UCLouvain

FACULTE DES SCIENCES

Faculté des sciences

Extreme Value Analysis of Mortality at Age Beyond 105

Auteure : Lixia YAN

Promoteur : Michel Denuit

Lecteur : Jean-François Walhin

Table des matières

1 Introduction

Since last century, the life expectancy is increasing across the world and the number of people attained the age of 100 years or more is expected to grow rapidly in the future. The United Nations (Ageing in the 21st Century n.d.) estimates that there were 343,000 centenarians worldwide in 2012, and by 2050, this figure is projected to reach 3.2 million. How to assess improvements in mortality at extremely high ages are of great interest to pension funds, annuity providers, social care and health care providers and life insurance companies.

The analysis on the mortality at extremely high ages were few due to the difficulty to obtain reliable mortality data for very old people. It was until 2001 that the International Database on Longevity (IDL) was established at the University of Montpellier, aiming to provide validated data on the semi-supercentenarians (>105) and the supercentenarians (>110). Numerous researches have been done based on the data from IDL. However, the evolution of the mortality at extreme ages has been a contentious topic. Belzile et al. (2020) analyzed the mortality of Italian and French semi-supercentenarians as of October 2019 and concluded that beyond age 108, the mortality rates remain constant and no difference are found between countries and cohorts. Their results also suggested that the upper bound of human lifespan is too large to be approached. The findings are consistent with the previous study conducted by Barbi et al. (2018) that the estimated hazard rates approach a plateau beyond age 105 based on the inhabitants of Italy aged 105 and older between 2009 and 2015. Other studies show differing views as to whether the mortality rates stay constant at the late stages of life. Gbari et al. (2017) based the study on the Belgium population aged 95+ and concluded no evidence of levelling-off in the force of mortality at the oldest ages. Gavrilova, Gavrilov, and Krut'ko (2017) demonstrated that based on the 1884-1894 birth cohorts, the hazard rates after age 110 do not stay constant and a Gompertz Model fits better to the data than the exponential model (flat mortality). Their results also suggested the mortality deceleration at oldest ages is not a universal phenomenon.

Different methods have been used to analyze the mortality at the extreme ages. Many studies predict the mortality at extremely high ages via the hazard rates, others use a less populate method called the quantile regression method to detect the trend in the ages of death of a

chosen cohorts (Medford et al. 2019). But the extrapolation beyond the data is not possible with these methods. A more sophisticated method using the “extreme value theory (EVT)” has been applied in many recent studies. Li, Ng, and Chan (2011) has integrated the EVT in a threshold life table(TLT) method to study the old-age mortality risk for Australia and New Zealand populations. The same method has been used by Bravo and Corte-Real (2012) to model the longevity risk of the Portuguese and Spanish Population. Huang, Maller, and Ning (2020) modified the TLT into the “smoothed threshold life table” (STLT) model to make the mortality transition from non-extreme to extreme ages more smoothly. Gbari et al. (2017) has also applied the EVT to study the individual ages at death of the Belgium population who was born between 1886 and 1904. Similar approach can also be found in the study of the upper limit of lifetime distribution of the Japanese population (Hanayama and Sibuya 2016), the Dutch residents born in Netherlands died between the year 1986 and 2015 (Einmahl, Einmahl, and Haan 2019) and the annual deaths at advanced age in Canada from 1949 to 1997(Kathryn A. Watts MSc and Bruce L. Jones 2006).

In practice, there are 2 approaches to apply the EVT: the r-largest approach (Block maxima) and the peak-over-threshold (POT) approach. It is believed that the POT approach gives better estimation as it uses all the information beyond the high threshold while the r-large approach considers only the highest data information. That is the reason why in our study, we will concentrate on the POT approach. The difficulty of such approach is to choose the right threshold beyond which, the data’s behavior will follow a generalized Pareto distribution. A detailed selection methods can be found in (Scarrott and MacDonald 2012). If the graphical tools have been largely used to have an idea about the range in which the threshold can be situated, several studies prefer using an automated selection procedure which is more robust and objective.(G’Sell et al. 2015) (Bader, Yan, and Zhang 2018) (Gbari et al. 2017) Other methods based on the goodness-of-fit tests were also used to help to choose the thresholds.(Choulakian and Stephens 2001) In this study, we will use the graphical tools, the goodness-of-fit tests and the automated selection procedure to decide the final threshold.

The aim of the paper is to analyze mortality of semi-supercentenarians (105+) of France, Germany, Belgium, and Switzerland using the method of EVT. All the datasets are collected from the IDL databases as of December 2019. The outline of the paper is as follows. Section 2

gives a brief description about the data. Section 3 recalls the fundamental results of EVT and presents their application in the extreme ages mortality analysis. The analysis results are showed in section 4. Special intention is given to the problem of homogeneity when applying the EVT (i.e.: a categorical covariate “nationality” is introduced). Section 5 is devoted to an application of the model to estimate the one-year force of mortality and the survival probability at age beyond 105. Section 6 summarizes the main findings and briefly concludes.

2 Description of the data

2.1 Data source

The data of our study is composed of the records of deaths at 105+ of France, Germany, Belgium, and Switzerland. All the data are collected from the "the International Database on Longevity (IDL)" database as of December 2019. IDL provides thoroughly validated information on individuals who have attained extreme ages at death (i.e.: beyond age 105) for research purpose. For each of the 4 countries, there are two data files available:

- List of individuals dead at ages 110 and older;
- List of individuals dead at age 105-109.

Each record in the data file contains information including the individual's gender, birth date, death date, birth country, death country, etc.

Table 2-1 gives a general overview of the data. France has the largest number of individuals living to 105 and above. In each country, people surviving beyond 110 are much less than the people living to 105. There are only 4 observations of individuals living to 110 or more in Switzerland.

		Belgium	France	German	Switzerland
Death age above 110	Year of Birth	1878-1904	1875-1907	1883-1894	1881-1890
	Year of Death	1990-2015	1987-2017	1994-2005	1993-2005
	Counts	19	241	16	4
Death age between 105 to 109	Year of Birth	1870-1910	1870-1912	1881-1898	1864-1900
	Year of Death	1977-2015	1978-2017	1989-2005	1971-2005
	Counts	777	9612	928	236

Table 2-1. Observed death above 105 in Belgium, France, German, Switzerland from IDL (last updated: 15/12/2018)

Our analysis focus on the birth cohorts from 1881 to 1898 because all the four countries have the observations within these cohorts. Another reason for this choice is that these birth cohorts are extinct cohorts meaning that all the people born in these cohorts have already died, so no information is lost in terms of the numbers of death in these birth cohorts.

2.2 Data validation process

In terms of the data validation, according to the IDL, all the data have been validated in such a way that age-ascertainment bias is avoided. Indeed, the data has been assigned 2 validation

levels: level A indicates full validation that at least one of the early life document and a death certificate has been checked; level B is assigned to the individuals from the countries that do not document as thoroughly as full validation requires, but individual cases have been carefully checked. As a result, we can use all the data from the IDL and be confident in the reliability of the age of death recorded in the database.

2.3 Descriptive statistics

Firstly, the basic descriptive statistics of the datasets are given in Table 2-2

	Belgium		France		German		Switzerland	
	F	**M**	**F**	**M**	**F**	**M**	**F**	**M**
Number of observations	269	36	2823	255	835	109	165	31
Maximum age of death	112,51	110,22	115,12	111,88	112,99	111,76	112,39	108,74
Average age of death	106,49	105,98	106,58	106,34	106,37	106,39	106,44	106,22
Median age of death	106,12	105,67	106,19	105,98	106,07	105,97	106,01	105,86
Standard deviation	1,36	1,04	1,47	1,27	1,21	1,27	1,33	1,13

Table 2-2. Basic description statistics for age at death above 105, cohorts 1881-1898

For the birth cohorts 1881-1898, in each country, the number of observations for women is larger than the number of observations for men. The maximum age of death observed for women and men is 115,12 and 111,88 respectively, both in France. The average age of death is around 106 years old, no big difference is observed between the countries, neither between women and men. The median age of death for women is slightly higher for woman than that for men in all the counties. At last, the standard deviation shows that the datasets are stable. The boxplots classified by gender are displayed in Figure 2-1. (BEL stands for Belgium, FRA for France, DEU for German, CHE for Switzerland.)

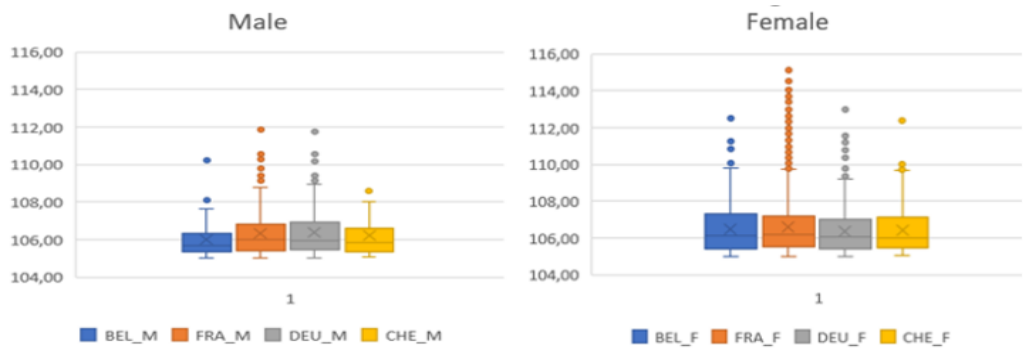

Figure 2-1. Boxplots for male and female datasets

For both female and male datasets, most of the death happened before the average and the median age, while the death beyond the average and the median death age spread on a wider range of value. This trend can also be seen on the histograms in Figure 2-2:

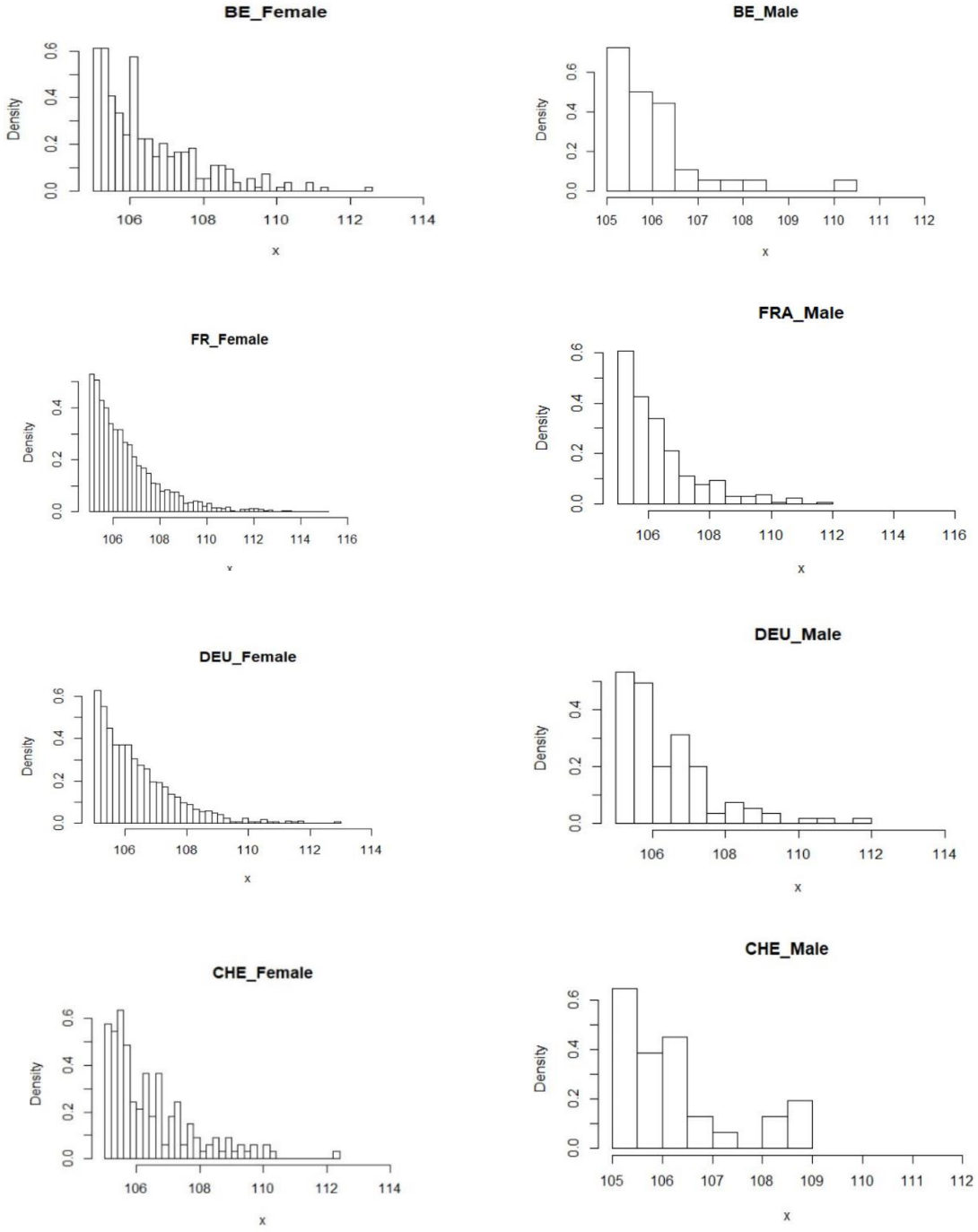

Figure 2-2. Histograms for women and men, classified by countries and by gender, birth cohorts 1881-1898

All the histograms show an extremely left skewed and long decreasing pattern in the right tail. A smooth decreasing trend can be observed on the histogram for French female and German female datasets, while all the other histograms show a certain irregularity of decreasing. The main reason may be due to the size of the sample which is not sufficiently large that the numbers observed on a relatively small scale can be volatile.

