alpha}_x + \sum_{i=1}^N \beta_x^{(i)} \kappa_t^{(i)} + \beta_x^{(0)} \gamma_{t-x} \right), \end{aligned} \quad (1.1.14)$$

the model is then calibrated by maximizing the Poisson log-likelihood

$$\log \mathcal{L}(\beta^{(i)}, \kappa^{(i)}, \beta^{(0)}, \gamma) \propto \sum_t \sum_x \left(d_{xt} \left(\hat{\alpha}_x + \sum_{i=1}^N \beta_x^{(i)} \kappa_t^{(i)} + \beta_x^{(0)} \gamma_{t-x} \right) - \hat{d}_{xt} \right). \quad (1.1.15)$$

To solve this system, we choose to set numerically the derivative of (1.1.15) with respect to $\beta_x^{(i)}, \kappa_t^{(i)}, \beta_x^{(0)}$ and γ_{t-x} to zero using a Newton-Raphson algorithm (Akram and Ann, 2015).

The Newton-Raphson iterative algorithm proceeds as follows. If the estimator of the number of deaths in the population at the j^{th} iteration is defined by

$$\hat{d}_{xt}^{(j)} = E_{xt}^c \exp \left(\hat{\alpha}_x + \sum_{i=1}^N \hat{\beta}_x^{(i)(j)} \hat{\kappa}_t^{(i)(j)} + \hat{\beta}_x^{(0)(j)} \hat{\gamma}_{t-x}^{(j)} \right), \quad (1.1.16)$$

with initial values $\hat{\beta}_x^{(i)(0)} = 1$, $\hat{\kappa}_t^{(i)(0)} = 0$, $\hat{\beta}_x^{(0)(0)} = 1$, and $\hat{\gamma}_{t-x}^{(0)} = 0$ we define the following

updated values for iteration $j + 1$:

$$\begin{aligned}
 \hat{\gamma}_t^{(i)(j+1)} &= \hat{\gamma}_t^{(i)(j)} + \frac{\sum_x (d_{xt} - \hat{d}_{xt}^{(j)}) \hat{\beta}_x^{(0)(j)}}{\sum_x \hat{d}_{xt}^{(j)} (\hat{\beta}_x^{(0)(j)})^2}, \\
 \hat{\beta}_x^{(0)(j+1)} &= \hat{\beta}_x^{(0)(j)} + \frac{\sum_t (d_{xt} - \hat{d}_{xt}^{(j)}) \hat{\gamma}_t^{(j+1)}}{\sum_t \hat{d}_{xt}^{(j)} (\hat{\gamma}_t^{(j+1)})^2}, \\
 \hat{\kappa}_t^{(i)(j+1)} &= \hat{\kappa}_t^{(i)(j)} + \frac{\sum_x (d_{xt} - \hat{d}_{xt}^{(j)}) \hat{\beta}_x^{(i)(j)}}{\sum_x \hat{d}_{xt}^{(j)} (\hat{\beta}_x^{(i)(j)})^2}, \quad i = 1, \dots, N, \\
 \hat{\beta}_x^{(i)(j+1)} &= \hat{\beta}_x^{(i)(j)} + \frac{\sum_t (d_{xt} - \hat{d}_{xt}^{(j)}) \hat{\kappa}_t^{(i)(j+1)}}{\sum_t \hat{d}_{xt}^{(j)} (\hat{\kappa}_t^{(i)(j+1)})^2}, \quad i = 1, \dots, N.
 \end{aligned} \tag{1.1.17}$$

The algorithm is stopped when the relative increase in the Poisson log-likelihood (1.1.15) becomes smaller than a pre-specified tolerance threshold. Finally, let us define $\bar{\kappa}^{(i)} = \frac{1}{n} \sum_t \hat{\kappa}_t^{(i)}$ and $\bar{\gamma} = \frac{1}{n+k-1} \sum_{c=t_1-x_k}^{t_n-x_1} \hat{\gamma}_c$. Our constraints estimators can then be obtained as follows:

$$\begin{aligned}
 \tilde{\alpha}_x &= \hat{\alpha}_x + \sum_{i=1}^N \hat{\beta}_x^{(i)} \bar{\kappa}^{(i)} + \hat{\beta}_x^{(0)} \bar{\gamma}, \\
 \tilde{\beta}_x^{(i)} &= \frac{\hat{\beta}_x^{(i)}}{\sum_x \hat{\beta}_x^{(i)}}, \quad i = 1, \dots, N, \\
 \tilde{\kappa}_t^{(i)} &= \left(\hat{\kappa}_t^{(i)} - \bar{\kappa}^{(i)} \right) \sum_x \hat{\beta}_x^{(i)}, \quad i = 1, \dots, N, \\
 \tilde{\beta}_x^{(0)} &= \frac{\hat{\beta}_x^{(0)}}{\sum_x \hat{\beta}_x^{(0)}}, \\
 \tilde{\gamma}_{t-x} &= (\hat{\gamma}_{t-x} - \bar{\gamma}) \sum_x \hat{\beta}_x^{(0)}.
 \end{aligned} \tag{1.1.18}$$

Once the model parameters have been estimated, it remains to fit ARIMA-type time series processes to the period indexes $\tilde{\kappa}_t^{(i)}, i = 1, \dots, N$ and the cohort index $\tilde{\gamma}_{t-x}$ in order to project mortality rates. A common approach in the literature is to use a random walk with drift; however, we refer the reader to Section 7 of Villegas et al. (2018) for further details.

1.2 Boumezoued et al.'s Model

The study by Boumezoued et al. (2022) aims to adapt the classical Lee-Carter model by adding a specific term designed to isolate excess mortality attributable to climatic causes on French and American data. Climatic data are incorporated through the construction of a climate index intended to capture exposure to extreme temperatures over the years. Periods of submortality are also accounted for in the model by explicitly modeling the *harvesting effect*. The population is divided into three age groups in order to accurately assess the impact of

heatwaves and extreme cold across different segments of the population. The model is first calibrated separately for each of the three age classes before reverting to continuous age for projection purposes.

1.2.1 Classical Lee-Carter Model with Construction of a Climate Index

Boumezoued et al. (2022) propose an adaptation of the RD Lee and LR Carter (1992) model (1.1.6) by adding a term that isolates mortality attributable to climatic causes. Still under a Poisson framework (1.1.1), we consider the following model:

$$\log(\mu_{xt}) = \alpha_x + \beta_x^{(o)} \kappa_t^{(o)} + \delta_x^{(c)} C_t, \quad (1.2.1)$$

where:

- α_x plays the same role as in the classical Lee-Carter model, namely capturing the general age-specific mortality pattern,
- $\beta_x^{(o)} \kappa_t^{(o)}$ aims to capture overall mortality not attributable to climatic conditions, with o standing for *other*, $\kappa_t^{(o)}$ as the period index representing general mortality dynamics, and $\beta_x^{(o)}$ modulating its effect across ages,
- $\delta_x^{(c)} C_t$ captures mortality directly related to climatic causes, with c for *climatic*, C_t being the climate index tracking temperature evolution over the years, and $\delta_x^{(c)}$ acting as an age-specific modulator.

The climate index C_t is constructed using a linear regression on a vector X_t of climate variables (e.g., the number of summer days during which the temperature exceeded 30°C) for year t , such that

$$C_t = a + b^T X_t, \quad (1.2.2)$$

where a and b are the parameters of the linear regression. The climate variables composing the vector X_t are selected based on an AIC criterion (Kuha, 2004) in order to identify the variables that best explain mortality rates associated with exposure to extreme environmental temperatures. Boumezoued et al. (2022) suggest using the **stepAIC** procedure available in the R programming language, which tests all possible combinations among the available climate variables and selects the combination yielding the lowest AIC. They then recommend validating the significance of the selected variables by conducting a p-value analysis, in order to retain only the most significant ones. This approach ensures a parsimonious model and helps limit overfitting in projection exercises.

1.2.2 Harvesting Effect and Age Range Constitution

When attempting to isolate excess mortality attributable to climatic causes, we are confronted with the harvesting effect as explained by Toulemon and Barbieri (2008). Indeed, the most vulnerable individuals in the population are more exposed to exceptional mortality events, which implies a period of excess mortality followed by a period of reduced mortality, as these individuals would have died in the days or weeks following the event regardless. To model this effect, Boumezoued et al. (2022) propose adjusting our climate index (1.2.2) by introducing

two parameters that account for both the current and previous year of observation. The climate index incorporating the harvesting effect therefore takes the following form:

$$CI_t = a + \pi(b^T X_t) + \rho(b^T X_{t-1}), \quad (1.2.3)$$

where π represents the proportion of the immediate impact of the climatic cause, and ρ the proportion attributable to the harvesting effect.

The final step proposed by Boumezoued et al. (2022) before proceeding to model calibration consists in dividing the population into three age groups, denoted by c_i , $i = 0, \dots, 2$, in order to reflect the impact of climate on mortality within these subgroups. Indeed, extreme temperatures do not affect all age groups in the population in the same way, making this subdivision necessary to properly calibrate the final model. These three groups remain broad enough to ensure a sufficient number of deaths in each category and are defined based on data availability.

1.2.3 Model Fitting

The calibration of model (1.2.1) is performed in three steps. We follow the methodology outlined in Boumezoued et al. (2022) as follows:

1. Estimation of climate-related mortality across predefined age groups:

$$\log(\mu_{c_i,t}) = \alpha_{c_i} + \beta_{c_i}^{(o)} \kappa_t^{(o)} + \delta_{c_i}^{(c)} C_t \quad (1.2.4)$$

- (a) Calibration of the three parameter vectors for the three age classes, α_{c_i} , is carried out using a Lee-Carter model fitted to available mortality data.
- (b) Calibration of C_t , the climate index defined in (1.2.2), is performed using linear regression between the mortality rate associated with climate-related causes and climate variables, following the method described in Section 1.2.1.
- (c) Calibration of the three age-specific climate modulators $\delta_{c_i}^{(c)}$ is done by minimizing the residuals without incorporating the harvesting effect:

$$R_{c_i,t} = \log(\mu_{c_i,t}) - \alpha_{c_i} - \delta_{c_i}^{(c)} C_t. \quad (1.2.5)$$

- 2. Integration of the harvesting effect.** This step consists in replacing C_t with CI_t defined in (1.2.3) by calibrating π and ρ . The harvesting parameters are obtained by minimizing the previous residuals (1.2.5) now including the harvesting effect:

$$R_{c_i,t} = \log(\mu_{c_i,t}) - \alpha_{c_i} - \delta_{c_i}^{(c)} (a + \pi(b^T X_t) + \rho(b^T X_{t-1})). \quad (1.2.6)$$

- 3. Calibration of the full model with continuous age.** The final step aims to convert all age-class parameters into age-continuous parameters to yield the final model equation:

$$\log(\mu_{xt}) = \alpha_x + \beta_x^{(o)} \kappa_t^{(o)} + \delta_x^{(c)} CI_t. \quad (1.2.7)$$

The final calibration involves applying a standard Lee-Carter model to the residuals

$$R_{xt} = \log(\mu_{xt}) - \delta_x^{(c)} CI_t \quad (1.2.8)$$

to estimate the parameters $\alpha_x, \beta_x^{(o)}$ and $\kappa_t^{(o)}$.

The calibration of this model using Belgian data, along with the analysis of the resulting outcomes, will be addressed in Chapter 2.

1.3 Milhaud and Gouthon's Model

The work of Milhaud and Gouthon (2024) aims to adapt an existing longevity model that incorporates climatic impacts by introducing a trigger event corresponding to situations of high temperatures on French data. This addition is intended to separate the effects of climate from other contributing factors. The trigger event is defined as the case where the average summer temperature for a given year exceeds the average summer temperature over the entire observation period. In order to isolate climatic effects, an age modulator is introduced based on a threshold age, which corresponds to the age beyond which high temperatures have harmful impacts on human health. This threshold age is determined through a linear regression between weekly temperature and weekly mortality rate; if the test is statistically significant, the corresponding age is retained. The resulting model can then be estimated using maximum likelihood methods, as described in Section 1.2.1.

1.3.1 SPO Model with Triggering Event

The model proposed by Milhaud and Gouthon (2024) is an adaptation of the existing SPO model introduced by Seklecka et al. (2017), which itself is an adaptation of the well-known model proposed by Plat (2009) in the longevity modeling literature. In all cases, these models belong to the family of GAPC models. We therefore adopt the notation and working framework described in Section 1.1. Let us begin by describing the SPO model in order to understand the subsequent extensions made to it. The evolution of the mortality rate is modeled as follows:

$$\begin{aligned} \log(\mu_{xt}) = & \alpha_x + \kappa_t^{(1)} + (\bar{x} - x)\kappa_t^{(2)} + (\bar{x} - x)^+\kappa_t^{(3)} \\ & + ((A - x)_+ + w_x(x - A)^+)^2 \kappa_t^{(4)} + \gamma_{t-x}, \end{aligned} \quad (1.3.1)$$

where α_x and γ_{t-x} are the static age function capturing the general trend of mortality by age and the parameter accounting for the cohort effect, as described in (1.1.2). The model also includes four period indices, denoted $\kappa_t^{(i)}$, $i = 1, \dots, 4$, which aim to capture trends in mortality over time, potentially modulated by other effects such as climatic impact. The first three terms are already present in the model of Plat (2009), and represent respectively the overall trend in mortality rates over time, the evolution modulated by age, and the evolution specifically modulated for younger ages. Seklecka et al. (2017) added $\kappa_t^{(4)}$, modulated by a quadratic term, to capture the non-linearity of mortality rates, with modulation based on an age threshold A . Below this threshold, only the quadratic age-related deviation is taken into account, while above it, a coefficient w_x , estimated outside the model, accounts for the Pearson correlation between temperature and mortality.

Milhaud and Gouthon (2024) highlight the following issues: in this formulation, $\kappa_t^{(4)}$ is designed to capture both the non-linear age-related effects and the temperature-related effects, which makes its interpretation difficult. Moreover, it is challenging to perform projections under different climate scenarios when the corresponding term does not isolate this specific cause. For this reason, the authors propose introducing a triggering event corresponding to situations in which temperatures exceed the average, thereby separating climatic effects from

other factors. The evolution of the mortality rate is then given by:

$$\begin{aligned} \log(\mu_{xt}) = & \alpha_x + \kappa_t^{(1)} + (\bar{x} - x)\kappa_t^{(2)} + (\bar{x} - x)^+\kappa_t^{(3)} \\ & + \left((x_1 - x)_+ + w(x - x_2)^+\right)^2 \kappa_t^{(4)} \\ & + (x - A_h)^+ \mathbb{1}_{\{I_t - \bar{I}_t > 0\}} \kappa_t^{(5)} + \gamma_{t-x}, \end{aligned} \quad (1.3.2)$$

where $\kappa_t^{(4)}$ now captures solely the non-linear effects in the mortality trend, and $\kappa_t^{(5)}$ describes the year-over-year evolution of mortality associated with climatic effects. Indeed, the indicator function $\mathbb{1}_{\{I_t - \bar{I}_t > 0\}}$ ensures that only years in which summers were warmer than average are considered, with I_t denoting the average summer temperature for year t , and $\bar{I}_t$ the average summer temperature over the entire observation period. Finally, the hyperparameters x_1 , x_2 , and A_h represent threshold ages estimated outside the model, with A_h in particular denoting the age at which humans become vulnerable to high temperatures. The construction of this threshold is detailed in Section 1.3.2. The estimation of the hyperparameters x_1 , x_2 , and w , a weight between 0 and 1, will be described in Section 1.3.3.

1.3.2 Age Threshold Construction

The model developed by Milhaud and Gouthon (2024) incorporates a hyperparameter, A_h , corresponding to the age from which extreme temperatures begin to impact human health. Since this model considers only the effects of high temperatures the proposed framework establishes a linear regression between the weekly temperature and the mortality observed during the same week of the summer period (June to August). This regression is made possible by the data available from the *Human Mortality Database* (HMD), in particular the *Short Term Mortality Fluctuation* (STMF) dataset, which provides weekly death counts and mortality rates for numerous countries from 2000 to 2025, as described in Section 3.1.1. More details on the data used are provided in our application to Belgian data in Chapters 2 and 3.

The STMF data are divided into five age groups: 0–14, 15–64, 65–74, 75–84, and 85+ years. These are the age groups we use for the linear regression, and thus to define the age threshold A_h . If we index each group by k , Milhaud and Gouthon (2024) propose defining the points $(d_j, \theta_j)_k$, where d_j and θ_j represent respectively the number of deaths and the average temperature recorded during the j^{th} week of the summer. This setup is used to perform a Pearson (1985) correlation test to identify the age group from which high temperatures begin to have an impact on longevity and thus set A_h .

1.3.3 Model Fitting

Milhaud and Gouthon (2024) propose focusing on data from ages 20 to 85+, where the 85+ age group represents all individuals aged 85 and above, and on the time period from 1980 to 2019. The authors also suggest dividing this period into a training period and a validation period, the latter spanning from 2012 to 2019. However, it is worth noting that the years 2012 to 2019 are among the warmer years; excluding them from the training phase could therefore lead to biased projections for future years.

The parameters of model (1.3.2) are estimated by maximizing the log-likelihood given in (1.1.15), as explained in Section 1.1. To address the identifiability issue inherent in the model,

Milhaud and Gouthon (2024) propose the following reparameterization:

$$\begin{cases} \tilde{\gamma}_{t-x} = \gamma_{t-x} + \psi_1 + (t-x)\psi_2, \\ \tilde{\kappa}_t^{(1)} = \kappa_t^{(1)} - \psi_1 - (d\bar{x} + t)\psi_2, \\ \tilde{\kappa}_t^{(2)} = \kappa_t^{(2)} - d\psi_2, \\ \tilde{\alpha}_x = \alpha_x + (1+d)x\psi_2, \end{cases} \quad (1.3.3)$$

where ψ_1 , ψ_2 , and d are constants. This issue is resolved by imposing the following identifiability constraints:

$$\sum_c \gamma_c = 0, \quad \sum_c c\gamma_c = 0, \quad \sum_t \kappa_t^{(3)} = 0 \quad \text{and} \quad \sum_t \kappa_t^{(4)} = 0. \quad (1.3.4)$$

Finally, Milhaud and Gouthon (2024) propose determining the hyperparameters w, x_1 and x_2 by minimizing the prediction error on the validation sample. Moreover, since non-linearity is more pronounced at younger ages and tends to evolve around age 50, the values are restricted to narrow intervals: $x_1 \in [45, 52]$, $x_2 \in [52, 60]$ and $w \in [0, 1]$.