Secondly, in terms of the trend throughout the birth cohorts from 1881 to 1898, Table 2-3 summarizes the numbers of observation above 105 in each country, grouped in every three years. A clear increasing trend can be seen on the numbers of observation throughout these cohorts for women. For men, this trend is not evident, especially when the sample size is small.

	Belgium		France		German		Switzerland	
	F	M	F	M	F	M	F	M
1881-1883	22	1	197	22	12	4	15	2
1884-1886	30	8	260	27	85	11	21	12
1887-1889	41	4	412	36	133	19	23	5
1890-1892	36	4	479	46	178	27	31	3
1893-1895	64	11	665	58	232	29	29	6
1896-1898	76	8	810	66	195	19	46	3
Total	269	36	2823	255	835	109	165	31

Table 2-3. Numbers of observations of death age above 105 in each country, cohorts 1881 to 1898

Thirdly, the shifted expectancy and the maximum death age of the birth cohorts 1881-1898 by countries and by gender is displayed in Figure 2-3. A linear trend line is added to the shifted expectancy and to the maximum age of death. Again, due to the limited size of the men sample from Belgium, Germany and Switzerland, their maximum and the average are remarkably close to each other. The linear trend line of the maximum age of death attained show a certain degree of increase for almost all the countries over the cohorts except for the Switzerland due to the high observations in 1881, while the linear trend line of the average age of death beyond 105 are nearly horizontal for all the countries and gender from 1881 to 1898.

The main finding according to the basic statistics can be summarized as below:

- The age of death for women and men attaining 105 or above within each country observed has a long right tail distribution with the mean around 106 years old.

- It seems no important improvement in terms of the average age of death for people attaining 105 or above through the birth cohorts 1881 to 1891.
- The observed maximum age of death seems to show a slightly increasing trend for the female groups, but not that evident for the male groups.

Figure 2-3. Shift expectancy and maximum of age of death by county and by sex, birth cohorts 1881 – 1898

Does the maximum age of human will be continually improved in the future? Or does there exist a maximum age that could never be surpassed? Based on the basic statics on the observations, these questions cannot be solved. A more sophistic analyze method using “the extreme value theory” will help to answer these questions.

3 Methodology

In this section, the fundamental results of EVT are recalled in section 3.1. Its application in the context of lifetime analysis is addressed in section 3.2.

The studies of extremes events behavior can be found as early as in the beginning of the 20th century. The probabilistic extreme value theory was developed by M. Fréchet (1927), R. Fisher and L. Tippett (1928), R.von Mises (1936), and B.Gnedendo (1943). The statistical theory was initiated by J.Pickands III (1975).

The extreme value theory (EVT) is the principal theory used in analyzing the behavior of the data far in the tails of distributions. Based on the part of the sample carrying the information about the extremal behavior (i.e. the largest or the smallest sample values), the EVT provides a solid theoretical basis and framework for extrapolation beyond that range of the data, where there are often very few observations or not at all. Indeed, the EVT has a wide variety of applications in different disciplines as far as the subject of interest concerns the extreme events. The earliest application can be found in 1877, the government of Netherlands wanted to know how high the dikes should be so that the probability of a flood within a given year is less than 0.0001 (de Haan and Ferreira 2006). Nowadays, as in many other disciplines such as climatology and hydrology, the EVT is a handy tool for actuaries to estimate the claims severity for the largest losses, to do the pricing of the excess-of-loss reinsurance treaties, to calculate the possible maximum loss (PML), etc.

3.1 Fundamental results of Extreme Value Theory

The extreme value theory (EVT) is considered as analogues to the central limited theory (CLT). Indeed, when the sample size grows, the EVT studies the behavior of "extreme" events beyond a certain threshold in the tail while the CLT examines the characters of the accumulation of the "normal" events.

Let $X_1, X_2, X_3 \ldots$ be a sequence of independent random variables with common distribution function $\mathrm{F(x)} = \mathrm{P(X \leq x)}$ for all $x \in \mathbb{R}$. Then The CLT identifies real sequences a_n and b_n such that the normalized partial sums $\frac{\sum_{i=1}^{n} X_i - a_n}{b_n}$ converge to the standard Normal distribution when the sample size grows to infinity. Precisely, for CLT:

$$\lim_{x\to\infty} P\left[\frac{\sum_{i=1}^{n} X_i - a_n}{b_n} \leq x\right] = \Phi(x) \qquad (1)$$

with $a_n = \mathrm{nE}[\mathrm{X}_1]$ and $\mathrm{b_n} = \sqrt{\mathrm{nVar}(\mathrm{X}_1)}$. Therefore, one important requirement of CLT is the existence of finite moment of second order (i.e. $E(X_1)^2 < \infty$).

As to the EVT, the subject of interest is the maximum of an independent and identically sequence of the random variables:

$$M_n = max[X_1, X_2, \dots X_n]$$

Without any normalization, M_n will always converge to the upper limit of the support:

$$M_n \to \omega = \sup\{\, \mathrm{x} \in R \mid F(x) < 1\}$$

Because the maximum order statistics indicate that:

$$P[max(X_1, X_2, \dots, X_n) \leq x] = \big(F(x)\big)^n$$

Thus the key question for EVT is the existence of the real sequences of $a_n > 0$ and $b_n$such that the properly normalized sequence $\frac{M_n - a_n}{b_n}$ converge in distribution of $H(x)$, a non-degenerate function(i.e. not concentrated on a single point) :

$$\lim_{x\to\infty} P\left[\frac{M_n - a_n}{b_n} \leq x\right] = \lim_{x\to\infty} \big(F(a_n x + b_n)\big)^n = H(x) \qquad (2)$$

If there exists such $a_n > 0$ and b_n for all the points of continuity of H, then H is a *"Generalized Extreme Value (GEV) distribution"*. The initial distribution F satisfying the equation (2) is called the *"Maximum domain of attraction"* of H. The famous *"Extreme value theorem"* describes the finding of H as following:

> ***Theorem 1*** *(Fisher-Tippett (1928), Gendenko(1943)) Let (X_n)be a sequence of i.i.d. random variables. If there exists a sequence of real constants a_n and b_n such that (2) holds for some non-degenerate (i.e. not concentrated on a single point) distribution function H, then H must be a GEV distribution, i.e.*
>
> $$H = H_\xi \; for \; some \; \xi$$

The parameter ξ is known as the extreme value index (or the tail index). Then H belongs to one of the three standard extreme value distributions depending on ξ.

The one-parameter Generalized Extreme Value (GEV) (Jenkinson (1955)) distribution function with parameter ξ is of the form

$$\mathrm{H}_{\xi}(x) = \begin{cases} \exp\left(-(1+\xi x)_{+}^{-1/\xi}\right), & \xi \neq 0 \\ \exp\left(-\exp(-x)\right), & \xi = 0 \end{cases} \quad (3)$$

where $(1+\xi x)_{+}$ refers to the positive part of $(1+\xi x)$.

The three standard extreme value distributions of H_{ξ} are:

- The Fréchet distributions if $\xi > 0$, with the support of $H_{\xi} \in (-1/\xi, \infty)$
- The Gumbel distributions if $\xi = 0$, with the support of $H_{\xi} \in (-\infty, +\infty)$
- The Weibull distributions if $\xi < 0$, with the support of $H_{\xi} \in (-\infty, -1/\xi)$

The initial distribution F underlying the random variables samples $[X_1, X_2, \ldots X_n]$ is said to be in the domain of attraction of Fréchet, Gumbel and Weibull distribution for $\xi > 0$, $\xi = 0$ and $\xi < 0$ accordingly. An illustration of the three GEV distributions is shown in Figure 3-1.

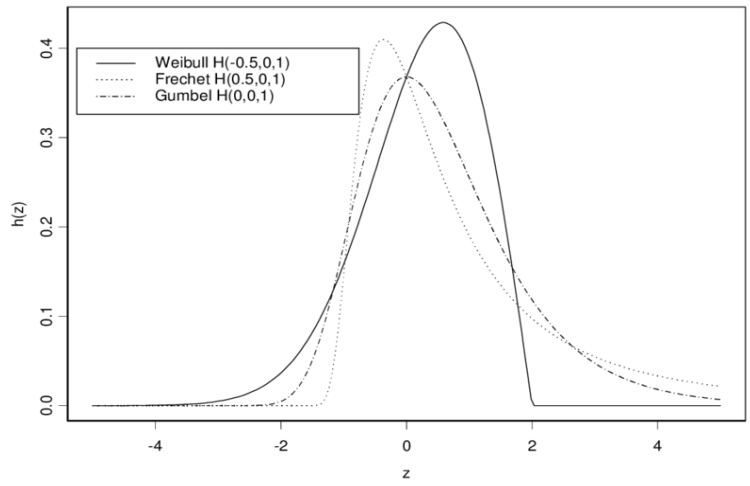

Figure 3-1. GEV distributions: Fréchet, Gumbel and Weibull distributions illustration

The Fréchet distributions enclose the heavy-tailed distributions while the Gumbel and the Weibull distributions refer to the distributions with lighter tail. In non-life insurance, the non-negligible probability of having an extreme large loss are often modeled by the heavy-tailed distribution as its upper bound is infinite. In the life insurance, the force of mortality in the end of lifespan is more likely to increase than to decrease. Thus, the distribution for modelling the mortality at the extreme age most likely belongs to the domain of attraction of the Weibull or Gumbel distributions. Moreover, the finite right endpoint of the Weibull distributions supports there is an age that can never be surpassed.

From this result, the Pickand's Theorem (1975) shows that the Generalized Pareto Distribution (GDP) gives a good approximation for $Y - u \mid \mathrm{Y} > \mathrm{u}$ when u is sufficiently large.

Theorem 2 *(Pickands-Balkema-de Haan Theorem) The Fisher-Tippett theorem holds with H_ξ if, and only if, there exists a certain function $\tau(u)$ such that*

$$log_{u\to\omega} \sup_{0\le y\le\omega-u} \mid F_u(y) - G_{\xi,\tau(u)}(y) \mid = 0 \qquad (4)$$

if and only if F belongs to the maximum domain of attraction of H_ξ.