The calibration of this model using Belgian data, along with the analysis of the resulting outcomes, will be addressed in Chapter 3.

Chapter 2

Boumezoued et al.’s Approach on Belgian Data

In this chapter, we will employ the methodology presented by Boumezoued et al. (2022), developed in Section 1.2, with the aim of integrating the impact of climate change into the modeling of Belgian mortality. We will first present the available data and how these will be used before constructing our climate index. Next, we will calibrate the modified Lee–Carter model on these data and conclude with a comparison of the performance against the classical model that does not isolate climate-related causes.

2.1 Available Belgian Data

In order to study the impact of heat waves on Belgian mortality data, we will require three complementary datasets. Indeed, in their study on French data, Boumezoued et al. (2022) propose using a dataset containing mortality data specific to climate-related causes, one containing the country’s overall mortality data, and, finally, one containing meteorological data recorded over the study period. We will proceed in the same manner for the Belgian data and present the following datasets.

2.1.1 Belgian Specific Mortality (Global Health Data, GHD)

The Global Health Data Exchange¹ is a database published by the Institute for Health Metrics and Evaluation (IHME), presenting the results from the 2021 Global Burden of Disease (GBD) study. This study provides estimates of mortality data by country, age group, sex, and cause of death from 1980 to 2021. The advantage of this database is that it allows us to isolate mortality data for specific causes. In our case, for example, we have access to the cause *Exposure to environmental heat and cold*, which is precisely what we aim to isolate in our study of the impact of climate change on mortality. Here, we choose to focus on data from 1980 to 2018 since data beyond 2018 do not appear reliable for Belgium; we explain this choice in more detail in the Section 2.2.2. We can graphically represent the selected data in Figure 2.1.1, on which the heat waves of July 2006 and August 2003 are clearly visible as peaks in mortality.

¹This database is accessible via <https://ghdx.healthdata.org/> and <https://vizhub.healthdata.org/gbd-results/>.

Extreme Environmental Temperature Exposure Death Rate

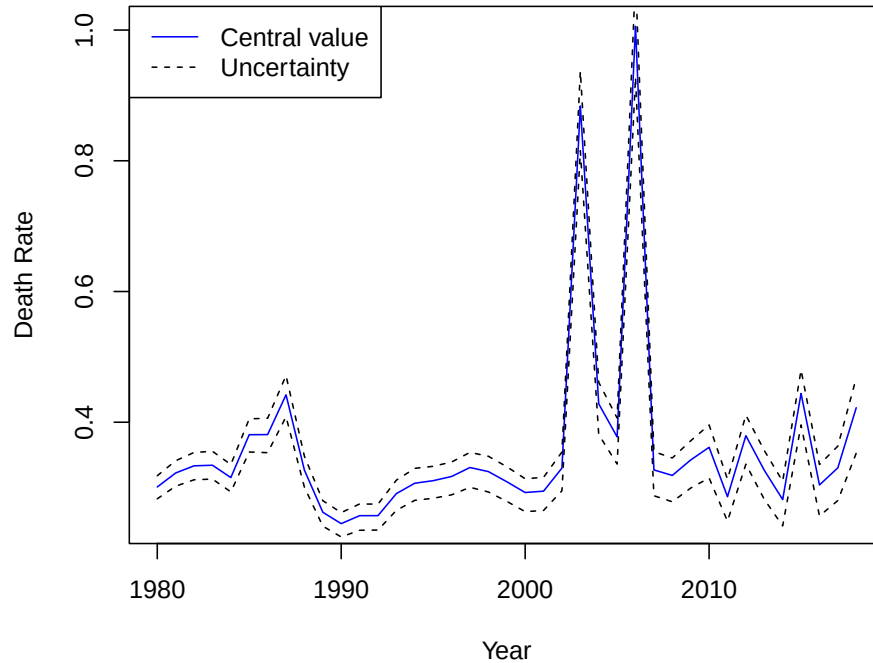


Figure 2.1.1: Belgian Death Rate Attributable to Extreme Environmental Temperature Exposure (both sexes, all ages, per 100,000 population).

2.1.2 Belgian Global Mortality (Human Mortality Database, HMD)

The Human Mortality Database (HMD) describes itself as the world’s leading scientific data resource on mortality in developed countries. The available life tables² provide all necessary information and serve as the standard reference within the actuarial community for analyzing the evolution of a country’s overall mortality. Moreover, the HMD format facilitates the practical use of the *StMoMo* package in R. We will employ the life tables for the period 1980–2018 across the full available age range.

2.1.3 Belgian Meteorological Data (Royal Meteorological Institute, RMI)

The Royal Meteorological Institute (RMI) offers daily datasets³ for each location in Belgium, segmented into *pixels*, containing various climate variables such as minimum, maximum, and mean temperatures, precipitation amounts, relative humidity, sunshine duration, and so forth. We will use these data to calculate national monthly averages of the provided climate variables and to create new ones for the period 1980–2018 in order to construct our climate index, as explained below.

²Belgian mortality data are accessible via <https://www.mortality.org/> and <https://www.mortality.org/Country/Country?cntr=BEL>.

³These datasets are accessible via the RMI’s new online platform <https://opendata.meteo.be/>.

2.2 Construction of a Climate Index

The aim of constructing a climate index is to capture the effect of exposure to extreme environmental temperatures on mortality. We will base our approach on the methodology of Boumezoued et al. (2022) explained previously in Section 1.2.1 for its construction.

2.2.1 Climate Variables

We begin by defining the climate variables that we will use in our analysis. The large volume of climate data provided by the RMI, detailed in Section 2.1.3, offers considerable flexibility in creating climate variables. In order to capture the effects of heat waves on mortality, we construct six climate variables over the four warmest months in Belgium (June, July, August, and September). We retain the following variables:

- $T_t^{\text{Max},s}$: the daily maximum temperature during summertime of year t ,
- $T_t^{\text{Min},s}$: the daily minimum temperature during summertime of year t ,
- $T_t^{\text{Avg},s}$: the daily average temperature during summertime of year t ,
- T_t^{30} : the average number of days with temperature exceeding 30 °C during summertime of year t ,
- T_t^{34} : the average number of days with temperature exceeding 34 °C during summertime of year t ,
- T_t^{38} : the average number of days with temperature exceeding 38 °C during summertime of year t .

As an extension of the methodology of Boumezoued et al. (2022), we also propose to capture the effects of cold spells by constructing six climate variables over the four coldest months in Belgium (November, December, January, and February). We retain the following variables:

- $T_t^{\text{Max},w}$: the daily maximum temperature during winter of year t ,
- $T_t^{\text{Min},w}$: the daily minimum temperature during winter of year t ,
- $T_t^{\text{Avg},w}$: the daily average temperature during winter of year t ,
- T_t^0 : the average number of days with temperature below 0 °C during winter of year t ,
- T_t^{-5} : the average number of days with temperature below -5 °C during winter of year t ,
- T_t^{-10} : the average number of days with temperature below -10 °C during winter of year t .

We thus obtain twelve climate variables aimed at explaining cause-specific mortality due to *Exposure to environmental temperatures*, for which we have data from 1980 to 2021 as explained in Section 2.1.1.

2.2.2 Period of Calibration

We decided to calibrate our climate index, and hence our model, over the period 1980–2018. This choice is motivated by the fact that mortality data attributed to exposure to extreme environmental temperatures appear to be biased for the year 2020 in Belgium.

The mortality rate peak in Belgium for 2020 reaches nearly 14, as shown in Figure 2.2.1, rendering all other years negligible for this cause of death. Yet it is well known that the 2003 European heatwave had a significant impact on Belgian mortality. To substantiate our concerns regarding data reliability, we plotted in Figures 2.2.1 and 2.2.2 the mortality rates attributable to the same cause for countries bordering Belgium, in order to observe the regional trend. As shown in Figure 2.2.2, Belgian mortality rates align with those observed in neighbouring countries for the years 1980 to 2018. However, Figure 2.2.1 clearly demonstrates that the 2020 mortality peak in Belgium vastly exceeds the trend of neighbouring countries, even though they share a relatively similar climate. Specifically, the peaks observed in 2020 do not exceed 4 in neighbouring countries, whereas the Belgian peak approaches 14 – approximately 3.5 times higher.

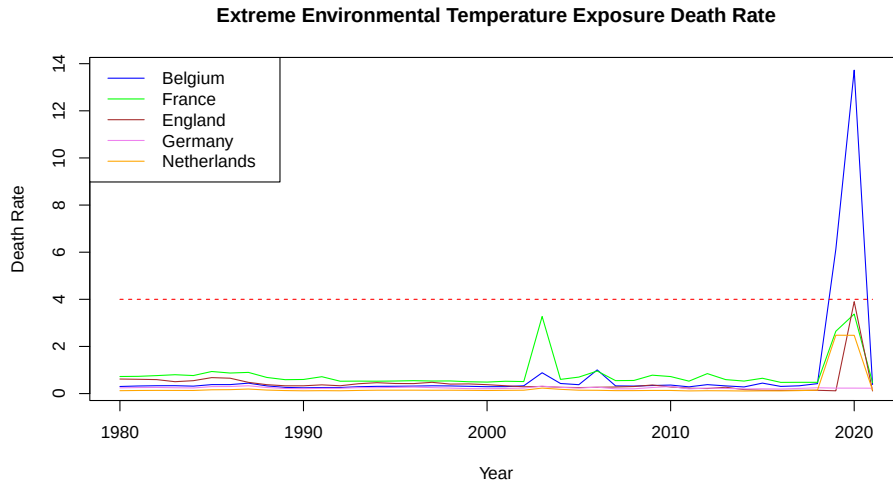


Figure 2.2.1: European Death Rate Attributable to Extreme Environmental Temperature Exposure Until 2021 (both sexes, all ages, per 100,000 population)

One possible explanation is that the 2020 Belgian data were biased due to excess mortality related to COVID-19, and that the small size of the country may have led to confusion in the reported figures. Based on these observations, we therefore chose to calibrate our model on the data that appear most reliable – namely the years 1980 to 2018 – despite excluding two notably hot years: 2019 and 2020.

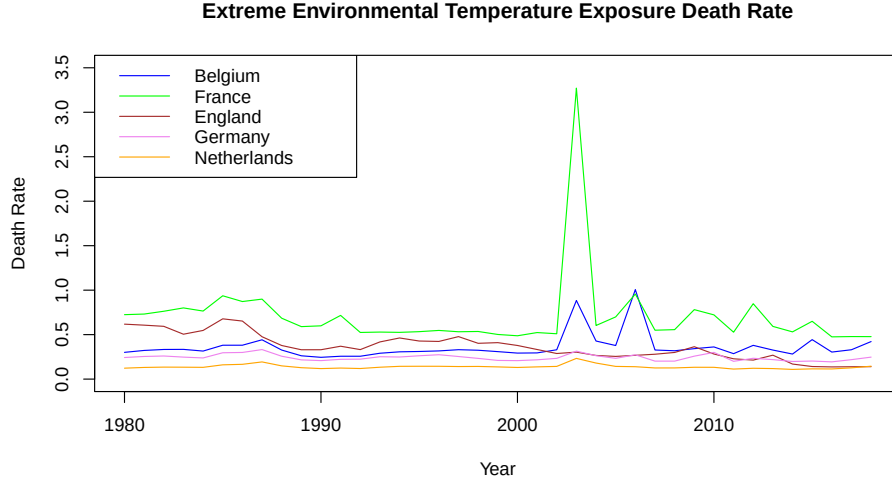


Figure 2.2.2: European Death Rate Attributable to Extreme Environmental Temperature Exposure Until 2018 (both sexes, all ages, per 100,000 population)

2.2.3 Fitting Procedure

As explained previously in Section 1.2.1, the objective is to select the best climate variables among the twelve constructed so as to most effectively explain climate-related mortality. To achieve this, we apply the **stepAIC** procedure in R (Zhang, 2016) on a linear regression between our vector of climate variables and our cause-specific mortality data (GHD). The results of the best linear model found (in terms of AIC) by this procedure are presented in Table 2.2.1. The variable T_t^{30} is not significant at the 5% level and is therefore excluded from the final model.

Variable	Estimate	Std. Error	t value	Pr(> t)	Signif.
Intercept	-0.9106	0.3913	-2.327	0.0262	*
$T_t^{\text{Min},s}$	0.0980	0.0337	2.907	0.0065	**
T_t^{30}	-0.0135	0.0089	-1.511	0.1404	
T_t^{34}	0.1875	0.0413	4.544	0.0001	***
T_t^{38}	-43.4336	13.2532	-3.277	0.0025	**
T_t^0	0.0032	0.0014	2.276	0.0294	*

Table 2.2.1: Best Linear Regression Results (AIC).

As a further examination, we can show that extending the methodology of Boumezoued et al. (2022) by incorporating climate variables accounting for cold days is beneficial. Indeed, we also applied the aforementioned procedure to models including only heat-related variables, only cold-related variables, and all climate variables. We observe in Table 2.2.2 that the lowest AIC corresponds to the model incorporating all climate variables. Furthermore, the variable T_t^0 remains statistically significant, as presented previously in Table 2.2.1.

Model	AIC
Cold Climate Variables	-36.19
Heat Climate Variables	-55.39
All Climate Variables	-59.35

Table 2.2.2: Akaike Information Criterion (AIC) for the Three Best Linear Models.

2.2.4 Selected Variables and Final Climate Index

Based on the analysis in the previous section, we retain the following climate variables for constructing our climate index:

- $T_t^{\text{Min},s}$: the daily minimum temperature during summertime of year t ,
- T_t^{34} : the average number of days with temperature exceeding 34 °C during summertime of year t ,
- T_t^{38} : the average number of days with temperature exceeding 38 °C during summertime of year t ,
- T_t^0 : the average number of days with temperature below 0 °C during winter of year t .

The results of the resulting linear regression are reported in Table 2.2.3. We observe that each variable contributes significantly to explaining the mortality rate attributable to the climate-specific cause. The AIC of this model is -58.74 , which remains lower (i.e., better) than that of the best model including only heat-related climate variables, as recalled in Table 2.2.2.

Variable	Estimate	Std. Error	t value	Pr(> t)	Signif.
Intercept	-0.6196	0.3469	-2.086	0.0430	*
$T_t^{\text{Min},s}$	0.0697	0.0285	2.442	0.0200	*
T_t^{34}	0.1418	0.0286	4.956	<0.0001	***
T_t^{38}	-37.9419	12.9829	-2.922	0.0061	**
T_t^0	0.0029	0.0014	2.073	0.0458	*

Table 2.2.3: Linear Regression Results for Selected Climate Variables.

We can also graphically present these four selected climate variables in Figure 2.2.3 to illustrate the information they convey in explaining the mortality rate.

Finally, the climate index constructed using the linear regression between the four climate variables and the climate-related mortality rates can be written, using the notation in (1.2.2), as a linear equation with five parameters:

$$C_t = a + b_1 T_t^{\text{Min},s} + b_2 T_t^{34} + b_3 T_t^{38} + b_4 T_t^0, \quad (2.2.1)$$

where a , b_1 , b_2 , b_3 , and b_4 are the regression parameters obtained in Table 2.2.3. We can graphically represent the resulting climate index in Figure 2.2.4. We observe that the trend in

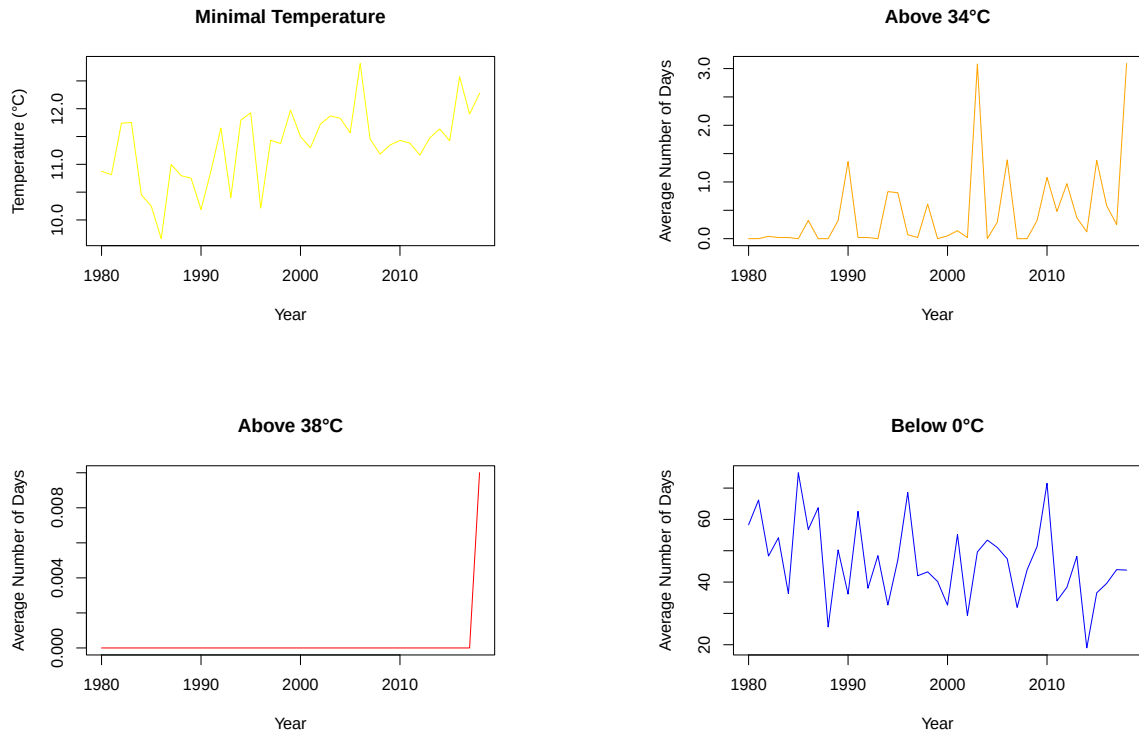


Figure 2.2.3: Evolution of the Selected Climate Variables.

climate-related mortality causes is captured rather well by our index. The fit is satisfactory, with a linear coefficient of determination (Ozer, 1985) $R^2 \approx 64.07\%$. Despite several attempts to integrate additional climate variables, we were unable to achieve a better fit for the Belgian data.

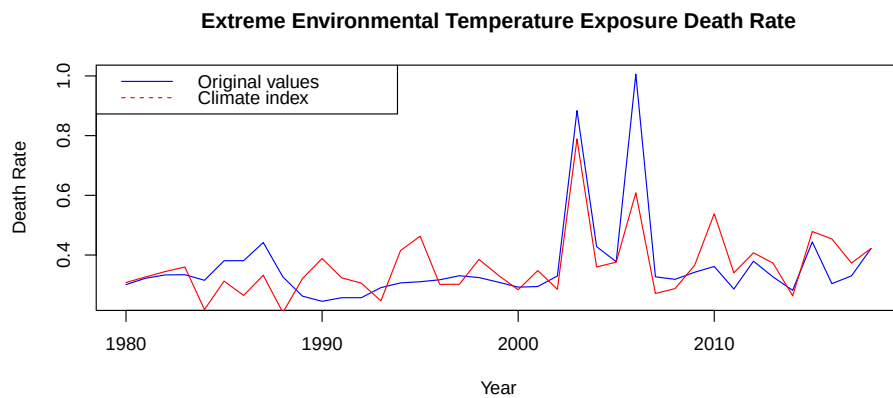


Figure 2.2.4: Climate Index and Original Death Rate in Belgium (both sexes, all ages, per 100,000 population).

2.3 Model Calibration

Our objective is to follow the methodology of Boumezoued et al. (2022) detailed in Section 1.2 to calibrate model (1.2.1), which is a modified RD Lee and LR Carter (1992) model integrating climate-related mortality causes via our climate index (2.2.1) constructed in the previous section. We will first estimate the model for three predefined age classes: 0–25 years (young people), 26–65 years (working-age population), and 65+ years (seniors), as suggested by the authors. Next, we will account for the harvesting effect, and finally we will calibrate the full model across all ages. For further details regarding the calibration steps, we refer the reader to Section 1.2.3.

2.3.1 Estimation of Climate-Related Mortality Across Age Groups

Using the notation from Section 1.2.3, we have the following relationships for our three age classes $c_0 = [0, 24]$, $c_1 = [25, 64]$, and $c_2 = [65, 105]$:

$$\log(\mu_{c_i,t}) = \alpha_{c_i} + \beta_{c_i}^{(o)} \kappa_t^{(o)} + \delta_{c_i}^{(c)} C_t. \quad (2.3.1)$$

We use the general mortality data from the HMD described in Section 2.1.2 for the period 1980–2018 to calibrate the three vectors α_{c_i} for each age class. To do so, we employ the **StMoMo** package in R, which fits a RD Lee and LR Carter (1992) model to each age class separately using the HMD data, as explained by Villegas et al. (2018). Note that similar results could have been obtained by applying the general fitting method described in Section 1.1.3. The estimated models therefore take the form

$$\log(\mu_{c_i,t}) = \alpha_{c_i} + \beta_{c_i} \kappa_t, \quad (2.3.2)$$

from which we retain only the estimates of α_{c_i} , since our goal is subsequently to decompose the term $\beta_{c_i} \kappa_t$ in order to isolate climate-related mortality from other causes. The resulting estimates for the three age classes are shown in Figure 2.3.1. The results display higher values for older ages as expected, along with the well-known pattern for class c_0 , notably the accident hump and the elevated values corresponding to infant mortality.

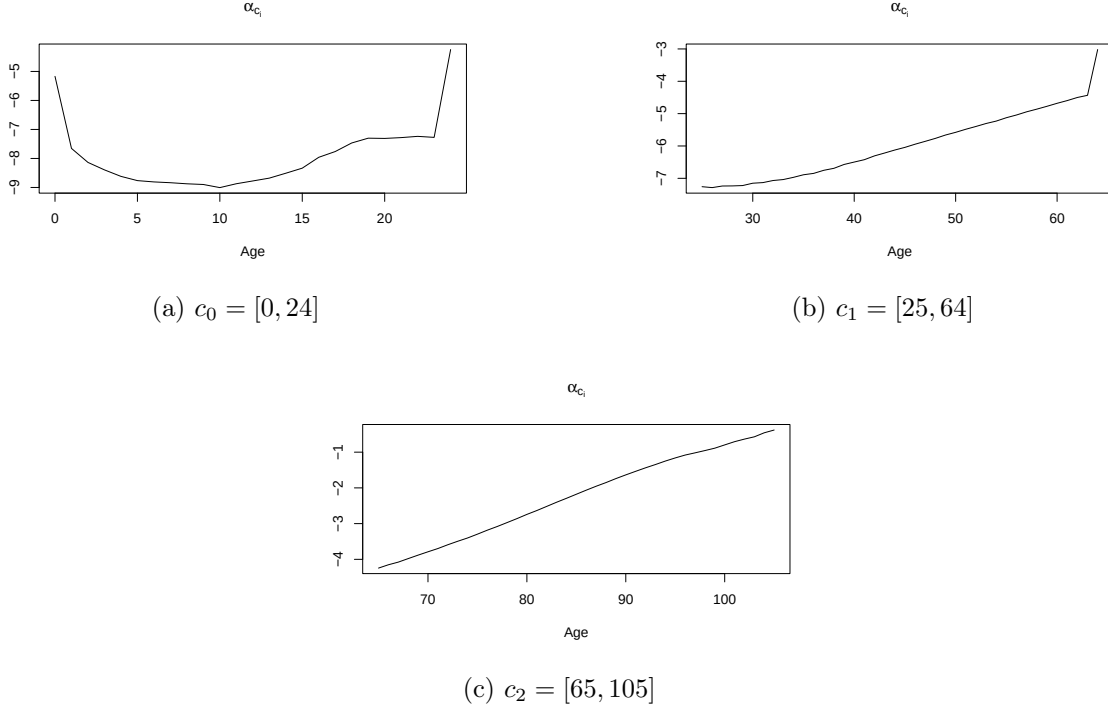
The calibration of the climate index C_t was already performed in Section 2.2 using model (2.2.1), and the results were presented in Figure 2.2.4.

As explained in Section 1.2.3, the calibration of the three age-specific climate modulators $\delta_{c_i}^{(c)}$ is performed by minimizing the residuals without incorporating the harvesting effect:

$$R_{c_i,t} = \log(\mu_{c_i,t}) - \alpha_{c_i} - \delta_{c_i}^{(c)} C_t. \quad (2.3.3)$$

To this end, we use the `optimize()` function in R, which applies a combination of golden section search and successive parabolic interpolation⁴. The results obtained for the three age classes are presented in Figure 2.3.2. The estimates remain relatively low and constant for the first two age groups, while they are higher and increasing for the third age group, indicating the presence of a pivot age beyond which exposure to heat significantly increases mortality risk.

⁴Some insights about this R function are available at <https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/optimize>.

Figure 2.3.1: Age-specific Mortality Patterns α_{c_i} for $i = 0, 1, 2$.

2.3.2 Integration of the Harvesting Effect

As recalled in Section 1.2.3, integrating the harvesting effect into the model calibration on our Belgian mortality data consists of replacing the previously calculated climate index, C_t , with CI_t , defined by

$$CI_t = a + \pi (b^T X_t) + \rho (b^T X_{t-1}), \quad (2.3.4)$$

and calibrating π and ρ . These parameters are obtained by minimizing the residuals in (2.3.3), now including the harvesting effect, i.e., by minimizing:

$$R_{c_i,t} = \log(\mu_{c_i,t}) - \alpha_{c_i} - \delta_{c_i}^{(c)} (a + \pi (b^T X_t) + \rho (b^T X_{t-1})). \quad (2.3.5)$$

Figure 2.3.3 overlays the resulting index with and without the harvesting effect, as previously shown in Figure 2.2.4. We observe that the climate index values accounting for the harvesting effect are slightly lower than those without it. This difference arises because, when omitting the harvesting effect, one obtains periods of excess mortality followed by periods of reduced mortality, as the frailest individuals would have died in the subsequent weeks regardless of the heat wave. Therefore, it is important to incorporate this effect into our climate index to achieve results more consistent with reality and avoid attributing excessive mortality to climatic causes.

2.3.3 Calibration of the Full Model with Continuous Age

The final step before obtaining our modified RD Lee and LR Carter (1992) model that accounts for climate-specific mortality causes is to move to continuous age. To this end, we convert all

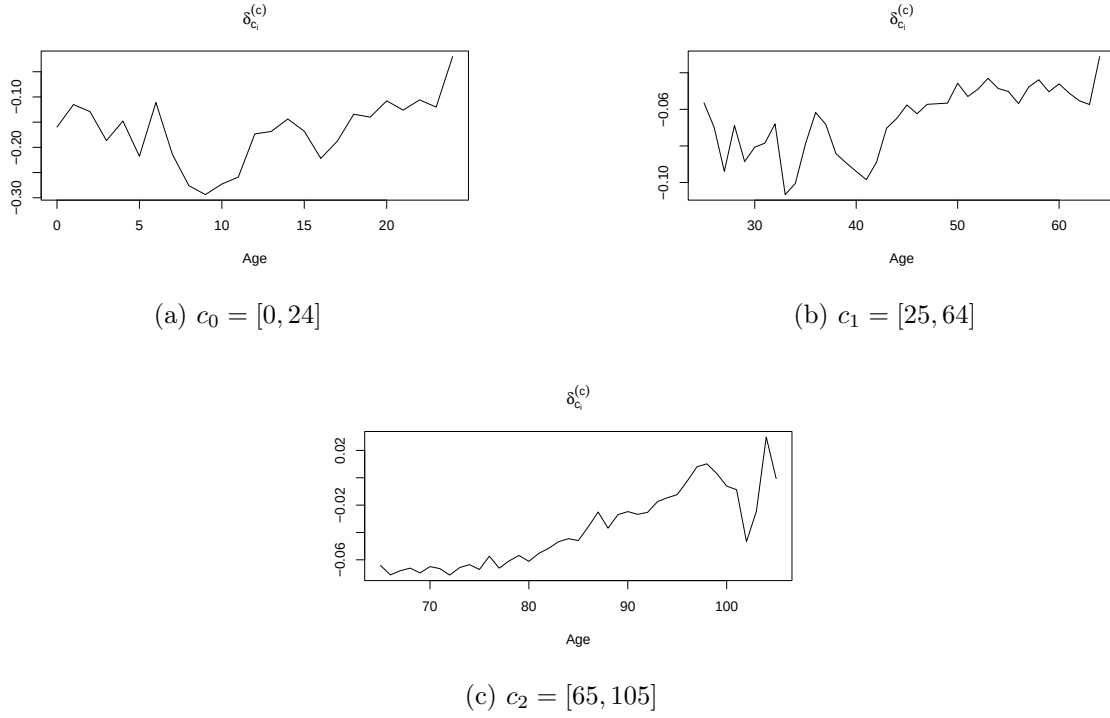


Figure 2.3.2: Age-specific Climate Modulators $\delta_{c_i}^{(c)}$ for $i = 0, 1, 2$.

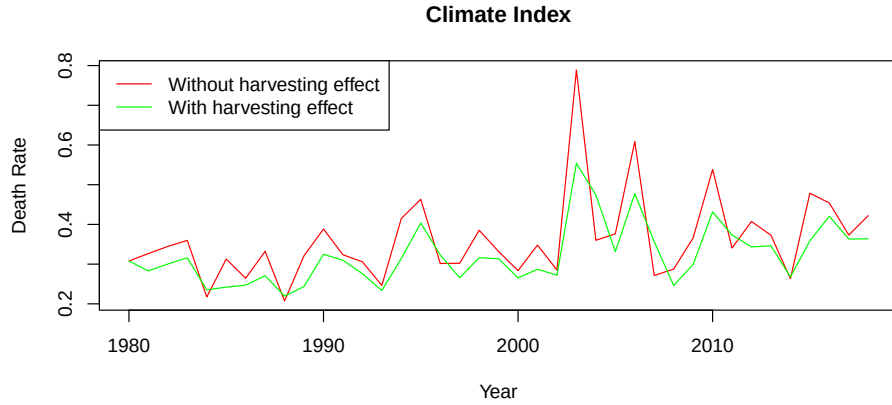


Figure 2.3.3: With and Without Harvesting Effect Climate Index in Terms of Death Rate in Belgium (both sexes, all ages, per 100,000 population).

previously obtained age-class parameters into age-continuous parameters, yielding the final model equation:

$$\log(\mu_{xt}) = \alpha_x + \beta_x^{(o)} \kappa_t^{(o)} + \delta_x^{(c)} CI_t. \quad (2.3.6)$$

This final calibration is performed as before by fitting a standard Lee–Carter model to the residuals

$$R_{xt} = \log(\mu_{xt}) - \delta_x^{(c)} CI_t, \quad (2.3.7)$$

in order to estimate the parameters α_x , $\beta_x^{(o)}$, and $\kappa_t^{(o)}$. We employ the same methodology as for the previous calibrations, namely the *StMoMo* package in R, whose specifics are described by Villegas et al. (2018). Similar results can be obtained by applying the general methodology outlined in Section 1.1.3. The fitted values of the parameters for the final model (2.3.6) are presented in Figure 2.3.4, and the climate index accounting for the harvesting effect is shown in Figure 2.3.3.

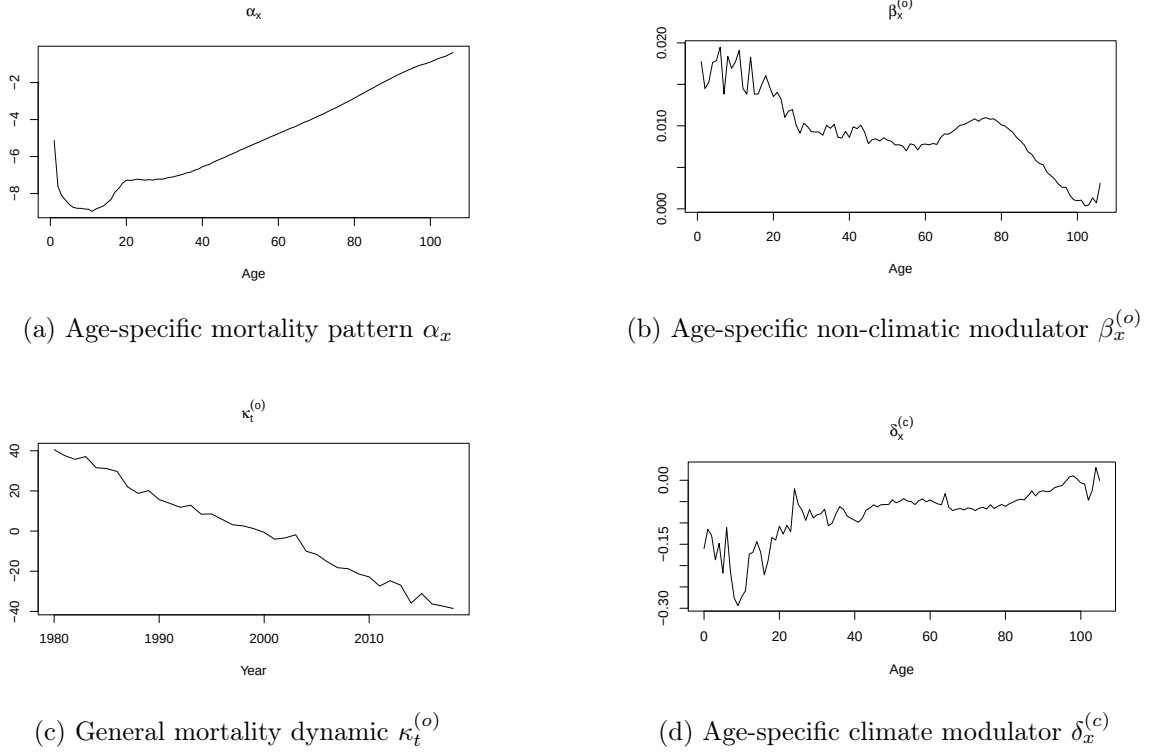


Figure 2.3.4: Estimated Parameters of the Modified Lee-Carter Model in Continuous Age.

The combined effect of the age-dependent increasing values of $\delta_x^{(c)}$ and the heat peaks captured by CI_t effectively highlights the phenomenon of climate change-related heat waves increasing mortality rates among older individuals, thereby counteracting the generally declining mortality trend captured by $\beta_x^{(o)} \kappa_t^{(o)}$, a pattern well documented in the literature.

2.4 Performance Comparison

In this section, we compare the performance of our final model (2.3.6), which accounts for climate-specific mortality, with the classical RD Lee and LR Carter (1992) model given by equation (1.1.6) under the assumption of a log-link function. Both models, calibrated on Belgian data using the techniques described above, are assessed by comparing various goodness-of-fit indicators.

As an example, Figure 2.4.1 plots the mortality rate at age 85 for the observed data, the

classical Lee–Carter model, and our climate-adjusted Lee–Carter model. As expected, the two model trajectories exhibit a very high linear correlation, with $R^2 = 99.97\%$, since the models remain similar despite the separation of climate-related mortality causes. Figure 2.4.1 also shows that our climate-adjusted Lee–Carter model achieves a slightly better fit to the observed data, with $R^2_{CLC} = 98.41\%$ compared to $R^2_{LC} = 98.33\%$ for the classical model. These results are satisfactory and mirror the findings of Boumezoued et al. (2022) on French mortality data, which reported $R^2_{CLC, FR} = 99.75\%$ versus $R^2_{LC, FR} = 99.73\%$.