The distribution function of the Generalized Pareto distribution $\mathcal{GP}ar(\xi,\beta)$ is given by

$$G_{\xi;\beta}(y) = \begin{cases} 1 - \left(1 + \xi\frac{y}{\beta}\right)_+^{-\frac{1}{\xi}} & if\ \xi \neq 0 \\ 1 - \exp\left(-\frac{y}{\beta}\right) & if\ \xi = 0 \end{cases} \qquad (5)$$

where $\beta > 0$ *and the support of* y *belongs to* $\begin{cases} (0,+\infty) & if\ \xi \geq 0 \\ \left(0,-\frac{\beta}{\xi}\right) & if\ \xi < 0 \end{cases}$

ξ is the shape parameter (or the tail index), and β is the scale parameter. The corresponding probability density function is

$$g_{\xi;\beta}(y) = \begin{cases} \frac{1}{\beta}\left(1 + \xi\frac{y}{\beta}\right)_{+}^{-\frac{1}{\xi}-1} & if\ \xi \neq 0 \\ \frac{1}{\beta}\exp\left(-\frac{y}{\beta}\right) & if\ \xi = 0 \end{cases} \quad (6)$$

Actually, the simple pareto distribution $\mathcal{P}ar(A,\alpha):\ F(y) = \left[\frac{y}{A}\right]^{-\alpha}\ (\alpha > 0)$is a special case of GPD when $\xi\ > 0$ and $\begin{cases} \xi = \frac{1}{\alpha} \\ \beta = A/\alpha \\ \mu = A \end{cases}$. Also, a type II pareto distribution is found in case of $\xi < 0$; When $\xi = 0$, the GPD becomes an exponential distribution.

Provided $\xi < 1$, the mean of $\mathcal{GP}ar(\xi,\beta)$ is $\frac{\beta}{1-\xi}$ and the variance equals to $\frac{\beta^2}{(1-\xi)^2(1-2\xi)}$ for $\xi < 1/2$.

3.1.1 Useful properties

Threshold stability: If $X \sim \mathcal{GP}ar(\xi,\beta)$, then the excess of X over a higher threshold u is also a generalized Pareto distribution with the same shape parameter ξ and a linearly transformed scaling parameter $\beta(u) = \beta + \xi u$.

Proof:

$$\begin{aligned} \bar{F}_u(x) &= \frac{\overline{F(u+x)}}{\overline{F(u)}} \\ &= \frac{\left(1+\xi\frac{u+x}{\beta}\right)^{(-1/\xi)}}{\left(1+\xi\frac{u}{\beta}\right)^{(-1/\xi)}} \\ &= \left(\frac{\beta+\xi(u+x)}{\beta+\xi u}\right)^{(-1/\xi)} \\ &= \left(1+\xi\frac{x}{\beta+\xi u}\right)^{(-1/\xi)} \end{aligned}$$

(end)

Linearity of the GPD mean excess function beyond threshold u *:* Given a random sample $X_1, X_2, \dots, X_n$ of size n with initial common distribution F, and a high threshold u, $X_1 - u, X_2 - u, \dots X_n - u$ are considered as the exceedances over u provided $X_i > u$. Then the mean excess function of a GPD is of a linear form.

Proof:

The common distribution function of the excesses of the variable over a high threshold u, given that the threshold u is reached, is

$$\begin{aligned} F_u(x) &= P[\, X - u \le x \mid \mathrm{X} > \mathrm{u}\,] \\ &= \frac{F(x+u) - F(u)}{1 - F(u)} \\ &= 1 - \frac{\overline{F(u+x)}}{\overline{F(u)}}, \qquad x \ge 0 \end{aligned}$$

Provided that $E[X] < +\infty$, the mean excess function of X is defined as

$$e_X(u) = E[\, X - u \mid \mathrm{X} > \mathrm{u}\,] = \int_0^{+\infty} \overline{F}_u(x)dx$$

Then the mean excess function of a $\mathcal{GP}ar(\xi, \beta)$ is given by

$$e(u) = \frac{\beta(u)}{1-\xi} = \frac{\beta}{1-\xi} + \frac{\xi}{1-\xi}u \qquad (7)$$

where $u \ge 0$ if $0 \le \xi < 1$ and $0 \le u \le -\frac{\beta}{\xi}$ if $\xi < 0$. (end)

3.2 Extreme value theory in the context of analyzing the mortality at extreme ages

The extreme value theory is a natural candidate to model the mortality of the extreme age at death as

- The data observed in the very end of the lifespan is limited.
- The common distributions used to fit the mortality for whole lifespan fit badly the tails.

Moreover, is there an age limit that cannot be surpass for human-being? The question can be treated by the EVT if the lifetime of a human is considered as a random variable with a certain distribution, then the question becomes does there exist a finite endpoint of this distribution.

The subject of analysis in our study is the mortality at the extreme ages. Let T_i represent the lifetime of an individual i. The probability that the individual dies at age x can be modeled by a distribution function:

$$\mathrm{F}(x) = {}_xq_0 = P[T_i \leq x], \ \ x \geq 0, \tag{8}$$

The remaining lifetime for the individual attained age u can be written as

$$F_u(x) = {}_xq_{\mathrm{u}} = P[T_i - \mathrm{u} \leq \mathrm{x} \mid \mathrm{T_i} > u] \tag{9}$$

The ultimate age that the individual can attend theoretically is denote as ω:

$$\omega = sup\{\, x \geq 0 \ \mid \ {}_xq_0 < 1 \,\} \tag{10}$$

We would like to know if the ultimate age ω is finite or not. Suppose there are n individuals with a common lifetime distribution function F. Let M_n be the observed oldest age of these individuals. It is obvious that when the sample size n increases, the observed maximum age M_n will approach to the real ultimate age ω:

$$P[M_n \leq x] = \left({}_xq_0\right)^n \rightarrow \begin{cases} 0 & if\ x < \omega \\ 1 & if\ x \geq \omega \end{cases} \quad when\ n \rightarrow \infty \tag{11}$$

But recall Theorem 1 that if there exists a real sequence of $a_n > 0$ and b_n such that the properly normalized sequence $\frac{M_n - a_n}{b_n}$ will converge to one of the three GEV distributions $\left(H_\xi\right)$. This corresponds to the following setting:

$$\lim_{x \rightarrow \omega} P\left[\frac{T-u}{\beta(.)} > x \,\middle|\, \mathrm{T} > \mathrm{u}\right] = 1 - G(x) \tag{12}$$

In other words, if there exists some appropriate function $\beta(.)$, the remaining lifetime at a sufficiently high threshold age u can be approximated by a GPD according to the approach POT: ${}_xq_{\mathrm{u}} \approx G_{\xi;\beta(.)}(x)$

The mortality at extreme age can be modeled in the following way: given the threshold age u^*, ${}_xq_{u^*}$ represents the probability that an individual dies before reaching age $x + u^*$. Then the exceedances $x - u^*$ can be approximated by a GPD:

$$q_{x-u^*} = \begin{cases} 1 - \left(1 + \xi \dfrac{(x-u^*)}{\beta}\right)^{-\frac{1}{\xi}} & if\ \xi \neq 0 \\ 1 - \exp\left(-\dfrac{(x-u^*)}{\beta}\right) & if\ \xi = 0 \end{cases} \quad (13)$$

Recall the instantaneous risk to die (or "hazard rate") for an individual at age x is quantified by the force of the mortality μ_x:

$$\mu_x = \lim_{\Delta x \to 0} \frac{P[\, x < T < x + \Delta x \mid \mathrm{T} > \mathrm{x}\,]}{\Delta x} = \frac{\dfrac{d}{dx}\, {}_xq_0}{{}_xp_0} \quad (14)$$

Then the expression $1/\mu_x$ can be considered as a parameter to quantify the force of resistance to die at age x. When x approaches the ultimate age, the relationship between the tail index ξ and the force of resistance to mortality at age x can be expressed as

$$\lim_{x \to \omega} \frac{d}{dx}\left(\frac{1}{\mu_x}\right) = \xi \quad (15)$$

If the force of resistance to die is decreasing with the attained age at the end of life span, then a negative ξ is expected and the initial distribution F underlying the observed lifetimes will belong to the Weibull family, which indicates that there will be an ultimate age ω which can never be surpassed as the support of the Weibull distributions has a right finite endpoint ($u - \frac{\beta}{\xi}$). This gives the estimated ultimate age

$$\widehat{\omega} = u^* - \frac{\hat{\beta}}{\hat{\xi}} \quad (16)$$

In our example, all the individuals died beyond age 105, and we are going to select a threshold u^* such that the remaining lifetime $(x - u^*)$ can be approximated by a GPD.

4 Case study results

4.1 Homogeneity

Before applying the extreme value theory to the mortality at extreme ages, a delicate question is that whether the data of different countries can be grouped together as for certain groups the numbers of observations are extremely limited due to the fact that the death age in our analysis is in the very end of human lifespan. Grouping can create a larger size sample which will increase the statistical power compared to a limited size sample. However, grouping can lead to the problem of homogeneity:

- the homogeneity among the 18 birth cohorts from 1881 to 1898
- the homogeneity among the different countries and
- the homogeneity between female and male.

4.1.1 Homogeneity among the birth cohorts

Regarding the homogeneity among the 18 birth cohorts, no visible upward trend of the average age beyond 105 at death for the birth cohorts from 1881 to 1898 has been observed in all countries and it is stagnated at around 106 to 107 for both men and women (see figure 3 of section 2). We may suggest that the underlying distribution of the individual mortality at extreme age is invariant over the 18 cohorts from 1881 to 1898. Meanwhile, an upward trend of the maximum age can be observed in almost all countries except in Switzerland. This can be explained by the increase in the number of births in the birth cohorts from 1881 to 1898 that the maximum observed age at death approaches to the ultimate death age.

We have not adopted the test statistic of Cramer-Von Mises to test the homogeneity between different birth cohorts because the test result is very sensible to the subjective criteria under which the groups are formed. Indeed, the grouping by 3 birth cohorts or 6 birth cohorts leaded to confusing results. Instead, the Mann-Kendall test has been used to access the presence of any monotone trend about the average age of death beyond 105 of cohorts from 1881 to 1898 for all the countries. The null hypothesis is that the data come from a population with independent realization and are identically distributed; the alternative hypothesis is that the data follows a monotonic trend. No significant mortality trends throughout cohorts is

detected in our datasets (Table 4-1), which supports that the 18 birth cohorts can be grouped together for both female and male.

Mann-Kendall Trend test p-value		
	Female	Male
Belgium	0,3233	0,7868
France	0,1386	0,1846
German	0,0578	0,6207
Switzerland	0,0696*	0,6522

*Table 4-1. Mann-Kendall Trend test about the average age of death beyond 105 for birth cohorts 1881-1898, by sex and country. *The result of Switzerland female group is obtained after exclusion of the data of birth cohort 1881 which was exceptionally high.*

4.1.2 Homogeneity among the different countries by gender

As to the homogeneity among the different countries, since Belgium, France, German and Switzerland are geographically close to each other and the lifestyle in these countries is similar, we assume the individual mortality at extreme age for women and for men from these countries share a common distribution. A K-sample Anderson-Darling test has been performed to verify the assumption. The null hypothesis is

$$H_0: F_n(x)^{BE} = F_n(x)^{DEU} = F_n(x)^{FRA} = F_n(x)^{CHE}$$

Against the alternative hypothesis H_1: All the underlying distributions are not the same.

The p-value for female and for male is 0,02739 and 0,4664 respectively. The p-value suggests the rejection of the null hypothesis for the female datasets while non-rejection of the null hypothesis for the male datasets. So, the male datasets can be grouped together. A closer look at the female datasets is necessary.

By plotting the empirical cumulative distributions of the age at death (Figure 4-1), the female datasets from Belgium and from Switzerland are close to each other. The difference comes from the German female dataset and the French female dataset. Apparently, the empirical CDF of the French female is overall lower than that of the German female, which suggests that a larger proportion of French females live longer than the German female. What's more, considering the fact that there are 2823 observations of French female compared to 835 observations of German female, it would be better to introduce "nationality" as a covariate such that the true development of German female lifespan won't be masked by the French female datasets.

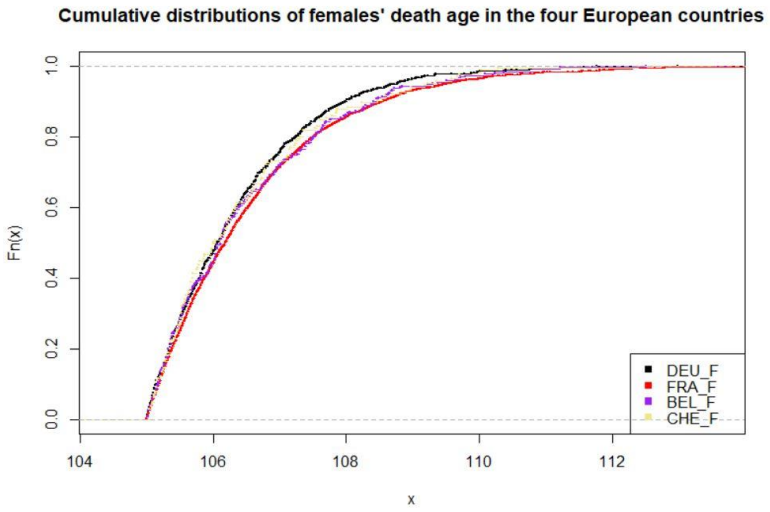

Figure 4-1. Cumulative distribution of age at death for the females in Belgium, France, Germany and Switzerland, cohorts 1881-1897

4.1.3 Homogeneity between female and male

In terms the homogeneity between female and male, the huge difference in the numbers of observations between men and women within each country suggests there must be a different level of mortality for men and for women. So even though the datasets of male observations are much sparser compared to the female observations, they will not be grouped together because they cannot be considered as a drawn from a common distribution. As a result, the analysis will be performed following 2 steps:

- Step 1: a first analysis on the group level of female and male datasets. The female group is composed of the females from Belgium, France, Germany and Switzerland for all the birth cohorts 1881 to 1898, and the male group is composed of the males from Belgium, France, Germany and Switzerland for all the birth cohorts 1881 to 1898 and
- Step 2: a second analysis will be performed on the subsets of the female group via the covariate “nationality”.

Figure 4-2 displays the boxplots and Figure 4-3 shows the histograms for the female datasets and the male datasets after grouping. The female group has a longer tail in the attained highest age at death than the male group.

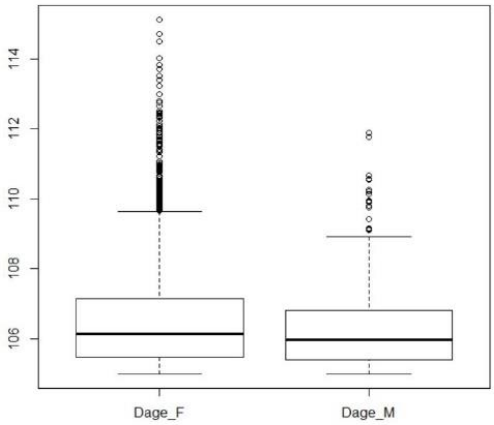

Figure 4-2 Boxplots for the female and male datasets composed of Belgium, France, Germany and Switzerland, cohorts 1881-1898

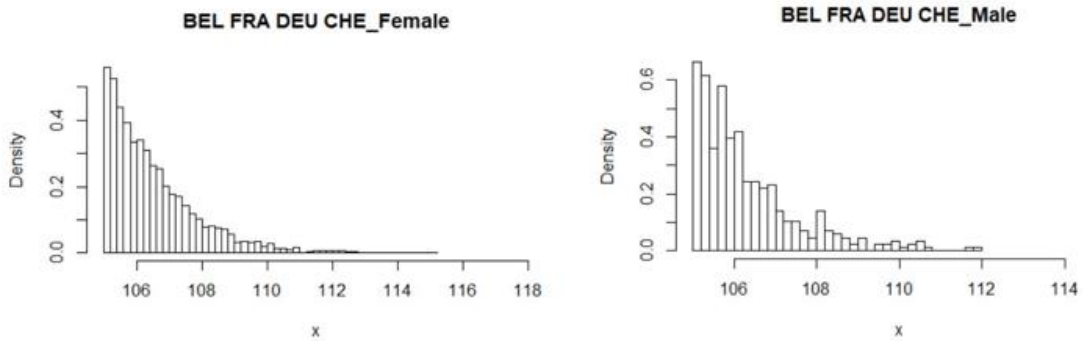

Figure 4-3 Histograms for the female and male datasets composed of Belgium, France, Germany and Switzerland. (Female on the left panel and male on the right panel)

4.2 Threshold selection

From which age of death, the distribution of the lifetime follows a generalized pareto distribution? It is critical to select an appropriate threshold in the POT approach. The threshold level will affect the accuracy of the parameters estimation when fitting a generalized Pareto distribution and the results of extrapolation. The principal is that the threshold should be high enough such that the exceedances over the threshold fit well a generalized Pareto distribution. But the threshold cannot be too high that the too small sample size can increase substantially the variance of the estimators. There are many threshold selection methods in the literatures. See (Scarrott and MacDonald 2012). As the subject of our study concerns the mortality at the extreme ages (above 105), the method of selection of the threshold will be limited to

- The graphical tools: the mean residual life plot and the shape plot;

- Raw p-value of the goodness-of-fit test and
- An Automated selection based on the test of an ordered series of thresholds via goodness-of-fit test combined with a forward stopping rule.

4.2.1 Graphic tools

4.2.1.1 The empirical mean excess function plot

The empirical mean excess function plot is based on the property of the "linearity of the GPD mean excess function beyond threshold u". The idea is to detect a sufficiently high threshold u beyond which the empirical mean excess function displays an approximately linear form. The empirical mean excess function $e_X(u)$ can be estimated from a random sample of claims $\{X_1, X_2, \dots, X_n\}$ by

$$\begin{aligned} \widehat{e_X}^n(u) &= \frac{\sum_{i=1}^{n} x_i \, 1[X_i > u]}{\#\{X_i : X_i > u\}} - u \\ &= \frac{\sum_{i=1}^{n} (x_i - u) \, 1[X_i > u]}{\#\{X_i : X_i > u\}} \end{aligned} \tag{17}$$

Denoting $X_{[1]}, X_{[2]}, \dots, X_{[n]}$ the observations arranged in ascending order, i.e. $X_{[1]} \leq X_{[2]} \leq \cdots \leq X_{[n]}$,

$$\widehat{e_X}^n\left(X_{[k]}\right) = \frac{1}{n-k} \sum_{j=1}^{n-k} \left(x_{[k+j]} - x_{[k]}\right) \tag{18}$$

It is also proven that the plot of the empirical mean excess function will show

- an upward trend if the initial distribution F is a heavy-tailed distribution
- a downward trend if the initial distribution F is a short-tailed distribution
- an approximately horizontal line for exponentially distributed data

The empirical plot of the mean excess function (Figure 4-4) shows the relationship between the attained death age x and the expectancy of the remaining lifetime $\widehat{e_X}$. A visible downward trend can be seen on both the female and male group. According to the EVT, the initial distribution F underlying the observations should be in the domain of attraction of Weibull family, and it is a short-tailed distribution. This observation also supports that there is an ultimate age that can never be surpassed.

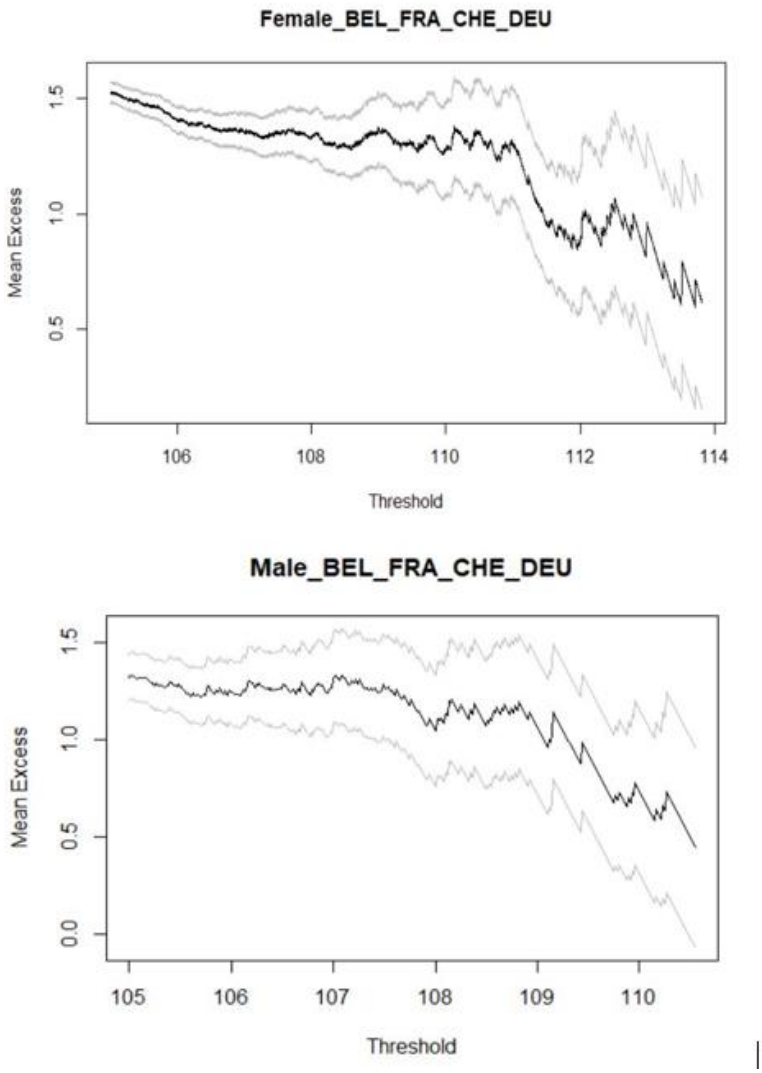

Figure 4-4 The empirical mean excess function plost for the female and the male datasets composed of Belgium, France, Germany and Switzerland, age of death above 105, cohort 1881 – 1898 (Female on the upper panel and male on the lower panel)

What is interesting is that at extreme ages, the resistance to the mortality does not decrease at the same speed as downward trend does not follow strictly a linear decrease trend from the beginning (i.e.: 105 years old) to the very end. Indeed, for the female group, at age 105, the expected remaining life is about 1,5 years and it decreases slowly until age 110, beyond that age, the expected remaining life decreases much faster, which also means the resistance to mortality falls much more quickly beyond 110 than before 110 years old. For the male group, the same pattern can be seen that the expected remaining life starts at 1,3 years for 105 years old and decreases slowly until 109 years old and the downward trend becomes more evident beyond 109. Recall that the resistance of mortality is quantified by the shape

parameter of the GPD distribution, we suspect the shape parameter would be different before and after the critical age.

The short-tailed behaviour of the lifetime distribution is also confirmed by the empirical plot quantile plot of the data against a negative Exponential distribution, the QQ-plot. For a short-tailed distribution, its empirical quantiles grow more slowly than that of a negative Exponential distribution, so it shows a concave pattern on the QQ-plot (Figure 4-5).

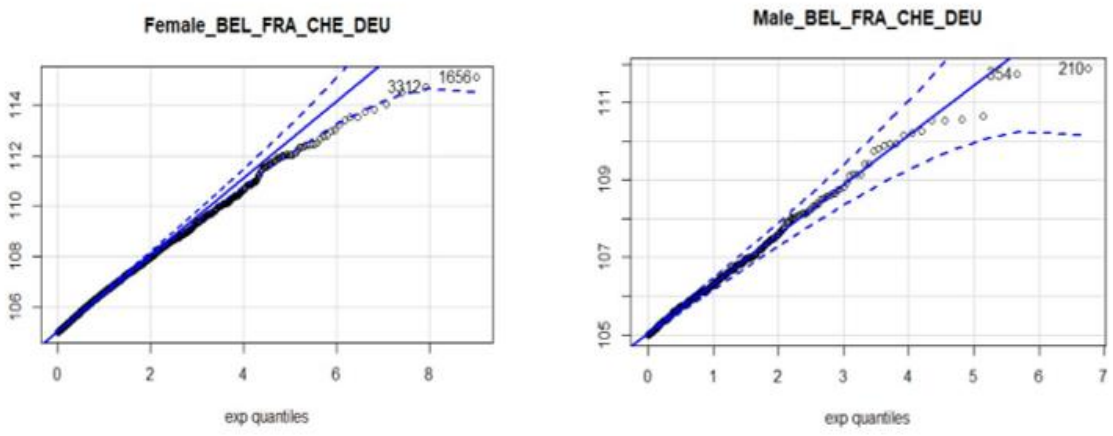

Figure 4-5 The empirical QQ-plot for the female and the male datasets composed of Belgium, France, German and Switzerland. age of death above 105, cohort 1881 – 1898 (Female on the left panel and male on the right panel)

4.2.1.2 The shape plots

Another graphical view "the shape plot" of how the estimated shape parameter interacts with the threshold levels and the increasing numbers of exceedances. It is constructed in the following way: by controlling the number of exceedances (or the thresholds level) and the number of consecutive GPD models to be fitted to the exceedances, the estimates of the corresponding shape parameters is obtained via the maximum likelihood method and plotted against the numbers of the exceedances in an ascending order. The corresponding level of thresholds are also marked on the graphics.