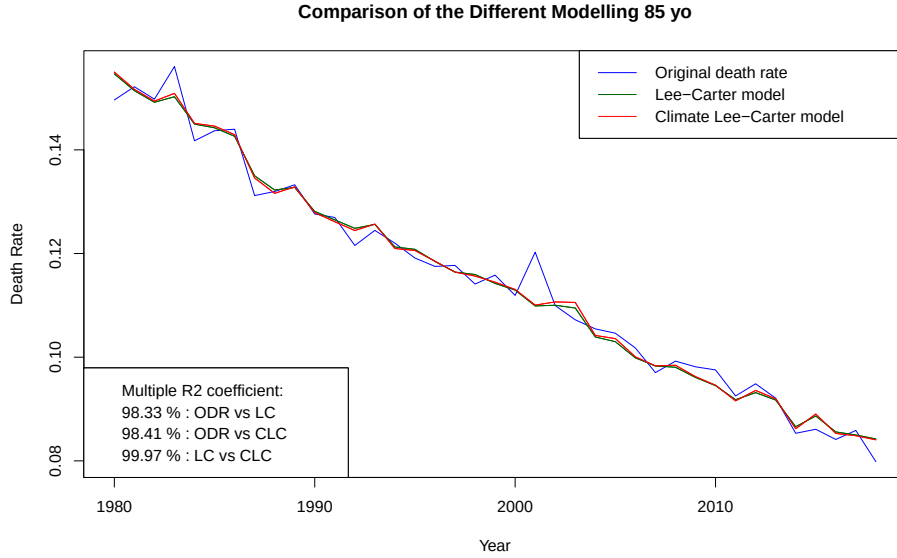


Figure 2.4.1: Comparison of the Mortality Rate at age 85 in terms of Linear Correlation Coefficients.

Still in terms of linear correlation coefficient R^2 , Figure 2.4.2 shows the coefficient values for both models across the entire age range. A brief inspection reveals that the climate-adjusted Lee–Carter model (2.3.6) outperforms the classical Lee–Carter model at older ages, notably from age 60 onwards. For example, the climate-adjusted model yields a higher R^2 for 31.13% of all ages, for 54.35% of ages above 60, for 69.44% of ages above 70, and for 96.15% of ages above 80. We note that, while our model is less performant on average, it achieves a better fit at higher ages, which is crucial for a model intended for life insurance pricing.

Since the ultimate goal is to produce projections of Belgian mortality using our climate-based model, we can also consider criteria that penalize overfitting. To this end, we retrieved the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) as explained by Kuha (2004), both of which penalize models with an excessive number of parameters. To compute these criteria, we require the maximized log-likelihood of the models. Specifically, we use the full Poisson log-likelihood, which underlies the expression given in (1.1.15), and is defined as:

$$\log \mathcal{L}(\hat{\theta}) = \sum_{x,t} \left[D_{x,t} \ln(E_{x,t} \hat{\mu}_{x,t}) - E_{x,t} \hat{\mu}_{x,t} - \ln(D_{x,t}!) \right], \quad (2.4.1)$$

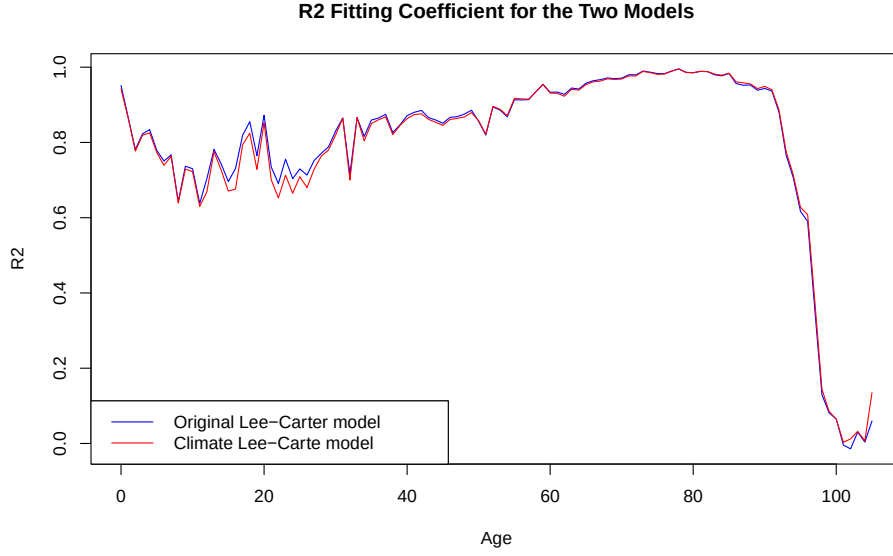


Figure 2.4.2: Comparison of Goodness-of-fit (R^2) for Both Models Across All Ages.

where $\hat{\mu}_{x,t} = \exp(\alpha_x + \beta_x^{(o)} \kappa_t^{(o)} + \delta_x^{(c)} CI_t)$ for the Climate Lee–Carter model (or without the $\delta_x^{(c)} CI_t$ term in the classical Lee–Carter model). Let k denote the total number of estimated effective parameters, and n_{obs} the number of observations. We then define:

$$\text{AIC} = -2\ell(\hat{\theta}) + 2k, \quad (2.4.2)$$

$$\text{BIC} = -2\ell(\hat{\theta}) + k \ln(n_{\text{obs}}). \quad (2.4.3)$$

Table 2.4.1 summarizes the results: the model incorporating climate-related drivers exhibits a slight improvement according to both selection criteria, consistent with the findings of Boumezoued et al. (2022) on French data.

Model	logLik	k	n_{obs}	AIC	BIC
Classical Lee–Carter	−19 911.90	249	4 134	40 321.79	41 897.22
Climate Lee–Carter	−19 457.82	357	4 134	39 629.64	41 887.21

Table 2.4.1: Comparison of AIC and BIC for the Fitted Classical Lee–Carter and Climate Lee–Carter Models.

Finally, we have demonstrated that the available Belgian data allow for an effective fit of a climate index, which can subsequently be used to calibrate an extension of the classical model of RD Lee and LR Carter (1992) following the methodology proposed by Boumezoued et al. (2022). The performance of the resulting model slightly surpasses that of the standard model without climate inputs, as shown in Figure 2.4.1 and Table 2.4.1. These reassuring results motivate us to undertake a forward-looking analysis on Belgian data, aiming to construct mortality tables for the 2050 horizon and to assess the impact of climate change on the pricing of life insurance products. This prospective approach will be detailed in Chapter 4.

Chapter 3

Milhaud and Gouthon’s Approach on Belgian Data

In this chapter, we apply the methodology of Milhaud and Gouthon (2024), previously detailed in Section 1.3, to Belgian data in order to construct a GAPC model of mortality rates capable of explaining the effects of heatwaves on mortality. To this end, we first describe, as already done at the beginning of Chapter 2, the available Belgian data and their preprocessing, before determining the age threshold above which temperature has a notable effect on mortality in Belgium, as outlined by the authors. Finally, we calibrate the proposed climate model using the fitting methods discussed in Section 1.1.3.

3.1 Available Belgian Data

As with the approach already applied to the Boumezoued et al. (2022) model, we will require three complementary datasets in order to determine the pivot age at which Belgian mortality becomes sensitive to temperature, and subsequently to calibrate the full model.

3.1.1 Weekly Belgian Mortality (Short Term Mortality Fluctuation, STMF)

The first dataset aggregates weekly Belgian mortality data by sex (male, female, both sexes) over the period 2000–2024. This dataset, called Short Term Mortality Fluctuation (STMF)¹, extends the data originally described in Section 2.1.2, providing finer temporal granularity suitable for the methodology of Milhaud and Gouthon (2024). The data are disaggregated into death counts and death rates; we will use the latter in our estimates and calibrations, as it accounts for exposure to risk, which is appropriate for this type of actuarial calculation. We will employ the period 2000–2019 to determine our age threshold in order to avoid potential biases induced by COVID-19 from 2020 onward. Finally, note that ages are grouped into five classes: 0–14, 15–64, 65–74, 75–84, and 85+, which will limit the determination of our pivot age to five potential values, as detailed in Section 3.2.

¹The data are available at <https://www.mortality.org/Data/STMF>.

3.1.2 Belgian Global Mortality (Human Mortality Database, HMD)

As recalled in Section 2.1.2, the Human Mortality Database (HMD) describes itself as the world’s leading scientific data resource on mortality in developed countries. The available life tables² provide all necessary information and serve as the standard reference within the actuarial community for analyzing the evolution of a country’s overall mortality. We will employ the life tables for the period 1980–2019 across the age range 20–100, as recommended by Milhaud and Gouthon (2024), for calibrating the model that accounts for the impact of climate change on mortality.

3.1.3 Belgian Meteorological Data (Royal Meteorological Institute, RMI)

As noted in Section 2.1.3, the Royal Meteorological Institute (RMI) provides daily datasets³ for each location in Belgium, segmented into grid cells, containing various climate variables such as minimum, maximum, and mean temperatures, precipitation amounts, relative humidity, sunshine duration, and so forth. Unlike the approach of Boumezoued et al. (2022), we will use solely the weekly temperatures during the summer period for each year to perform a linear regression with the mortality data from the STMF dataset presented in Section 3.1.1, aiming to determine the age threshold above which temperature has a significant impact on mortality. We detail this procedure in Section 3.2.

3.2 Age Threshold Determination

The longevity model proposed by Milhaud and Gouthon (2024) uses a hyperparameter referred to as the age threshold, A_h , which corresponds to the age beyond which high temperatures have harmful impacts on human health, in order to isolate climatic effects. This threshold allows the model to incorporate climate-related mortality effects only from a certain age onward in the population. We therefore apply the methodology presented in Section 1.3.2 to Belgian data using the datasets introduced in Sections 3.1.1 and 3.1.3.

3.2.1 Empirical Evidence

The work of Gasparrini et al. (2015) highlights the existence of an age at which the human body becomes more vulnerable to extreme temperatures, particularly heatwaves. It is precisely this age that we seek to identify for the Belgian population. Using our weekly Belgian mortality data, we can, for example, plot the death rate against the week of the year, focusing on the warmer weeks, for three years with well-known climatic trends.

For instance, we know that the heatwaves of August 2003 (weeks 31–33) and July 2006 (weeks 28–30)⁴ had a substantial impact on mortality in Belgium, whereas 2010 was relatively ordinary in terms of climate data. The mortality results are shown in Figures 3.2.1 and 3.2.2. We readily observe that the heatwaves of August 2003 and July 2006 manifest as two distinct peaks – more than 1.5 times higher than the 2010 mortality trend – for the 85+

²Belgian mortality data are accessible via <https://www.mortality.org/> and <https://www.mortality.org/Country/Country?cntr=BEL>.

³These datasets are accessible via the RMI’s new online platform <https://opendata.meteo.be/>.

⁴The week numbering is available via <https://www.meteo.be/fr/climat/climat-de-la-belgique/evenements-remarquables-depuis-1901/evenements-remarquables/evenements/canicule>.

age class, whereas these peaks are virtually absent for the age class assumed to be the most resilient, namely 15–64. These findings underscore the presence of an age threshold for the Belgian population as well, which we will attempt to determine more precisely in the following section.

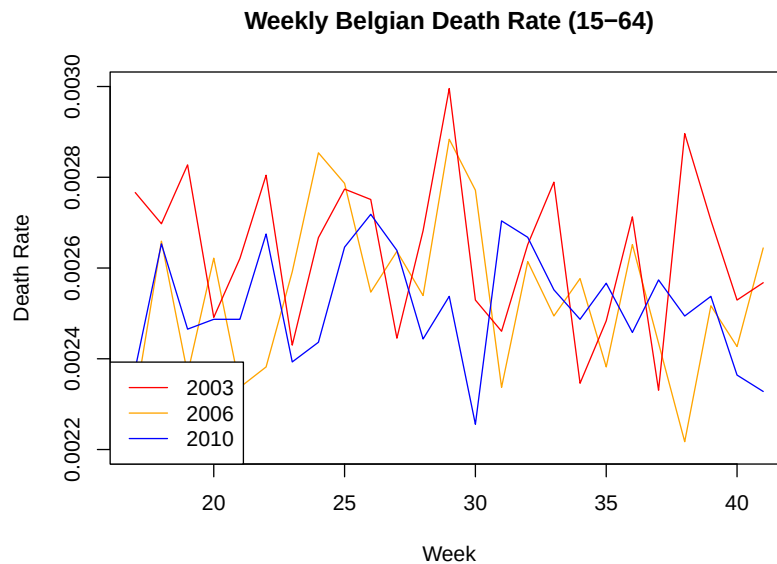


Figure 3.2.1: Belgian Weekly Death Rate in 2003, 2006 and 2010 for the 15-64 Age Class (both sexes.)

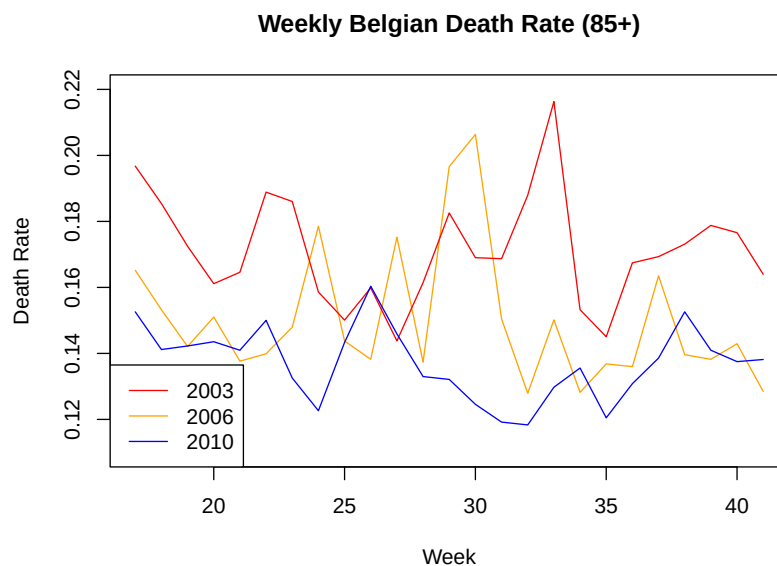


Figure 3.2.2: Belgian Weekly Death Rate in 2003, 2006 and 2010 for the 85+ Age Class (both sexes).

3.2.2 Linear Correlations

The empirical indications of an increased vulnerability age threshold to heatwaves in the Belgian population, as presented in Figures 3.2.1 and 3.2.2, motivate us to apply the methodology of Milhaud and Gouthon (2024) for estimating this age, A_h . The framework detailed in Section 1.3.2 consists of performing a linear regression between the weekly mean summer temperatures – constructed from the RMI dataset described in Section 3.1.3 – and the mortality observed during those same weeks, using the STMF dataset of Section 3.1.1. We consider the summer period (June to August) encompassing weeks 22 to 35 (with the first week of January numbered as week 1) for the years 2000 to 2019, to respect the limits of both datasets.

Using the notation of Section 1.3.2, and indexing each age class by $k = 1, 2, \dots, 5$, Milhaud and Gouthon (2024) define the points $(d_j, \theta_j)_k$ where

- d_j is the number of deaths recorded during the j^{th} week,
- θ_j is the mean temperature recorded during the j^{th} week.

From these definitions, we perform a Pearson correlation test (Pearson, 1985) to identify the first age class k for which high temperatures begin to impact longevity.

Applying exactly the methodology of Milhaud and Gouthon (2024) – i.e., testing the correlation between death counts and mean temperature within the same week – we obtain the results in Table 3.2.1. As shown, the results are difficult to interpret for determining an age threshold: the 5% significance level is met first for the 15–64 group, then not met for 65–74, and then met again for 75–84 and 85+.

Age Group	Pearson Correlation	p -value
0–14	0.050891	0.396269
15–64	0.129331	0.030501
65–74	0.113846	0.057082
75–84	0.174602	0.003377
85+	0.332134	0.000000

Table 3.2.1: Pearson Correlations by Age Group (using death counts).

We therefore propose a slight modification to Milhaud and Gouthon (2024)’s approach by replacing the point $(d_j, \theta_j)_k$ with $(\mu_j, \theta_j)_k$, where μ_j represents the mortality rate in the j^{th} week. This change accounts for exposure to risk and yields more reliable and stable results. The outcomes are presented in Table 3.2.2, which indeed identify an age threshold for the Belgian population of $A_h = 75$, since this is the first class for which the p -value falls within the desired statistical tolerance. Note that, due to the five age classes in the STMF dataset, finer granularity in threshold determination is not possible.

This pivot age of 75 will thus serve to isolate the climatic impact on mortality in the longevity model proposed by Milhaud and Gouthon (2024). Indeed, as we will see in the following section, the proposed model relies both on the age threshold we have determined for the Belgian population and on a triggering event that “activates” the climate impact on mortality.

Age Group	Pearson Correlation	<i>p</i> -value
0–14	0.039005	0.515685
15–64	0.083152	0.165265
65–74	0.005785	0.923233
75–84	0.124351	0.041681
85+	0.279776	0.000002

Table 3.2.2: Pearson Correlations by Age Group (using death rates).

3.3 Model Calibration

The approaches of Boumezoued et al. (2022) and Milhaud and Gouthon (2024) differ primarily in how they integrate climatic variations into their longevity models. Whereas Boumezoued et al. (2022) proposes to construct a climate index that reflects the impact of climate on mortality over a year based on the incidence of various climate variables, Milhaud and Gouthon (2024) proposes to jointly define an age threshold – determined in Section 3.2 – beyond which exposure to heatwaves affects mortality (otherwise the effect is not considered), and a triggering event corresponding to years hotter than the average. This triggering event is defined by Milhaud and Gouthon (2024) as an indicator equal to 1 if the mean summer temperature of the year in question exceeds the average summer temperature over the entire study period and 0 otherwise. Thus, only for certain years and above a certain age are climate effects on mortality captured and applied in the longevity model. We will first recall the structure of the longevity model proposed by Milhaud and Gouthon (2024) and then attempt a calibration using Belgian data; however, we will show that the absence of observations for certain years will prevent us from capturing relevant information regarding the impact of climate on mortality.