The shape plots confirm that both the female group and the male group have a negative shape parameter, especially in the very end of the life span. For the female group, the shape parameter stabilizes at a level close to 0 for the death age from 105 to 109, while for the male group the shape parameter becomes stable and close to 0 between 106 and 107.(Figure 4-6)

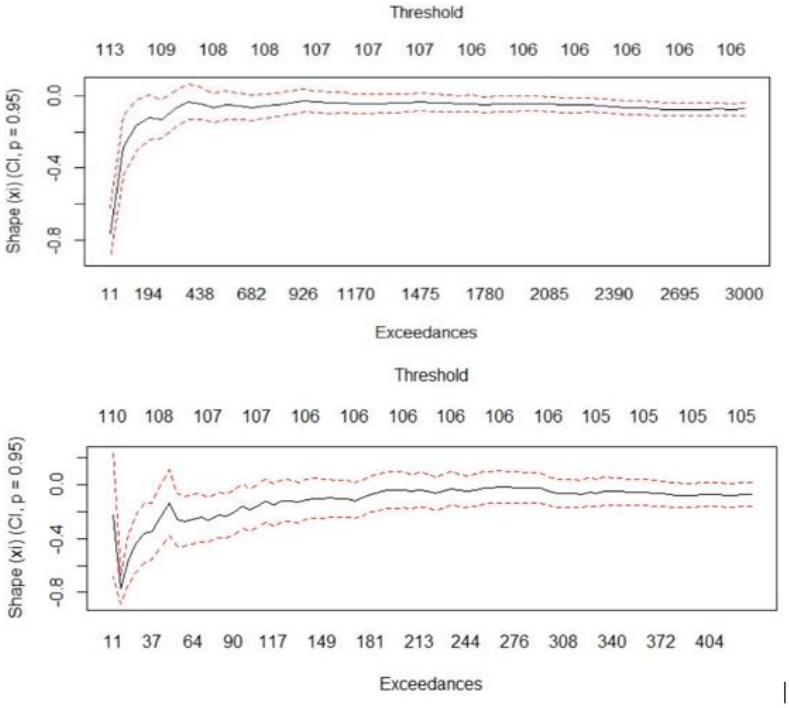

Figure 4-6 The shape plots for the female and the male datasets composed of Belgium, France, German and Switzerland, age of death above 105, birth cohorts 1881 – 1898 (Female on the upper panel and male on the lower panel)

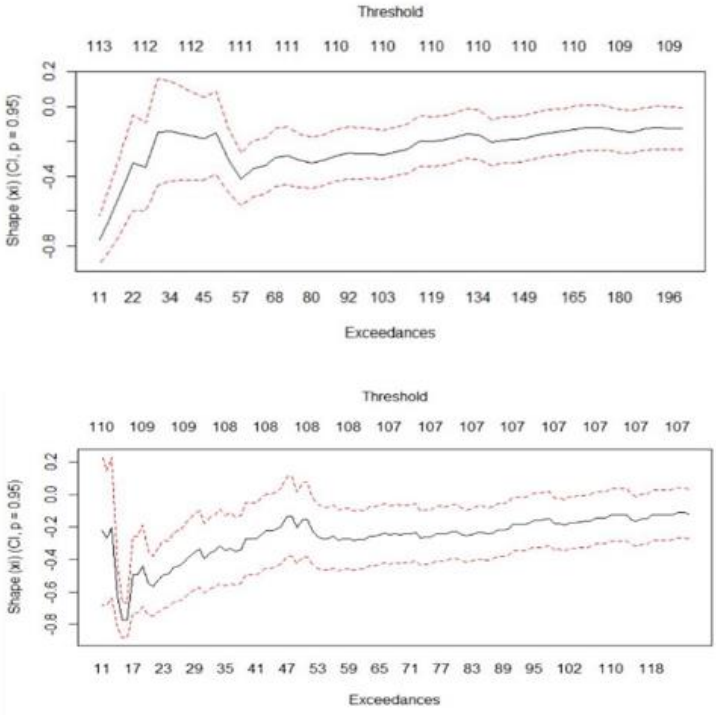

Figure 4-7 The shape plots for the female and the male datasets composed of Belgium, France, German and Switzerland, age of death above 109 for female and 107 for male, birth cohorts 1881 – 1898 (Female on the upper panel and male on the lower panel)

It might be tempting to conclude that the high volatility of the shape parameters beyond these critical ages (i.e.: 109 for women and 106-107 for men) is due to the small sample size. But the corresponding numbers of exceedances beyond these critical ages are still around 200 and 150 for women and men, respectively. To investigate more the shape parameters beyond the critical ages, a second shape plot is done with only 200 exceedances for the female group and 150 for the male group (see Figure 4-7). It seems that the shape parameters become really volatile when the numbers of exceedances are below 50. Beyond 50 exceedances, the shape parameters become stable again but at a lower level (i.e.: around -0,3 for female group and -0,2 for the male group) than the level of the shape parameters before the critical ages. These findings are in accordance with the observations from the empirical plots of excess mean function.

In conclusion, we believe for the birth cohorts 1881-1898, the force of resistance to mortality at the extremes ages beyond 105 is not homogeneous. Indeed, for the female group, the force of resistance to mortality stays stable until around 109 while for the male group, it stays stable until around 107. After these critical ages, the chance to die (hazard rate of death) is becoming much higher. As a result, the shape parameters quantifying the force of resistance to mortality could be materially different for the death ages before and after the critical ages. And we also believe the true information about the highest ages at death is carried more by the data beyond the critical ages than data before the critical ages. Using a same shape parameter to fit all the death ages from 105 to the end of the life span in our case may lead to a substantial bias about the estimation of the highest age. A formal discuss will be addressed in the section5.

4.2.2 Threshold selection via the goodness-of-fit test

A second way to select the threshold is based on the $p - value$ of the goodness-of-fit testing on a selection of ordered selected thresholds. The idea is based on the goodness-of-fit tests for the generalized Pareto Distribution proposed by Choulakian and Stepens (2001). In the article, it is shown how the tests can be used to help select the threshold in the POT model for the heavy tailed distributions. The technique of choosing the threshold is based on the stability property of the GPD.

For the female group, the initial threshold value is 105, which is the lowest death age in our dataset. There were 4092 exceedances. The value of the test statistics is Anderson-Darling (AD) statistic $A^2 = 1{,}08093$ and a Cramér-von Mises (CVM) statistic $W^2 = 0{,}15808$, the corresponding p-value is 0,04033 and 0,0484 (p value< 0.1) respectively, so the GPD does not fit well the exceedances at the threshold value of 105. Next, we raise the threshold by 0.01 until the GPD fits the new values of the remaining exceedances. The choice is the first threshold whose p value exceeds 10%. The results are resumed in Table 4-2.

	CVM		**AD**	
	Statistic	p-value	Statistic	p-value
Female	105,79	0,1664	105,89	0,2525
Male	105,00	0,6016	105,00	0,5750

Table 4-2 Thresholds selection via the goodness-of-fit tests

We notice the Anderson-Darling goodness-of-fit test always give a higher threshold compared to the CVM goodness-of-fit test, which is in accordance with the findings in the original article. But the caveat is that the raw p-values are not stable. Indeed, the plots of the p-values of the CVM test and the AD tests against the increasing level of the thresholds shows a high volatility about the p-values (see Figure 4-8).

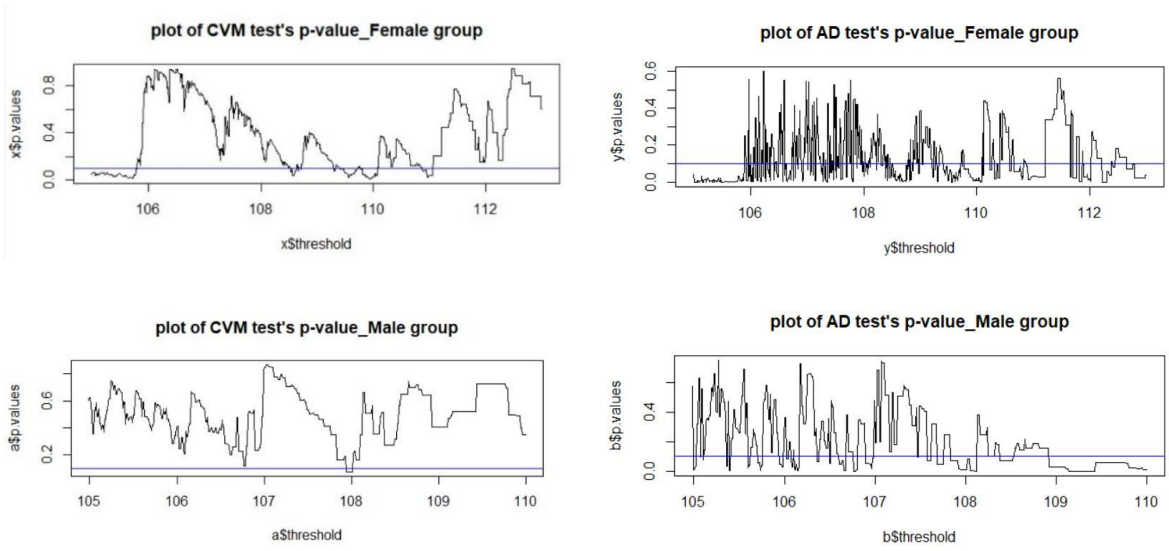

Figure 4-8 Plots of the p-values of the CVM and AD tests, the blue lines corresponds to 10%.

In fact, the stability property is not respected by choosing the first threshold whose p value exceeds 10%. Choulakian and Stephens in the original paper describes the problem as a "failure-to-reject" problem that the statistic test fails to reject the null hypothesis when the true distribution is not yet a "GPD" because of an acceptance may happen at a low threshold

by chance.(Choulakian and Stephens 2001) Besides, it seems the p-values of AD tests is difficult to be interpreted as it remains volatile until the very end of the data.

4.2.3 Threshold selection via an automated selection procedure

Bader, Yan, and Zhang (2018) has proposed an automated threshold selection procedure based on a sequence of goodness-of-fit tests with error control for ordered, multiple testing. This method is more robust and efficient compared to the raw goodness-of-fit test. It is proceeded in the following way: First, the selected sequence of thresholds will be tested in an ascending order. For each of the threshold, a goodness-of-fit test (i.e.: Anderson-Darling test, the CVM test, etc.) will be performed. This gives a sequential testing of ordered null hypothesis $H_1, \dots, H_l$ with the corresponding p-values $p_1, \dots, p_l$. Secondly, the raw p-value ordered null hypothesis $H_1, \dots, H_l$ will be transformed to a monotone sequence (G'Sell et al. 2015) such that the original method of Benjamini and Hochberg (Benjamini and Hochberg 1995) can be applied. The rejection rule will return a cutoff $\hat{k}$ such that all the null hypothesis before $H_{\hat{k}}$ are rejected. This method is called the "Forward Stopping rule". The cutoff point $\hat{k}_F$ is giving by

$$\hat{k}_F = \max\{\, \mathrm{k} \in [1, \dots, \mathrm{l}] : -\frac{1}{k}\sum_{i=1}^{k} log(1 - p_i) \leq \alpha \,\}$$

where α is a desired level for the false discovery rate (FDR).

It is shown that the false discovery rate (FDR) can be better controlled than based on a simple raw p-value in an ordered sequential testing. Thus, the combination of the raw p-value and the "forward stopping rule" will provide a robust and efficient automated selection procedure. With the FDR level α set to 10%, the choice of the thresholds is shown in Table 4-3.

	CVM + Forward Stopping rule		**AD + Forward Stopping rule**	
	Statistic	p-value	Statistic	p-value
Female	105,91	0,7447	106,86	0,3856
Male	<105	>0,1	<105	>0,1

Table 4-3 Thresholds selected by the automated selection procedure

4.2.4 Final choice for the thresholds

For the reason of robustness and efficacity, we favored the results given by the automatic selection procedure which combine the Anderson-Darling test with the forward stopping rule. Our final choice for the thresholds for female and male group is shown in Table 4-4

Selected Threshold u^* for datasets $\geq$ 105 age at death	
Female Group	106,86
Male Group	105,00

Table 4-4 Final choice for the thresholds

For the female group, the automatic selection procedure with the Anderson-Darling goodness-of-fit test gives the highest threshold. By choosing the highest results, the stability of threshold property is ensured.