3.3.1 Fitting Methodology

As recalled in Section 1.3.1, the longevity model proposed by Milhaud and Gouthon (2024) is specified as follows:

$$\begin{aligned}
\log(\mu_{xt}) = & \alpha_x + \kappa_t^{(1)} + (\bar{x} - x)\kappa_t^{(2)} + (\bar{x} - x)^+ \kappa_t^{(3)} \\
& + \left((x_1 - x)_+ + w(x - x_2)^+ \right)^2 \kappa_t^{(4)} \\
& + (x - A_h)^+ \mathbb{1}_{\{I_t - \bar{I}_t > 0\}} \kappa_t^{(5)} + \gamma_{t-x},
\end{aligned} \tag{3.3.1}$$

where $\kappa_t^{(5)}$ captures the year-over-year evolution of mortality attributable to climatic effects. The indicator function $\mathbb{1}_{\{I_t - \bar{I}_t > 0\}}$ ensures that only years with summer temperatures above the long-term average are considered, with I_t denoting the mean summer temperature in year t , $\bar{I}_t$ the average summer temperature over the entire observation period, and $A_h = 75$ the age threshold beyond which high temperatures begin to impact human health, as determined in Section 3.2. For further details on notation, the reader is referred to Section 1.3.1.

To initiate our fitting procedure, note that aside from the already estimated A_h , model (3.3.1) includes three additional hyperparameters that must be estimated separately, as explained in Section 1.3.3. For our initial calibration trials, we set $x_1 = 45$, $x_2 = 60$, and $c = 0.5$,

following the recommendations based on the French data. We also recall that Milhaud and Gouthon (2024) focus on ages 20 to 85+ (where the 85+ group represents all individuals aged 85 and above) and on the period from 1980 to 2019. Accordingly, we adjust our dataset described in Section 3.1.2 and deliberately forgot splitting into training and validation periods for this first fitting attempt, in order to capture as fully as possible the heatwaves of the recent years.

Last but not least, the presence of the indicator function for the triggering event in model (3.3.1) precludes the use of the `StMoMo` package in R, since it depends on the calendar year t and cannot be specified as an age-modulating function, which may only depend on age x . Consequently, it is necessary to implement a custom Newton–Raphson iterative algorithm (Akram and Ann, 2015), to employ the general methodology presented in Section 1.1.3, and to adhere to the constraints proposed by Milhaud and Gouthon (2024) as given in Equations (1.3.3) and (1.3.4).

3.3.2 Estimation Results

After applying the methodology described in the previous section, we obtain a rather disappointing fit for the Belgian data compared to the results reported by Milhaud and Gouthon (2024) for the French data. Although the final fit of the model in Figure 3.3.1 appears satisfactory in both illustrative cases, and the results in Figure 3.3.2 are coherent and align with known patterns for the age-specific mortality pattern, α_x , and the cohort effect, γ_c , the estimates intended to isolate climatic effects in Figure 3.3.3 do not capture the impact of heatwaves effectively. Note that we have used the notations of Section 1.1.1 for the parametric age-modulator functions $\beta_x^{(i)}$ for $i = 0, \dots, 5$.

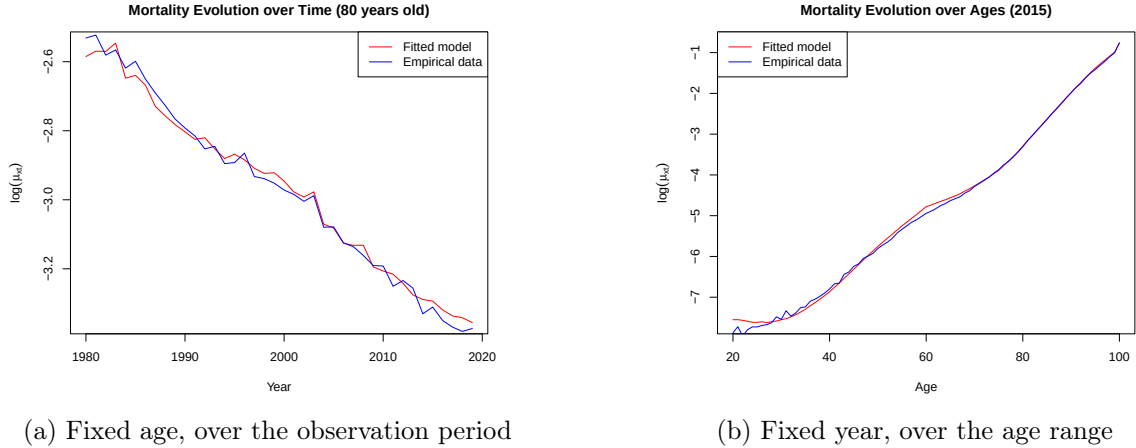


Figure 3.3.1: Visual Goodness-of-Fit of the Model on Belgian Data in two Illustrative Scenarios.

In fact, a closer look at Figure 3.3.3 reveals that the time series intended to capture the effects of heatwaves, $\kappa_t^{(5)}$, is simply an increasing curve at the points where the triggering event occurs – that is, when the mean summer temperature exceeds the overall summer average. This time series therefore fails to capture the specific effects of hotter years: one might have

expected to observe distinct peaks in $\kappa_t^{(5)}$ for 2003 and 2006, which are well-known heatwave years in Belgium, but this is not the case, unlike the results obtained by Milhaud and Gouthon (2024) with the French data. Consequently, the model fitted to the Belgian data does not capture sufficient climate-related information to serve as a model integrating the impacts of climatic variations on longevity in Belgium.

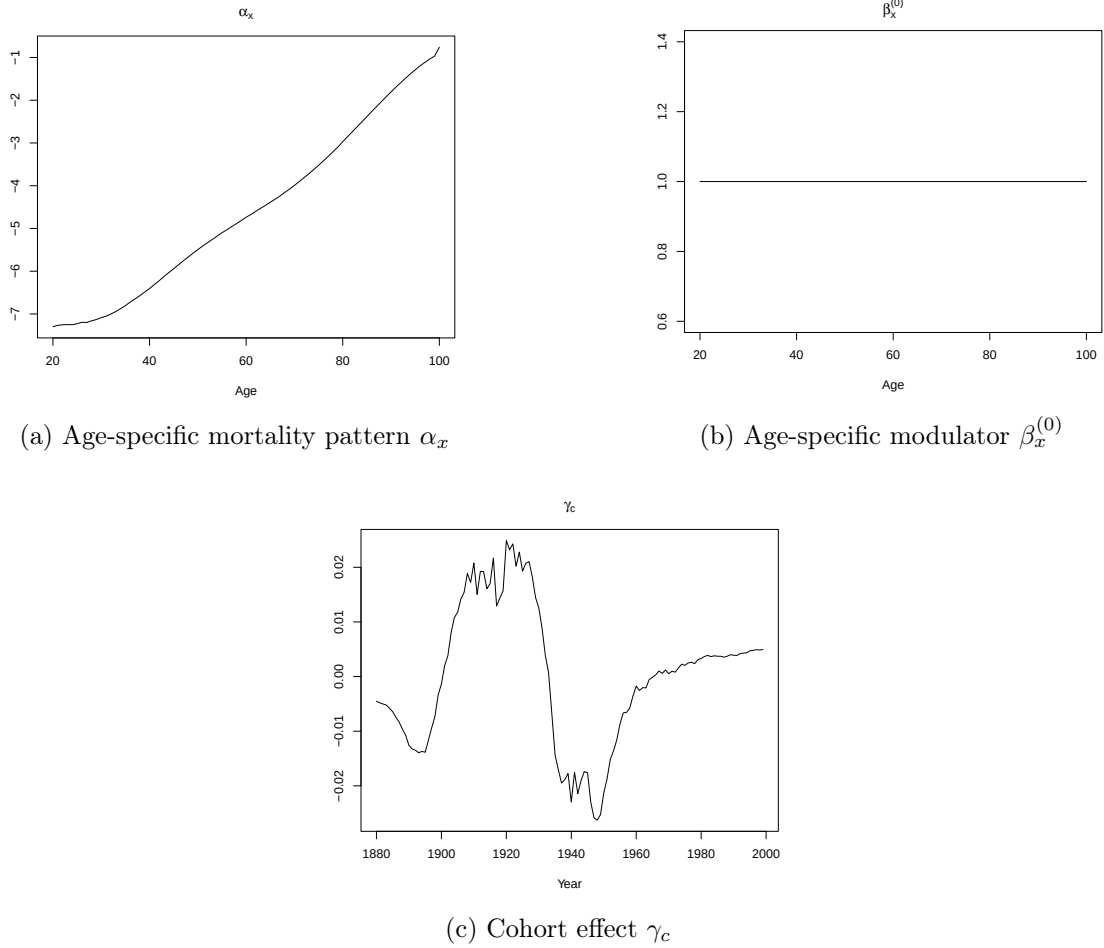


Figure 3.3.2: Key estimated parameters not related to climate.

Figure 3.3.4 displays the estimated age-specific modulators $\beta_x^{(1)}$ through $\beta_x^{(4)}$. Ideally, each $\kappa_t^{(i)}$ represented in Figure 3.3.5 should align with its corresponding $\beta_x^{(i)}$ in a way that they jointly capture a unique dimension of the mortality surface. However, our Belgian calibration shows substantial overlap: none of the $\kappa_t^{(i)}$ series isolates a single effect cleanly. Instead, they all appear to respond similarly over time, indicating that the model components are not disentangling the intended trends. This contrasts with the French results obtained by Milhaud and Gouthon (2024), where each $\kappa_t^{(i)}$ exhibited a more distinct dynamic. Consequently, the lack of clear separation among the κ_t trajectories suggests that the model as specified may not be well-suited to the Belgian context without further adaptation.

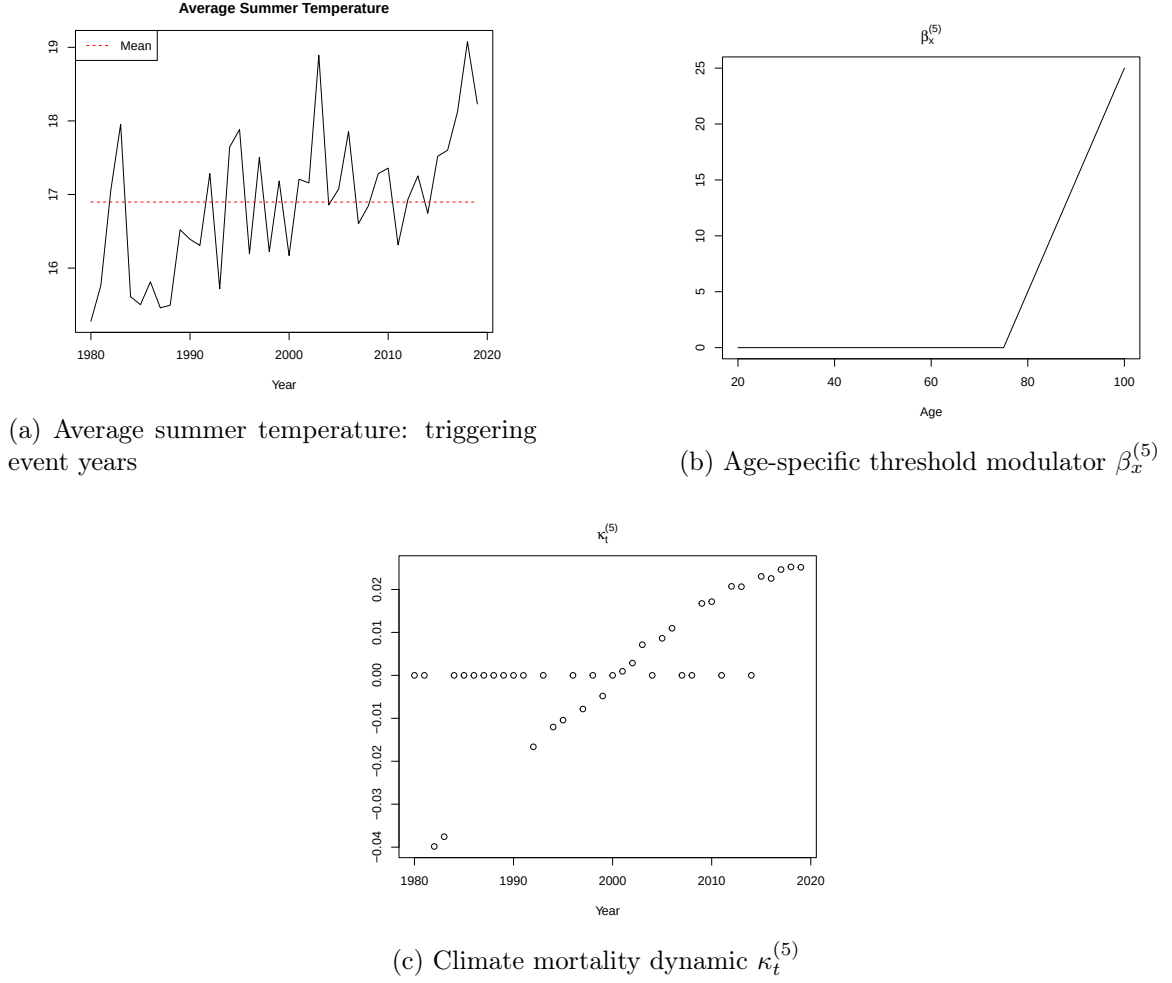


Figure 3.3.3: Key estimated parameters related to climate.

In conclusion, we have successfully applied the methodology of Milhaud and Gouthon (2024) to identify an age threshold of increased vulnerability to extreme temperature exposure within the Belgian population. However, despite several attempts to remove or add parameters, the fit obtained for the climate-effect component of the model proposed by the authors remains incoherent and does not capture the desired information for the Belgian data. We are unable to establish a specific linkage between heat and mortality with this model. Therefore, we will not pursue projections based on this model in Chapter 4; instead, we will leverage the age threshold results and explore how to integrate these findings and methodology into the framework of Boumezoued et al. (2022) to determine whether a hybrid model can achieve a better fit than the one analyzed in Chapter 2.

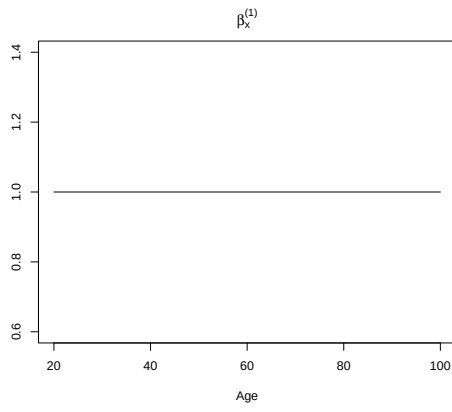
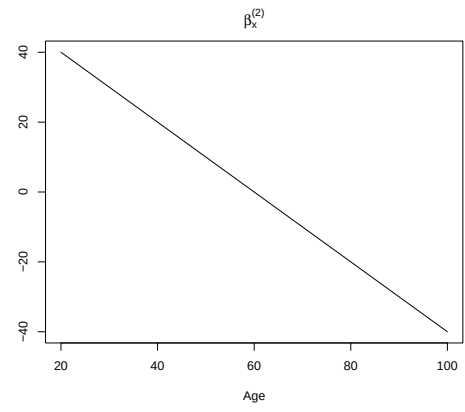
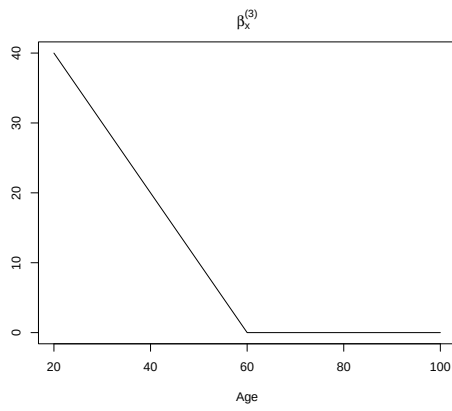
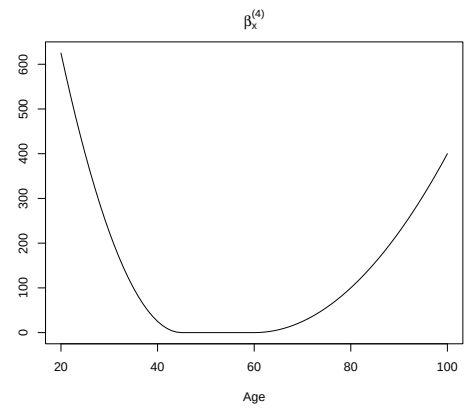
(a) Age-specific modulator $\beta_x^{(1)}$ (b) Age-specific modulator $\beta_x^{(2)}$ (c) Age-specific modulator $\beta_x^{(3)}$ (d) Age-specific modulator $\beta_x^{(4)}$

Figure 3.3.4: Estimated Age-Specific Modulators of the Milhaud and Gouthon (2024)'s Model