The true threshold for the male group is below 105 since the automated selection procedure returns "0" as the necessary step to take before reaching the forward stop criteria. This corresponds to the results obtained by the threshold selection via the goodness-of-fit test based on the raw $p - value$ method that at age 105, the raw $p - value$ is already far beyond the 0.1 desired level.

4.3 Estimation of the generalized Pareto distribution parameters

Now the thresholds for the female group and male group are chosen, the next step is to fit a generalized Pareto distribution to the exceedances of the threshold. Many methods of estimation are available for the estimation of the parameters, among which, the maximum likelihood method is the most used one due to its optimal statistical properties. The notation used here is from Gbari et al. (2017).

Notation:

- $t_i \in T_i$: the observed lifetime of individual i
- the number of survivors at age u^*: $L_{u^*} = t_i \mid t_i > u^*$
- $t_i - u^*$: the remaining lifetime over the threshold age u^*

The contribution to the likelihood of an individual who died beyond the threshold age of death u^* is its density distribution function

$$f(x) = \frac{1}{\beta}\left(1 + \xi\frac{(t_i - u^*)}{\beta}\right)^{-\frac{1}{\xi}-1} \qquad if\ \xi < 0 \tag{19}$$

The GP likelihood function for observed remaining lifetimes at beyond the death age u^* is given by

$$\mathcal{L}(\xi, \beta) = \prod_{i|t_i > u^*}\left(\frac{1}{\beta}\left(1 + \xi\frac{(t_i - u^*)}{\beta}\right)^{-\frac{1}{\xi}-1}\right)$$

The corresponding log-likelihood function to be maximized is given by

$$LL(\xi, \beta) = -ln\beta\ num.of\ \{t_i \mid \mathrm{t_i} > \mathrm{u^*}\} - \left(1 + \frac{1}{\beta}\right)\sum_{i|\mathrm{t_i}>\mathrm{u^*}} ln\left(1 + \xi\frac{(t_i - u^*)}{\beta}\right) \tag{20}$$

The starting value for the parameter ξ and β is given by their theoretical expressions using their moment conditions based on the sample mean and the sample variance:

$$\widehat{\xi^{MOM}} = \frac{1}{2}\left(1 - \frac{\bar{x}^2}{s^2}\right) \quad \text{and} \quad \widehat{\beta^{MOM}} = \frac{1}{2}\bar{x}\left(\frac{\bar{x}^2}{s^2} + 1\right)$$

where $\bar{x}$ and s^2 are the sample mean and the variance of the observed exceedances.

At the threshold $u_F^* = 106{,}86$ and $u_M^* = 105$, there are 1238 and 430 exceedances for the female and the male group, respectively. The results are summarized in Table 4-5:

	Female	**Male**
$\boldsymbol{u^x}$	106,86	105
$\boldsymbol{L_{u^*}}$	1238	430
$\hat{\boldsymbol{\xi}}$	-0,0445	-0,0706
$\boldsymbol{s.e.}\left(\hat{\boldsymbol{\xi}}\right)$	0,02853	0,0480
$\widehat{\boldsymbol{\beta}}$	1,4339	1,4116
$\boldsymbol{s.e.}\left(\widehat{\boldsymbol{\beta}}\right)$	0,05771	0,0960
AD p-value	0,3856	0,3155
CVM p-value	0,7447	0,6245

Table 4-5 GPD parameters estimation for the female and the male datasets

To check model fit, the diagnostic plots (Figure 4-9 and Figure 4-10)are

- The empirical and theoretical density plot
- The Q-Q plot
- The empirical and theoretical CDF plot
- The P-P plot

For both the female and the male group, the GP models fit well the exceedances.

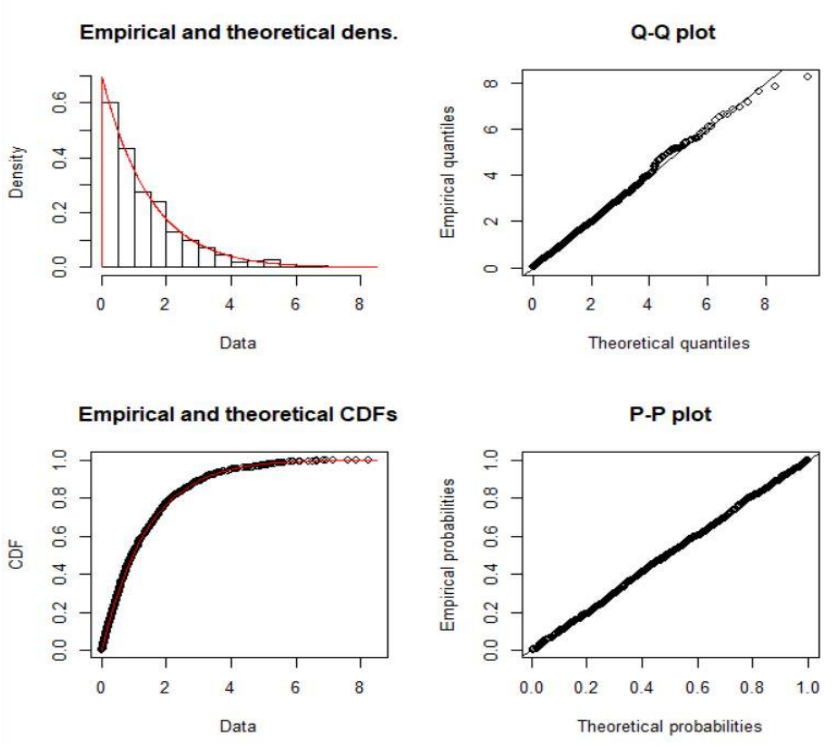

Figure 4-9 Model fitting diagnostic for the female datasets

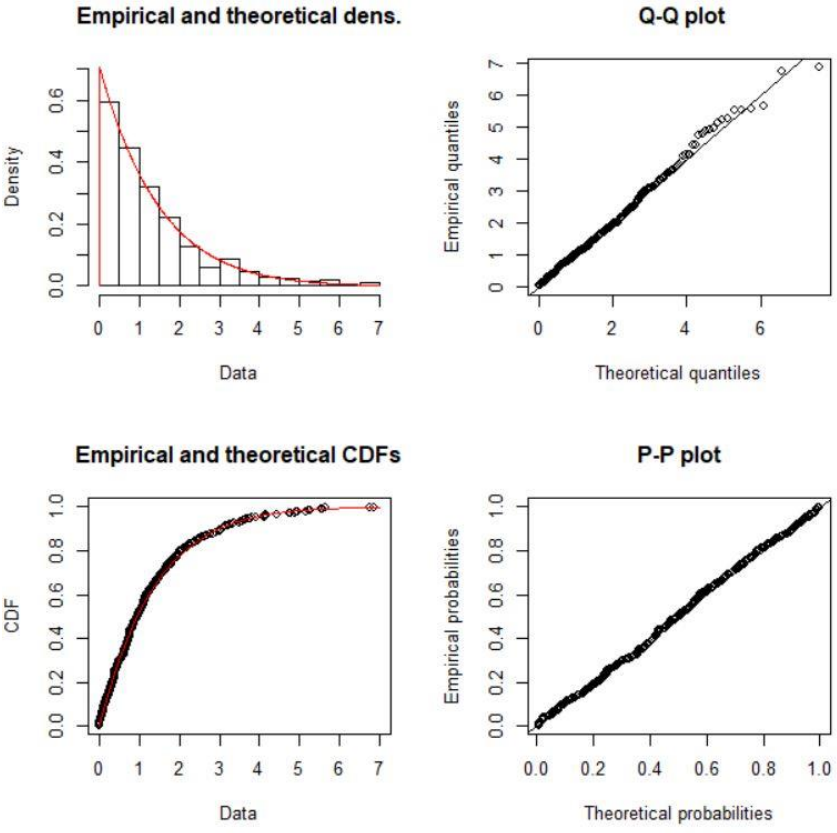

Figure 4-10 Model fitting diagnostic for the male datasets

4.4 Analysis on the female sub-datasets

In section 4.1, the female group has a problem of homogeneity concerning the datasets of France and the datasets of Germany. Besides, the large proportion of the French observations in the group sample may dominate the analysis results of other countries. To solve these problems, in this section we will perform the similar analysis on the subgroup of the female datasets. The analysis on the subsets of the male group is not necessary for two reasons: 1) The K-sample test shows that not enough evidence to reject the null hypothesis which means the datasets of the four countries can be considered having a common distribution; 2) The size of the subsets are so limited which will infect the power of the statistical tests.

Firstly, the empirical mean excess function (Figure 4-11) shows that for all the countries, there is a downward trend, which support a negative shape parameter and a finite endpoint of the distribution. Again, the French female datasets display a certain change in terms of the force of resistance of mortality. Indeed, for the French female dataset, the change point is around 110 years old. Indeed, the remaining lifetime expectancy decreases slowly until that age and drops dramatically afterwards. This phenomenon is not visible in the female datasets from Belgium and Switzerland. For the German female datasets, the decreasing trend can still be regarded as a linear trend except the bump observed around 109 to 109,5.

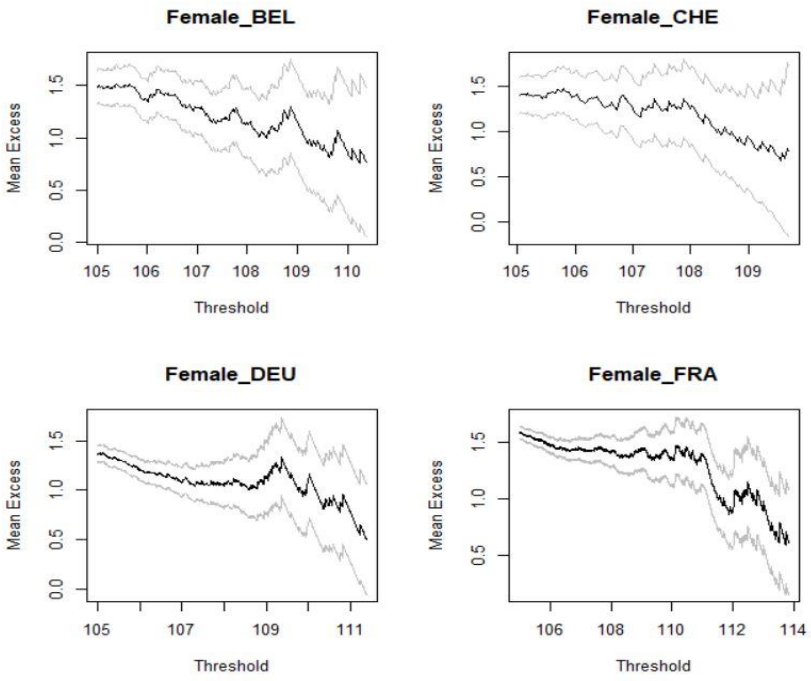

Figure 4-11 The empirical mean excess function plots for the female sub-datasets

Secondly, the shape parameter plots (Figure 4-12) also show a high volatility when the sample size becomes limited. Indeed, the shape parameter will become relatively stable when there are more than 50 exceedances. The shape plots also confirm negative shape parameters are expected.