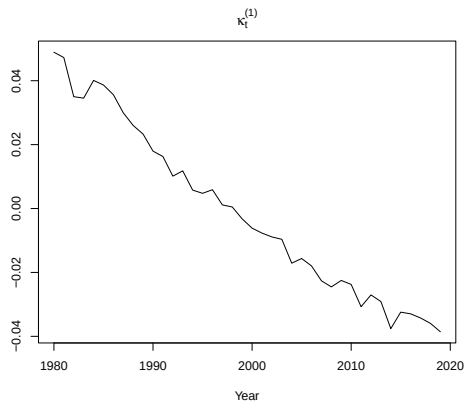
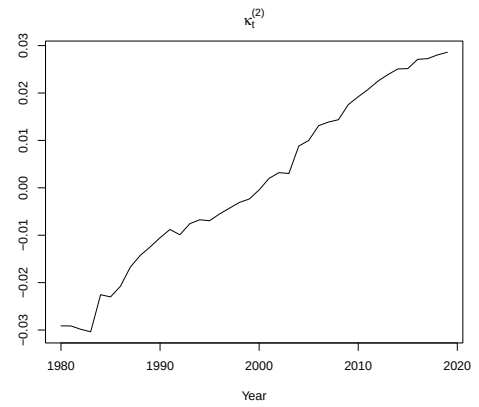
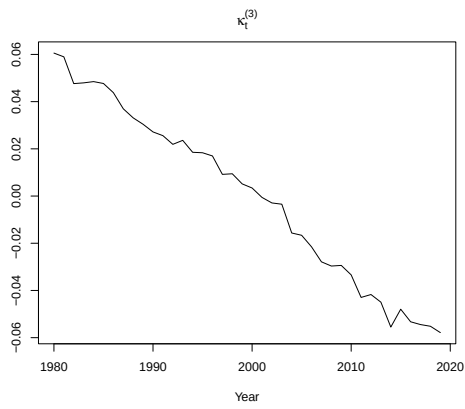
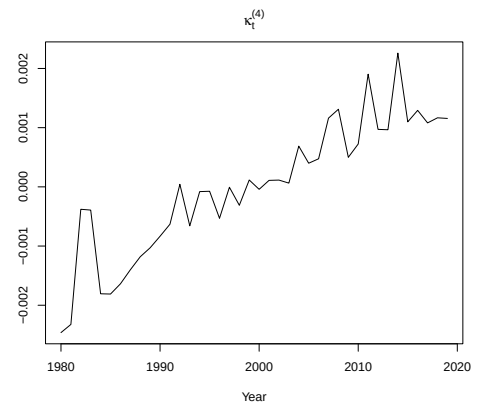
(a) Mortality dynamic time series $\kappa_t^{(1)}$ (b) Mortality dynamic time series $\kappa_t^{(2)}$ (c) Mortality dynamic time series $\kappa_t^{(3)}$ (d) Mortality dynamic time series $\kappa_t^{(4)}$

Figure 3.3.5: Estimated Mortality Dynamic Time Series

Chapter 4

Prospective Approach on Belgian Data

We conclude this thesis with a prospective approach based on Belgian data, aiming to quantify the impact of climate change on mortality risk by the year 2050. To this end, we begin by comparing the advantages and limitations of the approaches proposed by Boumezoued et al. (2022) and Milhaud and Gouthon (2024) for longevity modelling and propose a model inspired by those previously studied in Chapters 2 and 3. Once the fit of this model is established, we apply time series analysis methods to project mortality data up to 2050. This will allow us to construct prospective mortality tables, which will be used to assess the impact of climate change on life expectancy and on the pricing of basic life insurance products.

4.1 Final Prospective Model

As we have seen in Chapters 2 and 3, Boumezoued et al. (2022) and Milhaud and Gouthon (2024) propose two distinct approaches to building longevity models that incorporate climate risk into their equations. These two methodologies each present specific advantages and limitations when applied to Belgian data. For this reason, we propose to construct and study a third model that integrates selected elements from both approaches, in order to better suit the characteristics of the population under study.

4.1.1 Proposed Model

The objective here is to retain the most effective elements of both approaches, or at least those that yielded the most conclusive results when applied to Belgian data. As previously mentioned, the results presented in Section 3.2 demonstrate that the methodology of Milhaud and Gouthon (2024) for determining an age threshold from which mortality is affected by heatwaves is statistically effective in the Belgian population. Indeed, Table 3.2.2 highlights the age threshold $A_h = 75$ years based on our data. On the other hand, the methodology proposed by Boumezoued et al. (2022) for integrating climate data via a climate index, accounting for the harvesting effect, and calibrating the model using Belgian data has shown better performance than the classical model introduced by RD Lee and LR Carter (1992), as demonstrated in Section 2.4. Finally, we will rely on a projection methodology based on RCP scenarios, inspired by the Section 4 of Milhaud and Gouthon (2024), as we will see in Section 4.2.

We therefore propose to adopt the majority of the methodology introduced by Boumezoued et al. (2022), as detailed in Sections 1.2.3 and 2.3, while calibrating the model separately for two age groups:

- $c_R = [0, 74]$: the *climate-resistant* age group, i.e., the age range for which heatwaves did not have a statistically significant impact on mortality according to the analysis conducted in Section 3.2,
- $c_V = [74, 105]$: the *climate-vulnerable* age group, i.e., the age range for which heatwaves showed a statistically significant impact on mortality based on the same analysis.

This segmentation offers both a statistically justified partition of the population into distinct groups and a reduction in the number of parameters to estimate, which helps mitigate overfitting – an important consideration when performing long-term projections. Indeed, the absence of climate sensitivity in the c_R group implies that we can set $\delta_{c_R}^{(c)} = \vec{0}$.

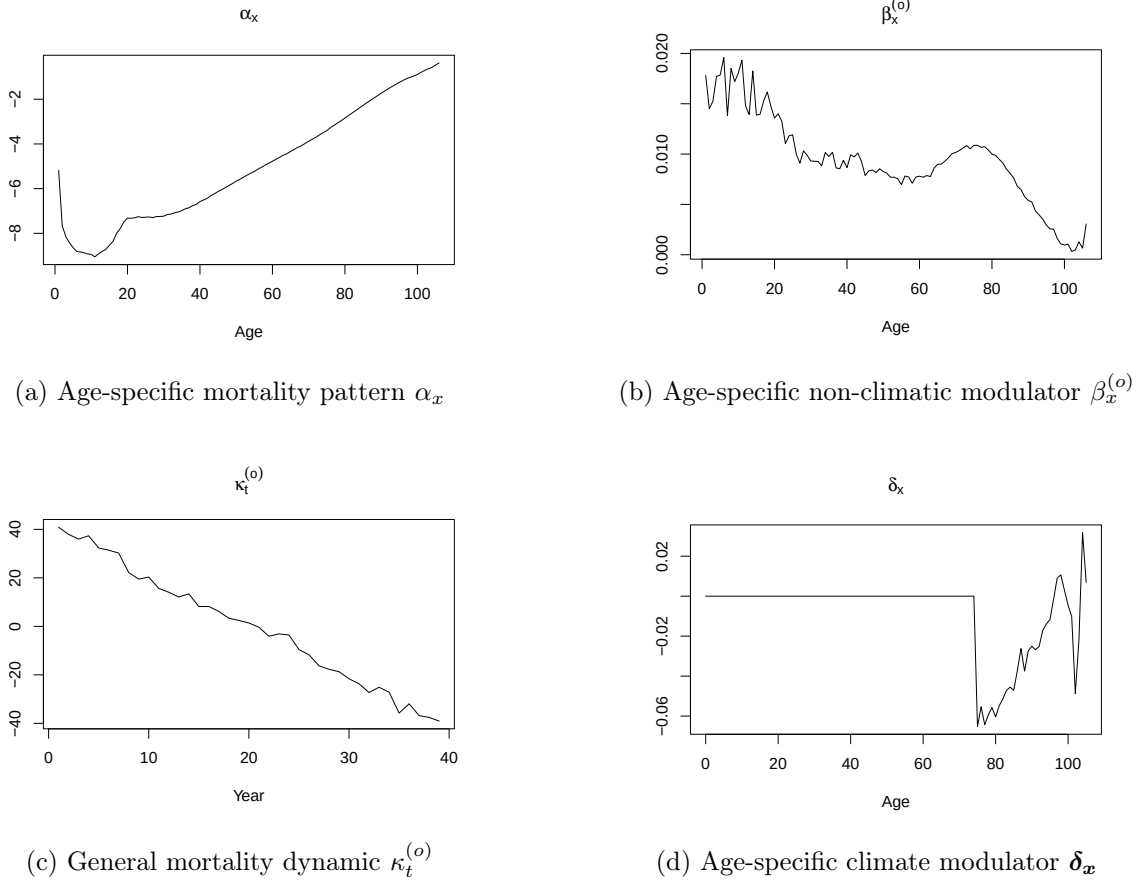


Figure 4.1.1: Estimated Parameters of the Final Modified Lee–Carter Model.

Following the notation and fitting steps introduced throughout Chapter 2, we can now present the final model that will be used for projection purposes:

$$\log(\mu_{xt}) = \alpha_x + \beta_x^{(o)} \kappa_t^{(o)} + \delta_x C I_t, \quad (4.1.1)$$

where $\delta_{\mathbf{x}} = (\vec{0}, \delta_{cV}^{(c)})$ since we consider the impact of the climate index to be null for ages $x < 75$. A graphical representation of the estimated parameters is presented in Figure 4.1.1, using the same climate index as in Figure 2.3.3, since it was estimated independently from the calibration of the Lee-Carter model.

4.1.2 Performance Comparison

The introduction of this slightly modified model is justifiable provided that its fitting performance exceeds that of the original model. We therefore proceed to compare the performance of the model proposed by Boumezoued et al. (2022), as obtained in Section 2.4, with the performance of the newly introduced model (4.1.1).

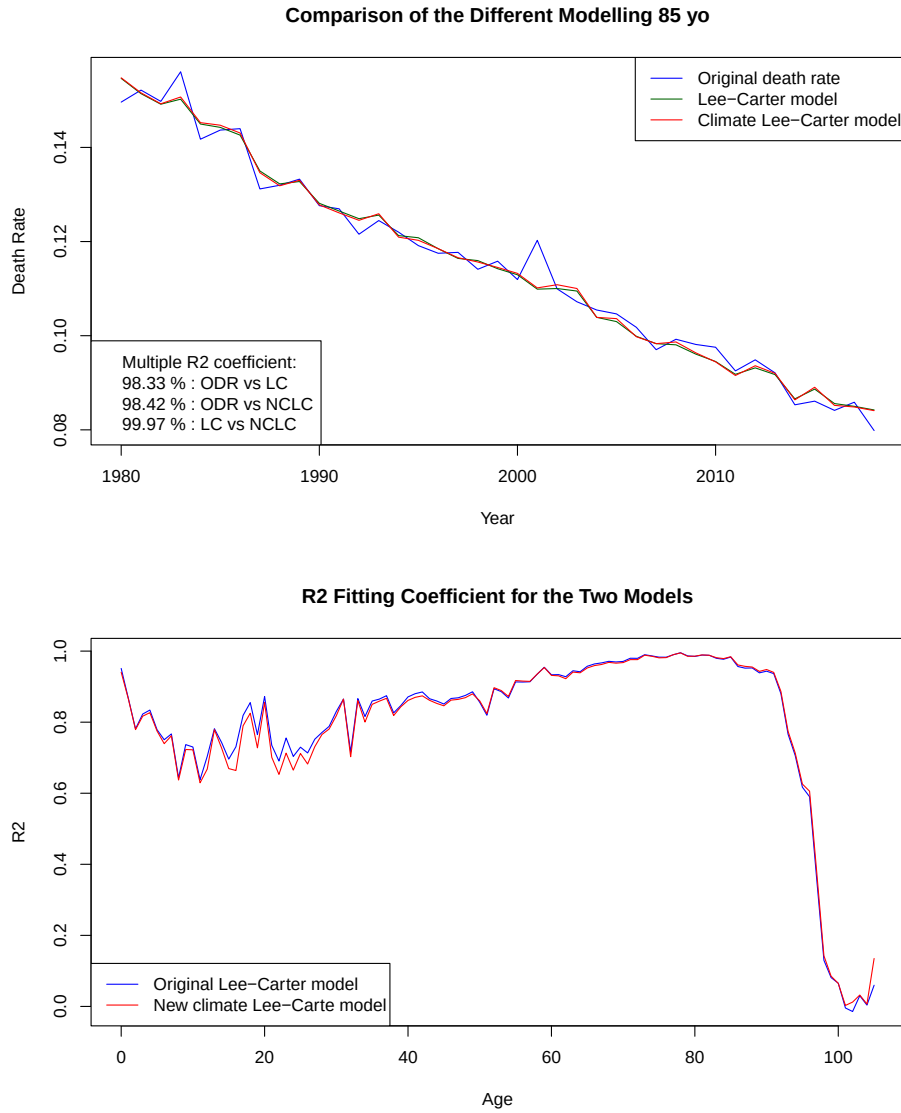


Figure 4.1.2: Comparison of the Mortality Rate in terms of Linear Correlation Coefficients.

We begin by comparing the models in terms of linear correlation coefficients, as done by Boumezoued et al. (2022). To this end, Figure 4.1.2 presents the same types of plots as in Section 2.4, for illustrative purposes. A quick examination of Figure 4.1.2 shows a fit of $R^2_{NCLC} = 98.42\%$ (*NCLC* standing for New Climate Lee-Carter, the proposed model), which is slightly better than the one obtained previously. Moreover, Table 4.1.1 confirms this trend: the newly proposed model achieves a higher R^2 than the original model by Boumezoued et al. (2022) across all ages, and it explains a greater share of the variance, particularly at older ages, compared to the classical Lee-Carter model, which does not isolate climate-related causes.

Age Group	Boumezoued et al.'s Model (%)	Proposed Model (%)
All ages	31.13	33.02
Ages > 60	54.35	56.52
Ages > 70	69.44	72.22
Ages > 80	96.15	96.15

Table 4.1.1: Comparison of Percentage of Age Groups where the Climate-Adjusted Boumezoued et al.'s Model and the Proposed Model Outperform the Classical Lee-Carter Model.

Finally, since our objective is to perform projections, it is important to consider information criteria. We therefore use the same calculation methods for AIC and BIC as in Section 2.4. The results presented in Table 4.1.2 show that, despite the proposed model having a slightly lower log-likelihood than the model by Boumezoued et al. (2022), the lower number of parameters to be estimated results in better AIC and BIC values. These results are encouraging and suggest that the proposed model is better suited for performing longevity projections.

Model	logLik	k	n_{obs}	AIC	BIC
Lee-Carter's Model	-19 911.90	249	4 134	40 321.79	41 897.22
Boumezoued et al.'s Model	-19 457.82	357	4 134	39 629.64	41 887.21
Proposed Model	-19 503.47	282	4 134	39 570.94	41 355.63

Table 4.1.2: Comparison of Lee-Carter and Climate-LC Models.

4.2 Impact at the 2050 Horizon

Mortality rate projection is particularly useful in actuarial calculations. Indeed, having projected rates allows us to compute life expectancy prospectively and to price insurance or pension products by anticipating mortality trends, thereby covering risk more precisely to avoid potential losses or to remain competitive in the market. We begin by recalling the projection methods that are well documented in the literature and apply them to the classical RD Lee and LR Carter (1992) model calibrated on Belgian data. We then propose an adaptation of these methods for projecting the model introduced in (4.1.1), before comparing the results obtained from these two projections in order to assess the incorporation of climate risk into standard life insurance products.

4.2.1 Classical Lee-Carter Model Projections using Random Walk with Drift

As we have seen, the models introduced in this thesis belong to, or are extensions of, the family of GAPC stochastic mortality models presented in Section 1.1.1. That is, as explained by Villegas et al. (2018), mortality dynamics are driven by the period indices $\kappa_t^{(i)}$, $i = 1, \dots, N$. Consequently, to simulate projections of mortality rates, it is necessary to rely on time-series analysis and forecasting techniques such as those presented by von Sachs (2024). We choose here to rely on the standard approach in the actuarial literature, notably Villegas et al. (2018), which assumes that the period indices follow a multivariate random walk with drift. Since the classical Lee–Carter model has only one period index, we can write

$$\kappa_t^{(1)} = \delta^\kappa + \kappa_{t-1}^{(1)} + \xi_t^\kappa, \quad \xi_t^\kappa \sim \mathcal{N}(0, \sigma_{\xi_t^\kappa}), \quad (4.2.1)$$

where δ^κ is the drift parameter and $\sigma_{\xi_t^\kappa}$ is the variance of ξ_t^κ , the univariate white-noise term. We obtain the estimates using the forecasting functionality of the **StMoMo** package in R, as reported in Table 4.2.1. Note that we could also have obtained similar results using the techniques and theoretical developments presented by von Sachs (2024).

Parameter	Value
$\sigma_{\xi_t^\kappa}$	5.1355
δ^κ	−2.020

Table 4.2.1: Estimates for the Random Walk with Drift Parameters of the Classical Lee–Carter.

Let

- t_n , the last calibration year, namely 2018,
- s , the projection horizon, $s = 1, \dots, 32$,
- $\dot{\mu}_{x,t_n+s}$, the mortality rate projected at horizon $t_n + s$,
- $\dot{\kappa}_{t_n+s}^{(1)}$, the projected period index at horizon $t_n + s$,

we then have the following relationship for the projection of the mortality rate:

$$\log(\dot{\mu}_{x,t_n+s}) = \alpha_x + \beta_x^{(1)} \dot{\kappa}_{t_n+s}^{(1)}. \quad (4.2.2)$$

Using the notation just defined and the results presented in Table 4.2.1, we can plot the graph of $\dot{\kappa}_{t_n+s}^{(1)}$ with its confidence intervals and the mean of the stochastic projections. Figure 4.2.1 clearly shows a decreasing trend in mortality over the years, driven by technological and medical advances.

From relation (4.2.2), we can thus reconstruct simulations of mortality rates for the coming years. Since these results are stochastic, each simulation in principle yields a different outcome. This is why, when we discuss pricing in Section 4.2.3, it is important to report the quantiles

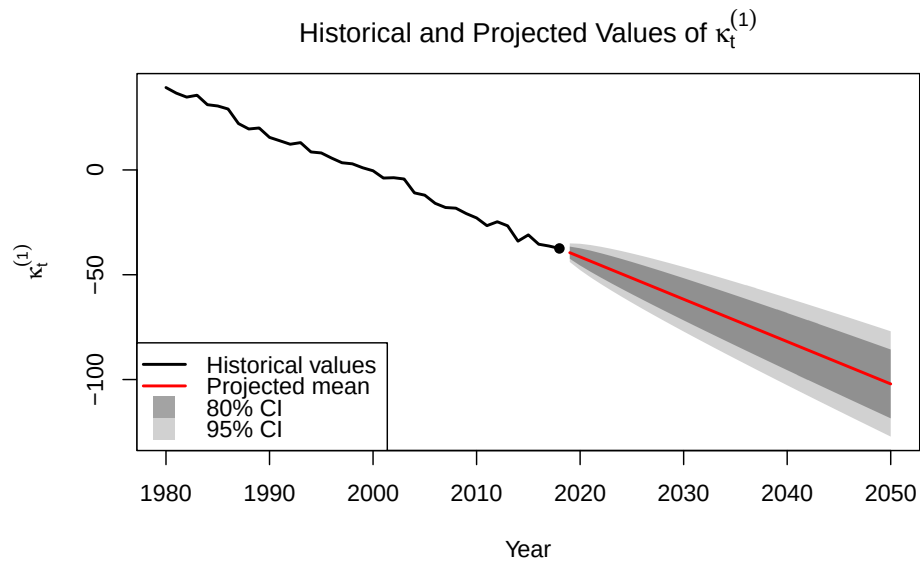


Figure 4.2.1: Historical and Projected Values of $\kappa_t^{(1)}$ using a Random Walk with Drift.

and the standard deviation of the prices obtained. As an illustration, Figure 4.2.2 presents the results of one projected simulation for the years 2019 to 2050. The rainbow-style presentation highlights the smooth decline in mortality rates from one calendar year to the next. This decline is less apparent for very young ages (infant mortality) and for advanced ages (95+).

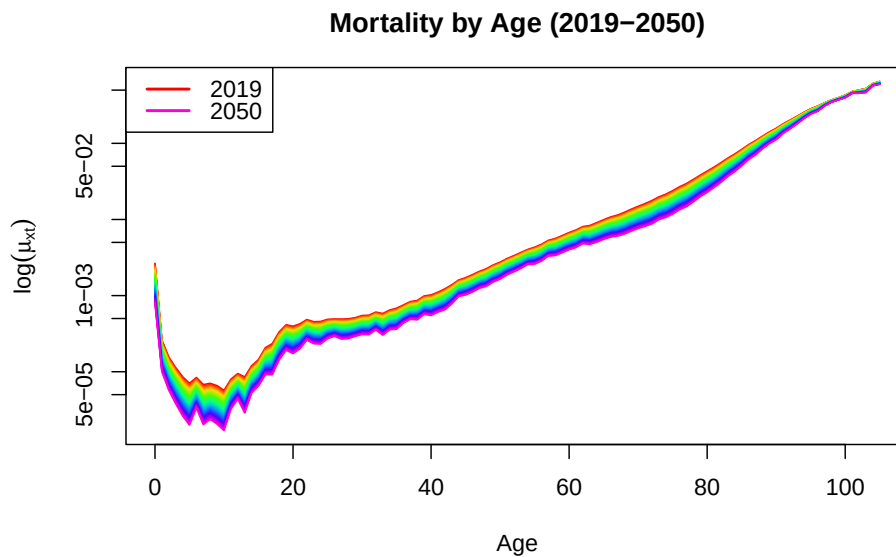


Figure 4.2.2: Projected log-Mortality by Age for the Classical Lee-Carter Model (2019 to 2050).

4.2.2 Climate Lee-Carter Model Projections Using RCPs

As a reminder, the final model proposed for projection is given by (4.1.1). We note that, for a projection to the 2050 horizon, we need – using the notation introduced in the previous section – the values of $\dot{\kappa}_{t_n+s}^{(o)}$ and $\dot{CI}_{t_n+s}$, representing respectively the projection in year $t_n + s$ of the period index capturing the general mortality trend not attributed to climate effects, $\kappa_t^{(o)}$, and the projection in year $t_n + s$ of the climate index incorporating the harvesting effect, CI_t .

For the period index $\kappa_t^{(o)}$, we follow the same methodology as in the previous section. Specifically, we model it as a random walk with drift:

$$\kappa_t^{(o)} = \delta^{\kappa^{(o)}} + \kappa_{t-1}^{(o)} + \xi^{\kappa_t^{(o)}}, \quad \xi^{\kappa_t^{(o)}} \sim \mathcal{N}(0, \sigma_{\xi^{(o)}}). \quad (4.2.3)$$

The resulting estimates are reported in Table 4.2.2, and the projected path of $\dot{\kappa}_{t_n+s}^{(o)}$ with confidence intervals is shown in Figure 4.2.3. We observe that the results are very close to those for the classical Lee-Carter model, which is expected given that $\kappa_t^{(o)}$ is intended to capture the general mortality trend and that mortality attributable to climate represents only a small proportion it.

Parameter	Value
$\sigma_{\xi^{(o)}}$	8.0590
$\delta^{\kappa^{(o)}}$	−2.0995

Table 4.2.2: Estimates for the Random Walk with Drift Parameters of the Modified Climate Lee-Carter.

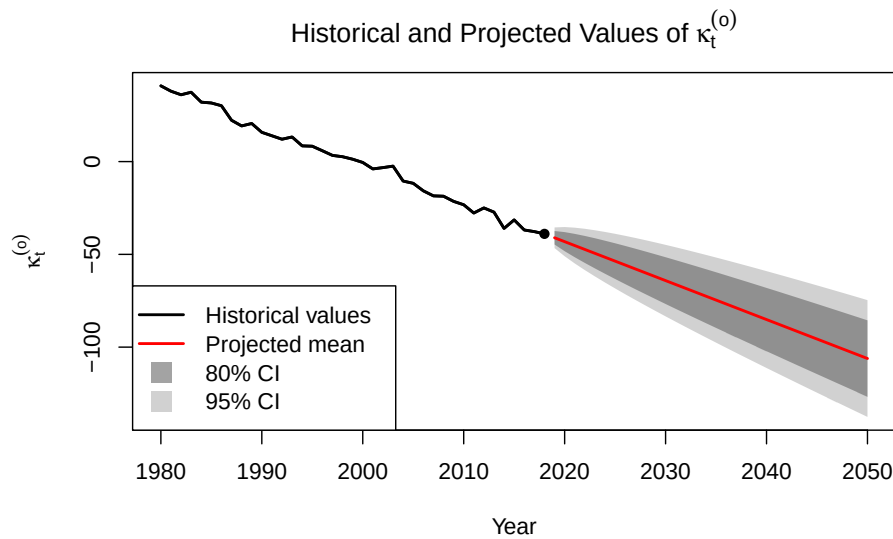


Figure 4.2.3: Historical and Projected Values of $\kappa_t^{(o)}$ using a Random Walk with Drift.

However, for the time series CI_t we chose not to fit a random walk with drift. The idea of fitting a random walk with drift to the time series C_t and CI_t was appealing, given that this was already the method used for the period indices. However, it discards information from the climate projections developed by climatologists – most notably the RCP models described by Jubb et al. (2013). We nevertheless attempted to fit such models to our data. The high variance in the climate indices (in particular due to the 2003 and 2006 heatwaves) renders the approach unusable, as the confidence intervals of the projections become far too wide as we can see in Figure 4.2.4.

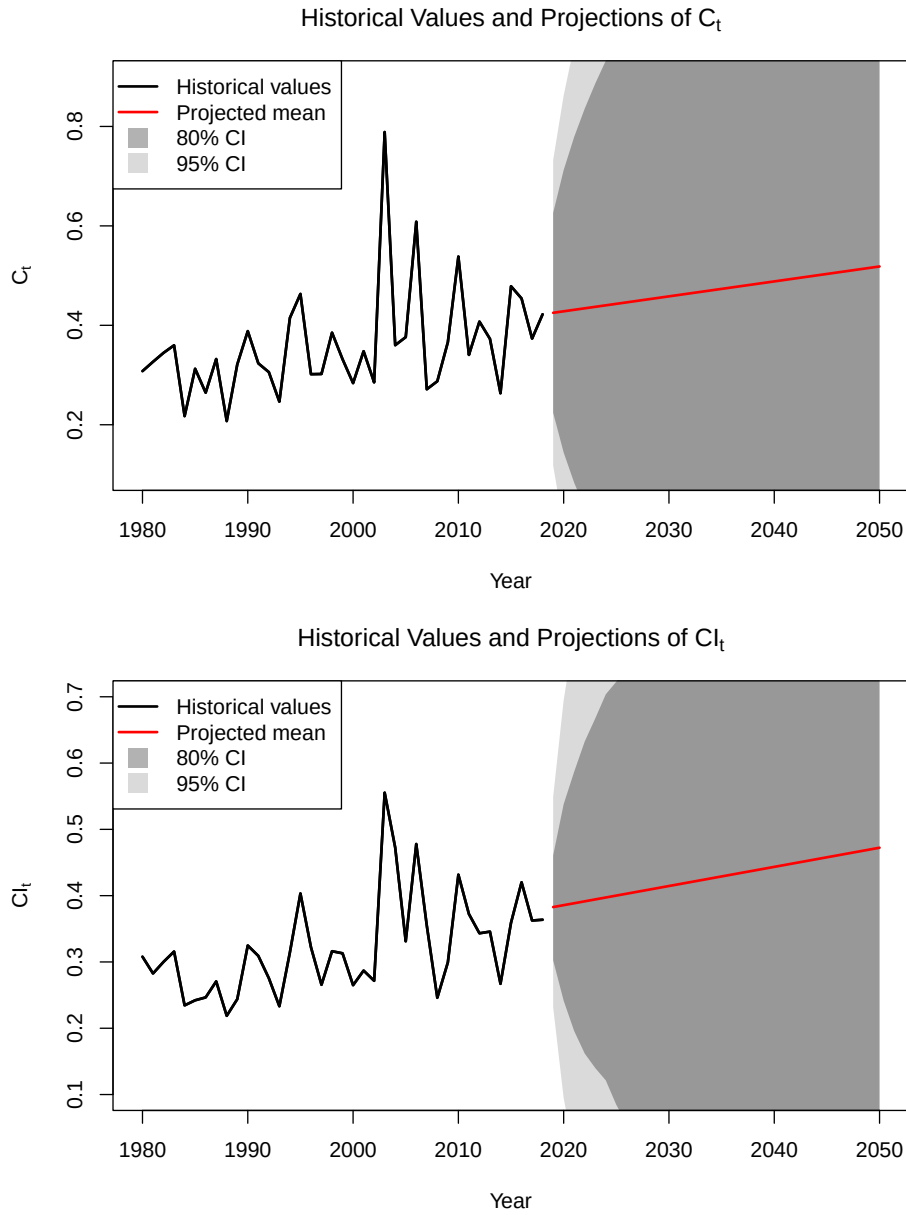


Figure 4.2.4: Historical and Projected Values of CI_t using a Random Walk with Drift.

We therefore choose to rely on a climate scenario proposed in the literature for the remainder of this thesis, viewing it as a simulation input to our model. We use data from a high spatial resolution of 0.025° (≈ 2.8 km) for the ALARO-0 regional climate model coupled to the ‘CNRM-CERFACS-CNRM-CM5’ global climate model, produced by RMIB-UGent¹. These projections are based on the RCP4.5 scenario. Representative Concentration Pathways (RCPs) are standardized climate-change scenarios used to project future atmospheric greenhouse gas concentrations, as explained by Jubb et al. (2013). The dataset provides projections of mean, minimum, and maximum temperature. Using these variables, we can reconstruct the climate variables introduced in Section 2.2 to build the projected climate index. In this spirit, rather than generating a large number of simulations from a random-walk-with-drift model, we use the climatological projection as the scenario mean to predict our climate index up to 2050, which we will then use to forecast mortality rates. Thus, by extending the construction in Section 2.2 using the same climate variables together with projected temperature data up to 2050, we obtain the climate indices shown in Figure 4.2.5. We observe that the projections have difficulty capturing events as rare as the 2003 heatwave; nevertheless, we still observe an upward trend in the climate indices, with or without the harvesting effect.

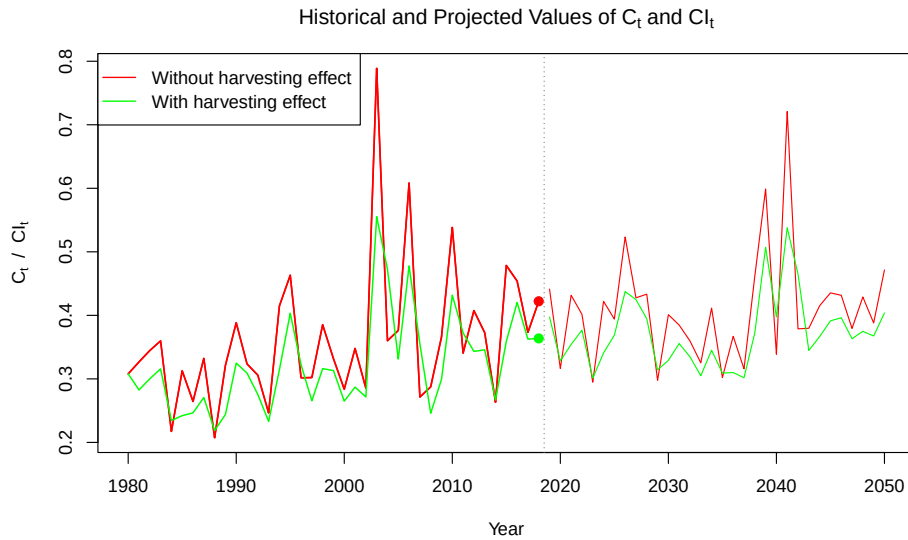


Figure 4.2.5: With and Without Harvesting Effect Projected Climate Index.

Having now the necessary projection data, we can reconstruct the mortality-rate projection from model (4.1.1), which yields:

$$\log(\dot{\mu}_{x,t_n+s}) = \alpha_x + \beta_x^{(o)} \dot{\kappa}_{t_n+s}^{(o)} + \delta_x \dot{C} I_{t_n+s}, \quad (4.2.4)$$

where δ_x was defined in Section 4.1.1. From this relation, we can reconstruct simulations of mortality rates for the coming years. Because the results are stochastic – at least for $\dot{\kappa}_{t_n+s}^{(o)}$ – each simulation, in principle, yields a different outcome. Accordingly, when we discuss pricing in Section 4.2.3, it is important to report the quantiles and the standard deviation of the

¹This climate dataset is openly available at <https://publish.geo.be/geonetwork/srv/api/records/22dc62f0-ca7f-11ee-9a71-847b573ec00f>.

resulting prices, as explained before. As in the previous section, Figure 4.2.6 presents the results of a single projected simulation for the years 2019-2050. The difference relative to Figure 4.2.2 is too small to be visually apparent in these figures; we therefore proceed with a numerical approach in Section 4.2.3.

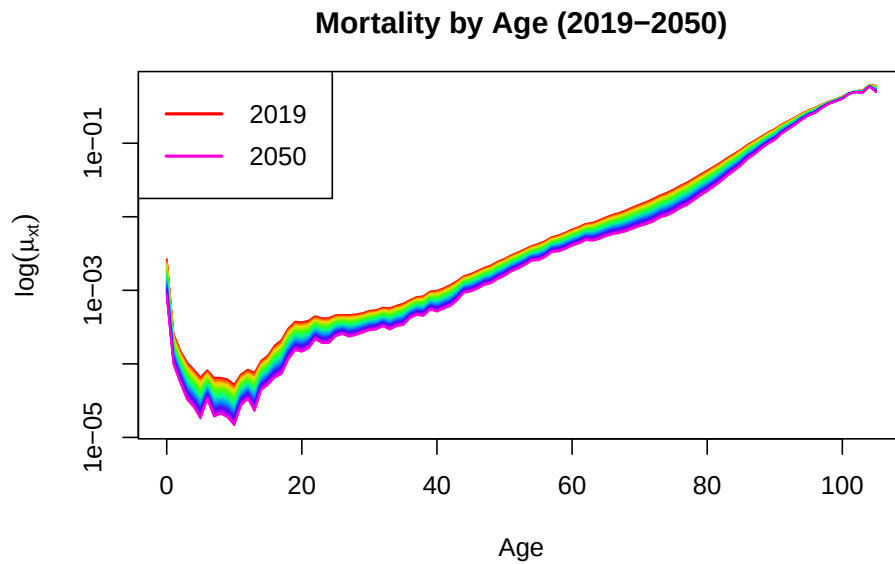


Figure 4.2.6: Projected log-Mortality by Age for the Proposed Model (2019 to 2050).

4.2.3 Impact on Longevity and Life Insurance Pricing

We begin by recalling a few actuarial notions from life insurance. We follow the definitions and notations in Hainaut (2024b). Let us first introduce some key quantities:

- ${}_tq_x$: the probability of dying before age $x + t$, with the convention $q_x := {}_1q_x$,
- ${}_tp_x = 1 - {}_tq_x$: the probability of surviving to age $x + t$, with the convention $p_x := {}_1p_x$,
- ω : the maximum human life span; here we take $\omega = 105$,
- $v = \frac{1}{1+i}$: the discount factor, where i is the contract's technical interest rate,

we will also use the following relationship:

$$\begin{aligned} {}_tp_x &= p_x \cdot p_{x+1} \cdot \dots \cdot p_{x+t-1} \\ &= (1 - q_x) \cdot (1 - q_{x+1}) \cdot \dots \cdot (1 - q_{x+t-1}) \\ &= \prod_{i=0}^{t-1} (1 - q_{x+i}). \end{aligned} \tag{4.2.5}$$

Because we assume, following Section 1.1.1, that the force of mortality is piecewise constant within each year, we have the approximation

$$p_x = \exp(-\mu_x) \iff (1 - q_x) = \exp(-\mu_x). \tag{4.2.6}$$

Combining (4.2.5) and (4.2.6) yields

$${}_tp_x = \prod_{i=0}^{t-1} \exp(-\mu_{x+i}), \tag{4.2.7}$$

an identity that will be useful for computing life expectancy and pricing life insurance products.