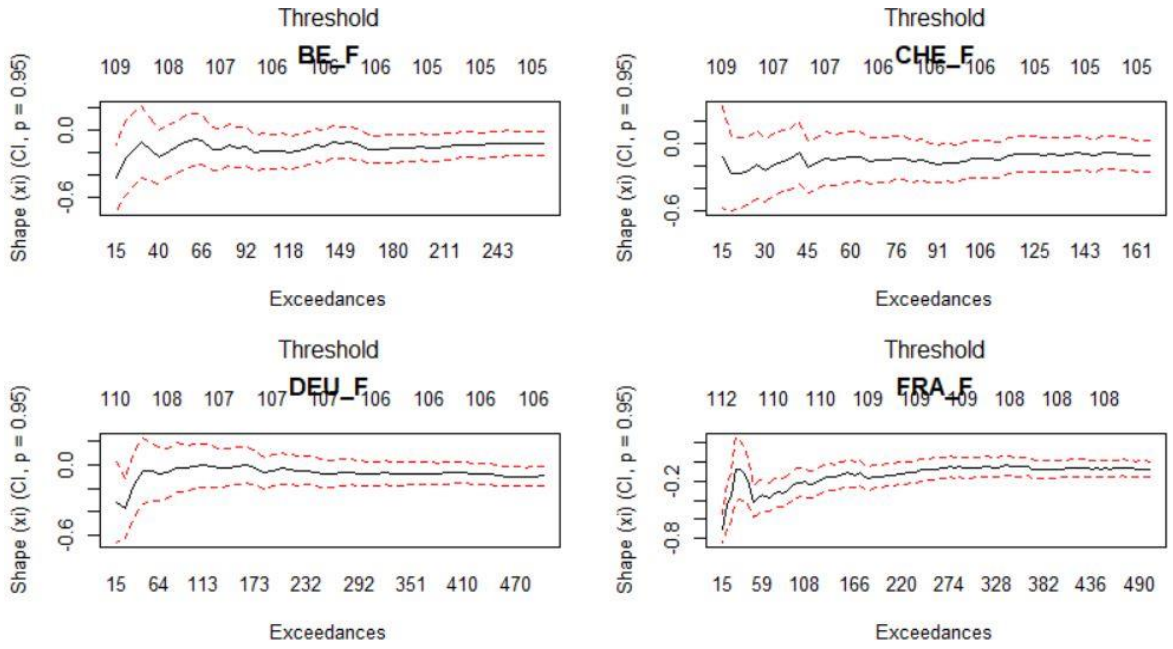

Figure 4-12 The shape plots for the female sub-datasets

Thirdly, via the automated selection procedure based on the AD goodness-of-fit tests combined the forward stop rule, the thresholds for the four countries female datasets are shown in Figure 4-13:

	AD + Forward Stopping Rule (Female sub-datasets)	
	Threshold	p-value
Belgium	<105	0,4096
France	106,21	0,1908
Germany	105,15	0,1027
Switzerland	105,04	0,7774

Figure 4-13 Thresholds selected by the automated selection procedure for the female sub-datasets

Compared with results obtained on the group level, the thresholds of the subsets are all lower than that of the group level. Among them, the French datasets have the highest threshold. And the true threshold for the Belgium datasets is lower than 105. It seems that by introducing a "nationality" covariate, lower threshold is needed to observe the GPD behavior. The same conclusion has been seen in Cebrián, Denuit, and Lambert (2003).

Lastly, the corresponding estimated parameters for each subset to fit a GPD distribution is summarized in Table 4-6:

	GPD parameters fitting by Maximum Likelihood Method				
	Belgium	France	Germany	Switzerland	Female group
$\boldsymbol{u^x}$	105,00	106,21	105,15	105,04	106,86
$\boldsymbol{L_{u^*}}$	269	1382	744	160	1238
$\hat{\boldsymbol{\xi}}$	-0,1219	-0,0353	-0,1346	-0,1033	-0,0445
$\boldsymbol{s.e.}\left(\hat{\boldsymbol{\xi}}\right)$	0,0585	0,0272	0,0284	0,0759	0,0285
$\hat{\boldsymbol{\beta}}$	1,6764	1,5048	1,5599	1,5908	1,4339
$\boldsymbol{s.e.}\left(\hat{\boldsymbol{\beta}}\right)$	0,1412	0,0576	0,0719	0,1739	0,0577
AD p-value	0,4096	0,1908	0,1027	0,7774	0,3856
AIC	754	3780	1953	439	3262

Table 4-6 GPD parameters fitting for the female sub-datasets

All the shape parameters are negative. The standard error of the parameter estimation is lower when there are more exceedances. According to the AIC, the model fitting quality has been improved for the Belgium, Germany, and Switzerland datasets, but not the France datasets.

5 Application

5.1 Maximum age at death

Recall the maximum age at death can be estimated by $\hat{\omega} = u - \frac{\hat{\beta}}{\hat{\xi}}$. The results are shown in the table below (Table 5-1 and Table 5-2). At the group level:

	Maximum age at death
Female group	137,23
Male group	124,99

Table 5-1 Maximum age at death estimated for the female datasets and the male datasets

At the subset countries level for the female group:

	Maximum age at death (female sub-datasets)
Belgium	118,75
France	148,79
Germany	116,74
Switzerland	120,44

Table 5-2 Maximum age at death estimated for the female sub-datasets

For the female group, the estimated maximum age at death is lower at the country level than the group level for Belgium, Germany, and Switzerland, for these countries the estimated maximum age at death is 118,75, 116,74 and 120,44, respectively. We believe these estimates are reasonable. However, the estimated maximum age at death at group (137,23) and of French female (148,79) are exceptionally high, which need more investigation.

Recall that in the empirical plots of the mean excess function at the group level and the French female datasets, we have observed a change in terms of the way that the remaining life expectancy decreases. The change points for the female group was around 110 years old, 108 for the male group and 110 for the French female datasets. Indeed, the remaining lifetime expectancy decreases slowly until these ages and drops dramatically afterwards. As the French female counts 69% of the total female group, the estimated results for the female group is largely impacted by the French female datasets. The decreasing trend can be measured by the force of resistance to mortality. When the force of resistance to mortality is strong, the decreasing trend of the remaining life expectancy will slow down, and the empirical plot of the remaining life expectancy will be close to a horizontal line. And when the force of resistance to mortality becomes weak, the remaining life expectancy drops. In the

GPD model, the force of resistance to mortality is quantified by the shape parameter ξ and the estimation of the ultimate age at death depends largely on choice of the threshold and the associated shape parameter. For the French female datasets, the selected threshold is 106,21 and the associated shape parameter is -0,0353, which situates in the range where the force of resistance to mortality is quite weak. As a result, the ultimate age at death projected based on this weak force of resistance will result in an overestimation. Even though the data beyond 106,21 can be considered as GPD distributed, the change in the shape parameters beyond the change point has not been taken into account for the estimation of the ultimate age. Same for the male group and the female group. A more reliable estimation would be based on the models starting from the change point: beyond 110 for the female group and the French subsets, beyond 108 for the male group. The corrected estimation is shown in the table below. The estimated ultimate age at death drops to 116,30, 116,79 and 117,00 for the French female subsets, the female group, and the male group, respectively (see Table 5-3) . These estimations are plausible compared with the maximum ages at death observed in our datasets, which is 115,12 and 111,88 for the female and the male, respectively. The AICs have also been improved for the new fitted models.

	Adjusted Maximum Age at Death		
	French female (>110)	Female Group level (>110)	Male Group level (>108)
$\boldsymbol{u^x}$	110,14	110,12	108,01
$\boldsymbol{L_{u^*}}$	85	109	46
$\hat{\boldsymbol{\xi}}$	-0,3093	-0,2555	-0,1347
$\boldsymbol{s.e.}\left(\hat{\boldsymbol{\xi}}\right)$	0,09695	0,08693	0,1879
$\hat{\boldsymbol{\beta}}$	1,9063	1,7037	1,2112
$\boldsymbol{s.e.}\left(\hat{\boldsymbol{\beta}}\right)$	0,26984	0,21639	0,286
AD p-value	0,3288	0,3119	0,0974
AIC	231	276	105,45
Ultimate age	116,30	116,79	117,00

Table 5-3 Adjusted Maximum Age at Death for the French female sub-datasets, the female group and the male group

5.2 One-year force of mortality at extreme ages

The one-year mortality rate q_x and the forces of mortality μ_x modeled by the generalized Pareto distribution beyond the high threshold age u^* can be estimated by

$$\hat{q}_x = 1 - \left[1 + \hat{\xi}\frac{1}{\hat{\beta} + \hat{\xi}(x - u^*)}\right]^{(-1/\hat{\xi})} \quad (21)$$

with the one-year force of mortality equals to

$$\widehat{\mu_x} = \frac{1}{\hat{\beta} + \hat{\xi}(x - u^*)} \quad (22)$$

The estimated one-year force of mortality for the female and the male datasets are displayed in Table 5-4, The estimated one-year force of mortality for the female sub-datasets from Belgium, France, Germany and Switzerland are displayed in Table 5-5.

	Estimated one-year force of mortality (group level)	
	Female group	Male group
106	0,509	0,555
107	0,520	0,576
108	0,532	0,629
109	0,544	0,681
110	0,524	0,742
111	0,600	0,809
112	0,699	0,882
113	0,824	0,951
114	0,960	0,994
115	∞	∞

Table 5-4 Estimated one-year force of mortality at group level

	Estimated One Year Force of Mortality (female sub-datasets)			
	Belgium	France	Germany	Switzerland
106	0,488		0,516	0,501
107	0,518	0,496	0,553	0,527
108	0,551	0,505	0,595	0,556
109	0,588	0,513	0,642	0,587
110	0,630	0,522	0,697	0,623
111	0,678	0,491	0,759	0,662
112	0,731	0,575	0,828	0,705
113	0,791	0,688	0,901	0,753
114	0,856	0,841	0,966	0,805
115	0,921	0,991	0,998	0,860
116	0,975	∞	∞	0,915
117	0,999			0,964
118	∞			0,994
119				∞

Table 5-5 Estimated one-year force of mortality for the female sub-datasets from Belgium, France, Germany and Switzerland

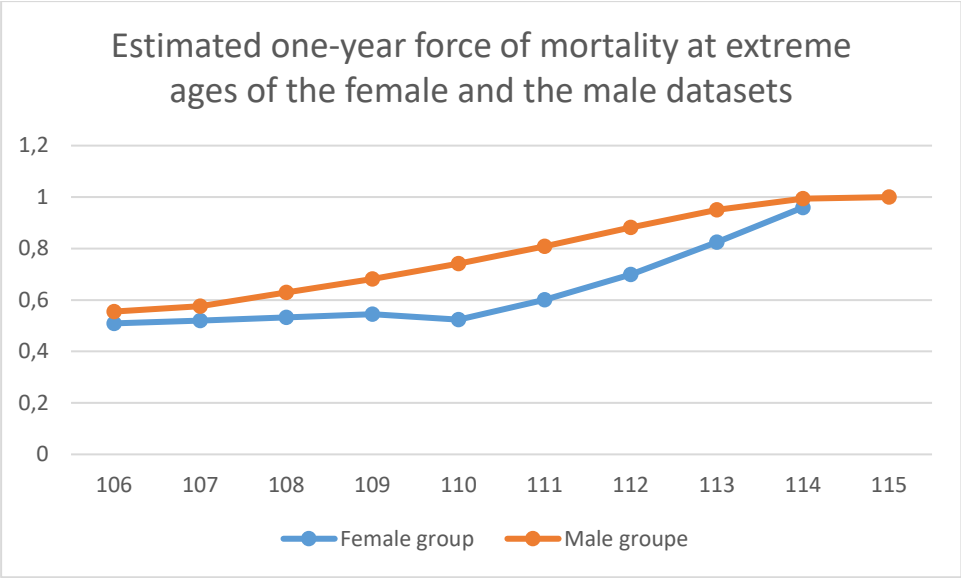

Figure 5-1 Estimated one-year force of mortality at extreme ages of the female and the male datasets

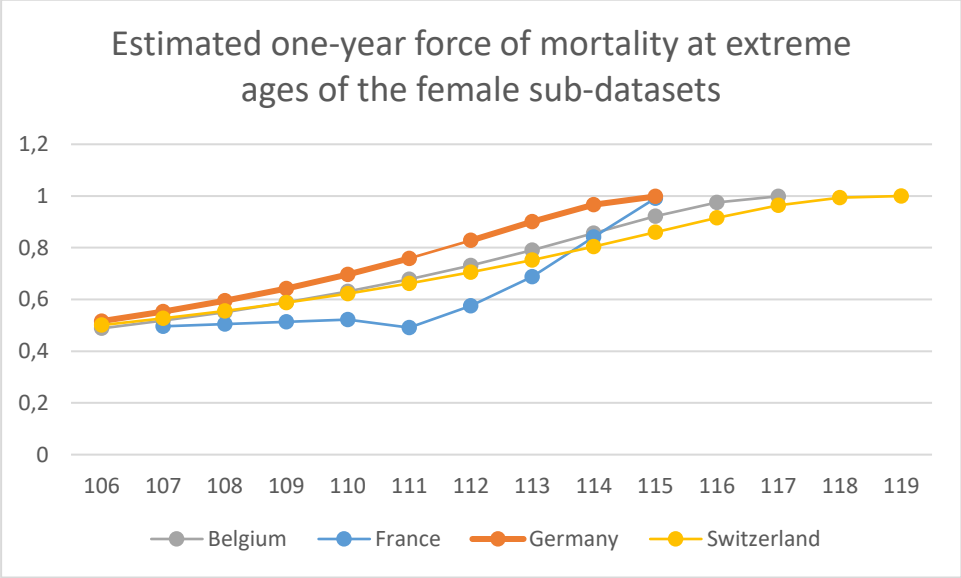

Figure 5-2 Estimated one-year force of mortality at extreme ages of the female sub-datasets