Following Hainaut (2024b), under the assumption that μ_{x+t} is piecewise constant on $[t, t+1]$ for $t \in \mathbb{N}$ and using (4.2.7), the life expectancy at age x , denoted $\mathring{e}_x$, is approximated by:

$$\begin{aligned} \mathring{e}_x &= \sum_{t=1}^{\omega-x} {}_tp_x \\ &= \sum_{t=1}^{\omega-x} \prod_{i=0}^{t-1} \exp(-\mu_{x+i}). \end{aligned} \tag{4.2.8}$$

From there, using our projection models defined in (4.2.2) and (4.2.4), we simulate a large number of paths, 10 000, to determine the mean life expectancy at ages 50, 60, and 70 for calendar years 2025–2050 under both models. The results of these projections are reported in Table 4.2.3. We observe for both models an increase in life expectancy over calendar years, driven by ongoing technical and medical improvements. However, it is noteworthy that this increase is slower for the modified Lee–Carter model that incorporates the climate index and thus climate effects on mortality. In particular, life expectancy is slightly lower under this model, with a difference that is somewhat more pronounced at older ages. We next examine

Year	Classical LC			Climate LC		
	$\check{e}_{50}$	$\check{e}_{60}$	$\check{e}_{70}$	$\check{e}_{50}$	$\check{e}_{60}$	$\check{e}_{70}$
2025	33.62	24.75	16.53	33.58	24.70	16.47
2026	33.76	24.88	16.64	33.74	24.84	16.59
2027	33.90	25.01	16.75	33.88	24.97	16.70
2028	34.05	25.14	16.86	34.02	25.10	16.80
2029	34.19	25.27	16.97	34.15	25.21	16.89
2030	34.33	25.40	17.07	34.29	25.35	17.00
2031	34.47	25.53	17.18	34.44	25.48	17.12
2032	34.61	25.65	17.28	34.58	25.61	17.22
2033	34.75	25.78	17.39	34.71	25.73	17.32
2034	34.89	25.91	17.49	34.86	25.86	17.43
2035	35.02	26.03	17.59	34.99	25.98	17.53
2036	35.16	26.15	17.70	35.12	26.10	17.63
2037	35.29	26.28	17.80	35.26	26.23	17.73
2038	35.43	26.40	17.90	35.40	26.36	17.85
2039	35.56	26.52	18.00	35.56	26.51	17.98
2040	35.69	26.64	18.10	35.67	26.61	18.06
2041	35.82	26.76	18.20	35.83	26.76	18.18
2042	35.95	26.88	18.29	35.95	26.86	18.27
2043	36.08	26.99	18.39	36.06	26.96	18.34
2044	36.20	27.11	18.49	36.19	27.08	18.44
2045	36.33	27.23	18.58	36.32	27.20	18.55
2046	36.46	27.34	18.68	36.45	27.32	18.64
2047	36.58	27.46	18.77	36.57	27.43	18.73
2048	36.70	27.57	18.87	36.69	27.54	18.83
2049	36.83	27.68	18.96	36.82	27.65	18.92
2050	36.95	27.79	19.05	36.94	27.77	19.02

Table 4.2.3: Projected Complete Life Expectations (in years) for Ages 50, 60, and 70 under the Classical LC and Proposed Model (2025–2050).

whether the impact on the pricing of a basic life insurance product is negligible or not.

Following Hainaut (2024b), an endowment with term n pays one monetary unit only if the insured is alive at age $x + n$. The net single premium is therefore:

$$\begin{aligned}
 {}_nE_x &= v^n {}_np_x \\
 &= v^n \prod_{i=0}^{n-1} \exp(-\mu_{x+i}).
 \end{aligned} \tag{4.2.9}$$

Consider a pure endowment with capital $C = 100\,000\text{€}$, technical interest rate $i = 2\%$, and term $n = 25$ years, offered to insureds aged 40, 50, and 60. We then simulate 10 000 projected paths under the two models (4.2.2) and (4.2.4) to estimate the net single premium,

reporting the sample mean, standard deviation, and selected quantiles. Note that we adopt a prospective calculation approach, i.e., we use the individual's mortality rate from the calendar year in which they will reach the relevant age. For example, for an individual aged 40 in 2025, to compute the death probability at age 60 we use the projected rate $\hat{\mu}_{60,2045}$ so as to make full use of our projections. The results are reported in Tables 4.2.4 and 4.2.5. These statistics appear coherent, given that the price of the pure endowment decreases with age in both models. Indeed, since the benefits are weighted by survival probabilities, the older the insured, the lower the net single premium. On the other hand, the results are very close for the two models. More specifically, Table 4.2.4 shows that prices are slightly lower for the modified Lee-Carter model starting at ages where heat begins to affect health (50 and 60 for a 25-year term). Conversely, Table 4.2.5 indicates that this price difference, expressed in relative terms, is below 1% for the three ages considered. A relative difference of less than 1% is often regarded as negligible.

Model	Age	Mean	sd	q5%	Median	q95%
Classical LC	40	56 907.60	318.64	56 360.13	56 924.82	57 409.36
Classical LC	50	52 496.19	740.97	51 215.94	52 539.76	53 654.82
Classical LC	60	40 917.55	1 605.85	38 168.50	40 990.50	43 480.31
Climate LC	40	56 910.17	385.96	56 236.26	56 933.53	57 514.81
Climate LC	50	52 475.10	897.77	50 900.92	52 534.41	53 868.32
Climate LC	60	40 571.48	1 971.48	37 161.44	40 671.74	43 693.04

Table 4.2.4: Net single premium statistics (pure endowment, $t_0 = 2025$, $n = 25$, $i = 2\%$, $C = 100,000$).

Age	Δ Mean (€)	Δ Mean (%)
40	+2.57	+0.00%
50	-21.09	-0.04%
60	-346.07	-0.85%

Table 4.2.5: Mean Premium Difference: Proposed Model – Classical LC.

Overall, although the modified Lee-Carter model that incorporates climatic variations does capture the slight loss of longevity – resulting in slower growth in life expectancy than the classical Lee-Carter model, as shown in Table 4.2.3 – the relative pricing difference for a basic life insurance product is negligible ($< 1\%$) in our analysis.

Conclusion

This thesis set out to assess whether global warming has a non-negligible impact on mortality in Belgium and, if so, to quantify that impact in a way that is usable for forward-looking actuarial work. The methodology followed a pragmatic route: start from established stochastic longevity models and integrate climate information in a parsimonious and identifiable manner, with the specific aim of isolating a climate-related signal robust enough to inform projections up to 2050. Two complementary approaches were considered. First, a climate-augmented Lee–Carter specification in the spirit of Boumezoued et al. (2022) embedded into the GAPC framework (RD Lee and LR Carter, 1992; Villegas et al., 2018). Second, a model based on the SPO family (Skelecka et al., 2017; Plat, 2009) following Milhaud and Gouthon (2024), which introduces a triggering mechanism for hot summers and an age threshold beyond which heat becomes detrimental.

A first result concerns the construction and signal content of the climate index tailored to Belgian conditions. From a broad set of meteorological covariates, we retained a parsimonious quartet – summer minimum temperature, counts of summer days above 34°C and 38°C, and the number of winter days below 0°C – that all enter significantly in the linear specification. The resulting index explains a substantial share of the variance in cause-specific mortality attributable to exposure to environmental temperatures, with an $R^2 \approx 64\%$. Despite several attempts to enrich the specification, this level proved difficult to improve without overfitting, which is consistent with the episodic nature of heatwaves and cold spells documented in multi-country evidence (Gasparrini et al., 2015). Incorporating a harvesting adjustment moderates the year-on-year oscillations of the index and avoids attributing to climate what partly reflects temporal displacement of deaths as shown in Figures 2.2.4 and 2.3.3.

When embedded in a Lee–Carter structure, the climate index improves the fit in a way that is both statistically disciplined and actuarially relevant. Relative to the classical benchmark, the climate model yields lower information criteria on Belgian data (1980–2018): the AIC decreases by about 692 points and the BIC by about 10 points, despite the penalty for additional parameters introduced by the climate term. At a finer level, the gain concentrates where it matters most for pricing: older ages. The proportion of ages for which the R^2 improves is 31.13% over all ages, but rises to 54.35% above age 60, 69.44% above age 70, and 96.15% above age 80. At age 85, both specifications track observed mortality very closely, with a slight but consistent edge for the climate model. These diagnostics support the view that isolating a climate-related component within a Lee–Carter framework is justified. (See Table 2.4.1 for AIC/BIC and Figures 2.4.1 and 2.4.2 for age-wise diagnostics.)

The second route, based on Milhaud and Gouthon (2024), complements this picture by separating the non-linear age structure of mortality from temperature-related excess mortality via an explicit trigger and an age threshold estimated from weekly summer mortality (STMF) and

temperature data. This construction is appealing for interpretation, in particular when stress testing sequences of atypically hot summers, and it connects naturally to the observation – already visible in the Lee–Carter extension – that the climate signal is primarily age-modulated and concentrated among seniors. In our Belgian application, we were able to identify a statistically significant threshold age $A_h = 75$ years using weekly data and Pearson-type tests, which confirms the existence of a pivot from which heat becomes detrimental. However, the subsequent calibration of the full trigger-based specification did not translate into relevant data: the hot-summer term $\mathbf{1}\{I_t - \bar{I}_t > 0\}\kappa_t^{(5)}$ failed to load robustly and the model did not capture the excess mortality of the warmest summers as Shown in Figure 3.3.3. The most plausible reasons are the short weekly sample and low counts at advanced ages (small-country setting). For these reasons, in projection we retain the climate-augmented Lee–Carter as the operational baseline, while viewing the trigger-based route as complementary and promising.

On the prospective side, we mapped a climate scenario (RCP) into trajectories for the climate index and then into mortality rates, while retaining a classical random walk with drift for the baseline period index in the absence of the climate term to maintain a clear reference point. The resulting projections to 2050 indicate that temperature-related mortality acts as a countervailing force against secular improvements: the older the age and the warmer the assumed path, the more pronounced the offset. Insurance quantities do not point to large effects. In our projections, the climate-augmented paths yield a slight reduction in life expectancy relative to the classical baseline, but the impact on premium factors for a basic life product remains negligible in actuarial terms – consistently below 1% across the scenarios examined as shown in Tables 4.2.3 and 4.2.5.

Several limitations follow directly from the data and the modelling choices, and they naturally suggest avenues for future work. First, the climate-augmented Lee–Carter (in the spirit of Boumezoued et al. (2022)) would work better for projection with a specification closer to Milhaud and Gouthon (2024) with parametric functions in age.

Second, this thesis focuses on temperature (heat and cold) as the primary channel, in line with the dominant literature on temperature–mortality relationships (Gasparrini et al., 2015). Other pathways – humidity and heat stress, air quality, the frequency and severity of natural disasters, and climate-sensitive diseases (e.g., mosquito-borne) – also matter. Extending the climate index to include such covariates, or building multi-component indices, would provide a more comprehensive view, provided parsimony and interpretability are preserved so that projections remain usable for pricing.

Third, matching weather and mortality is constrained by a scale trade-off. In large countries, mortality data are abundant but climate is heterogeneous; in small countries, climate is more uniform but mortality counts (especially by age and week) are lower. Belgium lies toward the latter end. This argues either for hierarchical or pooling strategies that borrow strength across regions while aligning exposures to local climate, or for carefully designed national indices that are robust to spatial aggregation.

A further limitation concerns cause-of-death measurement and sample choice. The cause-specific series used to isolate “*exposure to environmental temperatures*” carry measurement uncertainty, and we restricted the window to periods where Belgian data are most reliable. This avoids spurious precision but leaves open the question of stability under alternative samples and definitions. Sensitivity analyses to age ranges, training windows, and lag structures (including harvesting) should therefore accompany projection use.

Finally, there is scenario and dynamics uncertainty. While we illustrate the mapping from RCP narratives to mortality through the climate index, a broader set of climate scenarios should be tested. Scenario design ought to reflect not only median trends but also the clustering of hot summers, given their outsize contribution to excess mortality at older ages. In addition, the period indices that inherit climate inputs may warrant richer time-series dynamics than a random walk with drift.

For actuarial applications, the main takeaway is not that climate overturns the secular improvement in longevity, but that it introduces an age-dependent counter-signal whose materiality increases with warming and with clusters of hot summers. In pricing and reserving, this calls for explicit documentation of whether and how climate-sensitive mortality enters the projection basis, sensitivity analysis to the climate term, and monitoring of incoming summers' temperatures and excess mortality as an early warning for recalibration.

In conclusion, this thesis shows that climate information can be integrated into mainstream stochastic mortality models for Belgium in a way that is both statistically disciplined and operationally usable. The climate-augmented Lee–Carter and the trigger-based extension of the SPO framework capture complementary aspects of temperature impacts on mortality and, when coupled with a carefully constructed index, deliver projections to 2050 that translate into interpretable insurance metrics. Immediate extensions include broadening the climatic covariate set beyond temperature, testing alternative RCPs and time-series dynamics, and developing more parametric age structures Lee–Carter augmented. These steps would sharpen the quantification of climate impacts on longevity and strengthen the link between climate narratives and pricing, reserving, and risk monitoring for Belgian portfolios.

Bibliography

- Akram, S. and Ann, Q. U. (2015). Newton raphson method. *International Journal of Scientific & Engineering Research*, 6(7):1748–1752.
- Akritis, A. G. and Malaschonok, G. I. (2004). Applications of singular-value decomposition (svd). *Mathematics and computers in simulation*, 67(1-2):15–31.
- Atance, D., Lledó, J., and Navarro, E. (2025). Modelling and forecasting mortality rates for a life insurance portfolio. *Risk Management*, 27(1):3.
- Bador, M. (2016). *Les changements d’extrêmes de température en Europe: records, canicules intenses et influence anthropique*. PhD thesis, Université Paul Sabatier-Toulouse III.
- Boumezoued, A., Elfassihi, A., Germain, V., and Titon, E. (2022). Modelling the impact of climate risks on mortality. *Milliman White Paper*.
- Florides, G. A. and Christodoulides, P. (2009). Global warming and carbon dioxide through sciences. *Environment international*, 35(2):390–401.
- Gasparrini, A., Guo, Y., Hashizume, M., Lavigne, E., Zanobetti, A., Schwartz, J., Tobias, A., Tong, S., Rocklöv, J., Forsberg, B., Leone, M., De Sario, M., Belle, M. L., Guo, Y.-L. L., Wu, C.-F., Kan, H., Yi, S.-M., Coelho, M. d. S. Z. S., Saldiva, P. H. N., Honda, Y., Kim, H., and Armstrong, B. (2015). Mortality risk attributable to high and low ambient temperatures: a multicountry observational study. *The Lancet*, 386:369–375.
- Guibert, Q., Pincemin, G., and Planchet, F. (2024). Impact of climate change on mortality: An extrapolation of temperature effects based on time series data in france. *arXiv*, 2406.02054v2.
- Hainaut, D. (2024a). LACTU 2030: introduction to static and prospective mortality tables. Lecture slides, Université Catholique de Louvain. Accessed during the course LACTU 2030 : Actuariat de l’assurance-vie.
- Hainaut, D. (2024b). LACTU 2030: Life Insurance. Lecture slides, Université Catholique de Louvain. Accessed during the course LACTU 2030 : Actuariat de l’assurance-vie.
- Hainaut, D. (2024c). LACTU 2170: Financial valuation of actuarial liabilities. Lecture slides, Université Catholique de Louvain. Accessed during the course LACTU 2170 : Financial valuation of actuarial liabilities.
- Hales, S., Kovats, S., Lloyd, S., Campbell-Lendrum, D., Security, W. H. O. H., and Cluster, E. (2014). *Quantitative Risk Assessment of the Effects of Climate Change on Selected Causes of Death, 2030s and 2050s*. World Health Organization.

- Hitz, S. and Smith, J. (2004). Estimating global impacts from climate change. *Global Environmental Change*, 14(3):201–218.
- Jubb, I., Canadell, P., and Dix, M. (2013). Representative concentration pathways (RCPs). *Australian Government, Department of the Environment*.
- Kuha, J. (2004). Aic and bic: Comparisons of assumptions and performance. *Sociological methods & research*, 33(2):188–229.
- Milhaud, X. and Gouthon, J.-L. (2024). Impact of heat waves on mortality: extension of longevity models to account for global warming. *HAL open science*, hal-04831333.
- Ozer, D. J. (1985). Correlation and the coefficient of determination. *Psychological bulletin*, 97(2):307.
- Pearson, K. (1985). Vii. note on regression and inheritance in the case of two parents. *The Royal Society of London*, 58:240–242.
- Plat, R. (2009). On Stochastic Mortality Modeling. *Insurance: Mathematics and Economics*, 45(3):393–404.
- Seklecka, M., Pantelous, A. A., and O’Hare, C. (2017). Mortality effects of temperature changes in the United Kingdom. *Journal of Forecasting*, 36:824–841.
- RD Lee and LR Carter (1992). Modeling and forecasting U.S. mortality. *Journal of the American Statistical Association*, 87(419):659–671.
- Tol, R. S. (2018). The economic impacts of climate change. *Review of environmental economics and policy*.
- Toulemon, L. and Barbieri, M. (2008). The mortality impact of the august 2003 heat wave in france: investigating the ‘harvesting’effect and other long-term consequences. *Population studies*, 62(1):39–53.
- Villegas, A. M., Millossovich, P., and Kaishev, V. K. (2018). StMoMo: An R package for stochastic mortality modeling. *Journal of Statistical Software*, (15):1–38.
- von Sachs, R. (2024). LSTAT 2170: Time Series Analysis. Lecture slides, Université Catholique de Louvain. Accessed during the course LSTAT 2170 : Time Series Analysis.
- Zhang, Z. (2016). Variable selection with stepwise and best subset approaches. *Annals of translational medicine*, 4(7):136.