5.3 Survival probability at extreme ages

The survival probability between $[105, \hat{\omega})$ can be estimated by

$$ {}_{x}\hat{q}_{105} = \frac{L_{u^*}}{L_{105}}\left(1 - G_{\hat{\xi};\hat{\beta}(.)}(x - u^*)\right) \tag{23}$$

The survival probability beyond the thresholds for the female and the male datasets is displayed in Table 5-6

Estimated survival probability for female and male datasets		
	Female group	Male group
106		0,48348008
107	0,27373844	0,22474147
108	0,13436803	0,09999089
109	0,06443906	0,04499908
110	0,03014478	0,01670254
111	0,01527443	0,00532006
112	0,00727054	0,00137493
113	0,00290643	0,00026250
114	0,00087573	0,00003105
115	0,00015393	0,00000153
116	0,00000623	0,00000001

Table 5-6 Estimated survival probability for the female datasets and the male datasets

The survival probability beyond the thresholds for the female sub-datasets are displayed in Table 5-7

Estimated survival rates at country level (Female)				
	Belgium	France	Germany	Switzerland
106	0,538312225		0,505951454	0,520028593
107	0,275481637	0,287262562	0,244758937	0,259597570
108	0,132819572	0,144759765	0,109438621	0,122810051
109	0,059632235	0,071718139	0,044368647	0,054564615
110	0,024549437	0,034901682	0,015872901	0,022508143
111	0,009073740	0,014229828	0,004812807	0,008492113
112	0,002922471	0,005371944	0,001160550	0,002872402
113	0,000784698	0,001750019	0,000199513	0,000847293
114	0,000163803	0,000466705	0,000019762	0,000209494
115	0,000023582	0,000093375	0,000000676	0,000040897
116	0,000001855	0,000011935	0,000000001	0,000005724

Table 5-7 The survival probability at extreme ages for the female sub-datasets from Belgium, France, Germany and Switzerland

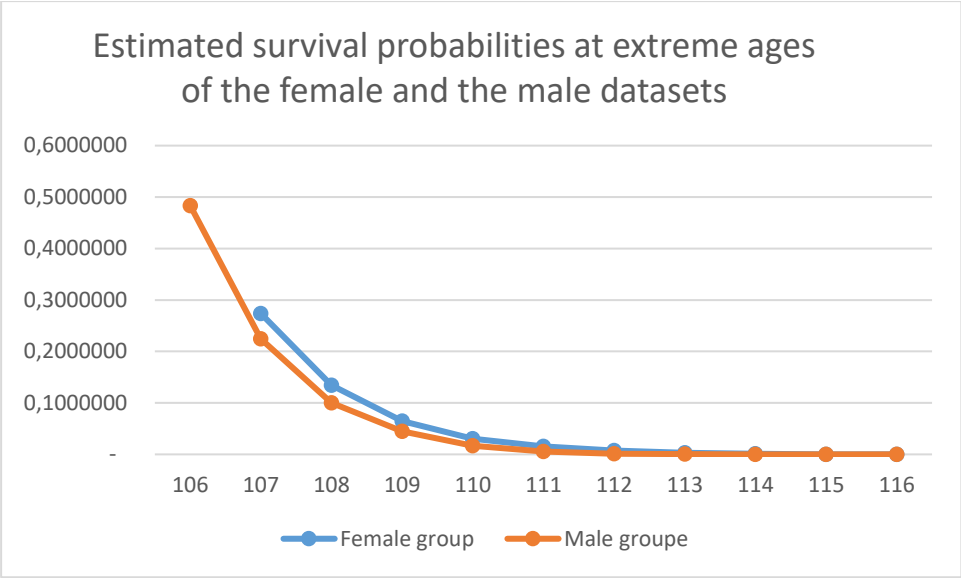

Figure 5-3 Estimated survival probabilities at extreme ages of the female and the male datasets

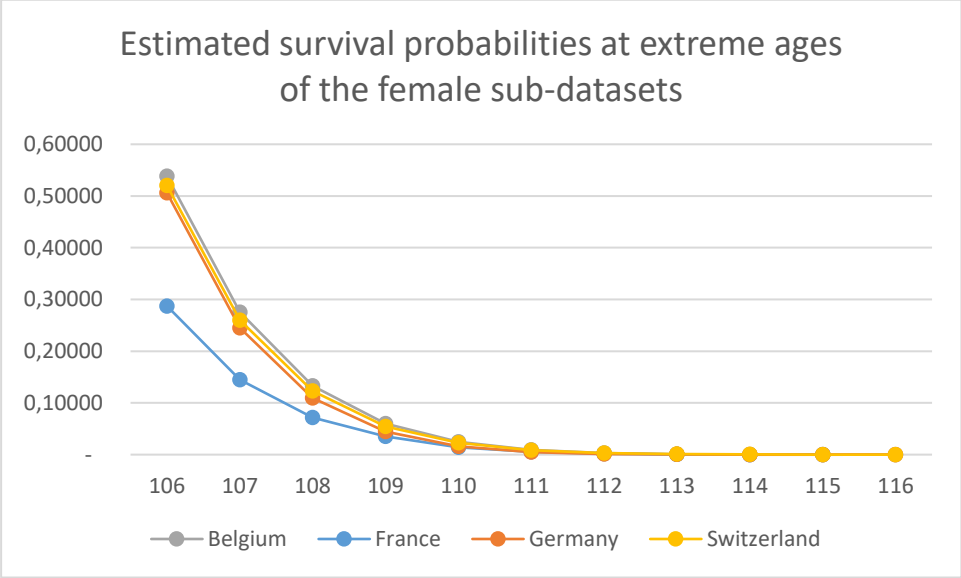

Figure 5-4 Estimated survival probabilities at extreme ages of the female sub-datasets

6 Conclusion

We have applied the method using the extreme value theorem to analyze the mortality at the extreme ages (>105) in Belgium, France, Germany, and Switzerland for the birth cohorts from 1881 to 1898. After examining the homogeneity, we analyzed the data in two steps: first at group level and secondly at subsets level for the female datasets. An automated selection procedure based on the sequential goodness-of-fit tests combined with forward stopping rule has been used to select the appropriate thresholds for each dataset. We find all the datasets have a negative tail index which supports the existence of an ultimate age. We also noticed that the estimated ultimate age is very sensible to the selection of the threshold and the associated shape parameter. Indeed, a change in the decreasing trend shown on the empirical mean excess plots around a certain age suggests the shape parameters may also change. It is the case in our French female data and the male data, for which the threshold selection method should be applied separately on the datasets before and after that changing point, otherwise the estimated ultimate age can be biased.

At the group level, the ultimate age at death for both female datasets and male datasets is around 116 years old. At the country level, the Switzerland female has the highest estimated age at death 120,44, after comes the Belgian female 118,75, the German female 116,74 and the French female 116,30. It seems the estimation of the ultimate age at death depends largely on the trend in the force of resistance to mortality carried by the data. In the Belgian and Switzerland female datasets, we have not observed an accelerated decreasing trend at the end of the extreme ages while this trend is more visible for the German and French datasets.

Lastly, the one-year force of mortality and the survival probability at extreme ages have been elaborated, which provides a practical view on the closing of the life table.

7 Bibliography

"Ageing in the 21st Century: Chapter 1: Setting the Scene." *HelpAge International*. https://www.helpage.org/resources/ageing-in-the-21st-century-a-celebration-and-a-challenge/ageing-in-the-21st-century-chapter-1/ (July 5, 2020).

Bader, Brian, Jun Yan, and Xuebin Zhang. 2018. "Automated Threshold Selection for Extreme Value Analysis via Ordered Goodness-of-Fit Tests with Adjustment for False Discovery Rate." *The Annals of Applied Statistics* 12(1): 310–29.

Barbi, Elisabetta et al. 2018. "The Plateau of Human Mortality: Demography of Longevity Pioneers." *Science* 360(6396): 1459–61.

Belzile, Léo R., Anthony C. Davison, Holger Rootzén, and Dmitrii Zholud. 2020. "Human Mortality at Extreme Age." https://arxiv.org/abs/2001.04507v1 (July 8, 2020).

Benjamini, Yoav, and Yosef Hochberg. 1995. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society. Series B (Methodological)* 57(1): 289–300.

Bravo, Jorge Miguel, and Pedro Corte-Real. 2012. "Modeling Longevity Risk Using Extreme Value Theory: An Empirical Investigation Using Portuguese and Spanish Population Data." *North American Actuarial Journal*: 31.

Choulakian, V, and M. A Stephens. 2001. "Goodness-of-Fit Tests for the Generalized Pareto Distribution." *Technometrics* 43(4): 478–84.

Einmahl, Jesson J., John H. J. Einmahl, and Laurens de Haan. 2019. "Limits to Human Life Span Through Extreme Value Theory." *Journal of the American Statistical Association* 114(527): 1075–80.

Gavrilova, Natalia S., Leonid A. Gavrilov, and Vyacheslav N. Krut'ko. 2017. "Mortality Trajectories at Exceptionally High Ages: A Study of Supercentenarians." *Living to 100 monograph* 2017(1B). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5696798/ (July 5, 2020).

Gbari, Samuel, Michel Poulain, Luc Dal, and Michel Denuit. 2017. “Extreme Value Analysis of Mortality at the Oldest Ages: A Case Study Based on Individual Ages at Death.” *North American Actuarial Journal* 21(3): 397–416.

G’Sell, Max Grazier, Stefan Wager, Alexandra Chouldechova, and Robert Tibshirani. 2015. “Sequential Selection Procedures and False Discovery Rate Control.” *arXiv:1309.5352 [math, stat]*. http://arxiv.org/abs/1309.5352 (July 3, 2020).

De Haan, Laurens, and Ana Ferreira, eds. 2006. “Limit Distributions and Domains of Attraction.” In *Extreme Value Theory: An Introduction*, Springer Series in Operations Research and Financial Engineering, New York, NY: Springer, 3–36. https://doi.org/10.1007/0-387-34471-3_1 (June 12, 2020).

Hanayama, Nobutane, and Masaaki Sibuya. 2016. “Estimating the Upper Limit of Lifetime Probability Distribution, Based on Data of Japanese Centenarians.” *The Journals of Gerontology: Series A* 71(8): 1014–21.

Huang, Fei, Ross Maller, and Xu Ning. 2020. “Modelling Life Tables with Advanced Ages: An Extreme Value Theory Approach.” *Insurance: Mathematics and Economics* 93: 95–115.

Kathryn A. Watts MSc, ASA, Debbie J. Dupuis PhD, and FCIA Bruce L. Jones PhD FSA. 2006. “An Extreme Value Analysis Of Advanced Age Mortality Data.” *North American Actuarial Journal* 10(4): 162–78.

Li, J. S. H., A. C. Y. Ng, and W. S. Chan. 2011. “Modeling Old-Age Mortality Risk for the Populations of Australia and New Zealand: An Extreme Value Approach.” *Mathematics and Computers in Simulation* 81(7): 1325–33.

Medford, Anthony, Kaare Christensen, Axel Skytthe, and James W. Vaupel. 2019. “A Cohort Comparison of Lifespan After Age 100 in Denmark and Sweden: Are Only the Oldest Getting Older?” *Demography* 56(2): 665–77.

Scarrott, C., and Anna MacDonald. 2012. “A Review of Extreme Value Threshold Estimation and Uncertainty Quantification.” *Revstat Statistical Journal* 10: 33–60.

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
Faculté des sciences
Place des sciences, 2 bte L6.06.01, 1348 Louvain-la-Neuve, Belgique | **www.uclouvain.be/sc**