Extrapolative Projections of Mortality: Towards a More Consistent Method. The Central Scenario and Study of Estimation Errors

Dalkhat M. Ediev
Vienna Institute of Demography

The paper presents an adjusted version of the method of direct extrapolation of mortality by age and sex. The method is supplemented by additional procedures in order to improve its efficiency in the short-run and preclude implausible mortality patterns in the long-run. The short-run efficiency is improved by building the forecast on data from the most recent periods of age-sex-specific duration, when mortality dynamics exhibits steady trend. In the long-run, the rates of mortality decline are assumed to converge to a plausible function of age and sex, which is derived from the data based on the assumption that it is a monotonic function of age. Efficiency of such improvements is supported by data and forecasting results. The framework proposed provides a natural basis for introducing uncertainty into the projection. Based on preliminary structure of the probabilistic model, simulations were run to study estimation errors of model parameters and autocorrelations involved. They indicate that estimates of some of the parameters are ineffective and need further research.

Table of contents

1. Introduction

In projecting life expectancy, it is often considered to be implausible to assume unchecked continuation of past improvements into future, because improvements in past were mainly due to decline in mortality at young ages, while recent trends as well as those expected in future, are attributed to decline in mortality at older ages in most of modern populations (e.g., Wilmoth 2005). Relative to the decline of mortality at young ages, decline at old ages has more moderate effect on life expectancy at birth (Keyfitz 1985) and might be limited if there are biological limits to human longevity[1].

A practical way to address these concerns is to disaggregate and extrapolate mortality by age, sex and, possibly, other relevant variables. Although, individual mortality rates may also experience turning points, the approach is likely to be more consistent as it explicitly reflects roles of dynamics of mortality rates at different age groups. After a proper transformation, individual age-sex specific mortality rates do show consistent trends, which may be extrapolated rather robustly. The approach has been used in many applications and model frameworks. The widely used extrapolative method by Lee and Carter (1992) (the method was also proposed by Gómez de León (1990); see also Continuous Mortality Investigation 2007), method based on extrapolating Wang-

[1] The dispute on these limits and on future growth in life expectancy is not yet resolved, however (e.g., Olshansky et al. 2001, Oeppen and Vaupel 2002, De Grey 2006).

transforms of death rates (de Jong and Marshall 2007) and methods based on extrapolating parameters of mortality models (e.g., McNown and Rogers 1989, Lutz et al. 1997) or on shifting the age profile of mortality rates (Bongaarts and Feeney 2003, Bongaarts 2005) also represent approach to mortality projection by age-sex groups (see also some other models in Cairns et al. 2008). Further decomposition of mortality, e.g., by causes of death, has not found enough justification as it leads to unjustified complication of the method and makes projection assumptions less transparent, while potentially resulting in internal inconsistencies with respect to interrelations between components of mortality (Wilmoth 2005).

Limitations of projecting mortality decomposed into age-sex-specific rates are linked to a possible lack of correspondence between *projected* individual trends. Assuming, for example, mortality to decline at every age by a constant rate observed from the past is likely to result in erratic developments in age structure of mortality, including crossovers of projected age-specific mortality trends. In such case, special arrangements are to be made in order to avoid inconsistent developments in projected trends. U.S. Census Bureau, for example, has once used the following conditions imposed on projected rates:

1. No 2050 death rate was allowed to be higher than it was in 1994
2. No male rate was allowed to ever be lower than the equivalent female rate
3. Within a given race-sex group, the death rates must steadily rise from age 25-29 to 100+
4. No death rate was permitted to improve more than 3 percent per year during the 1994 to 2050 period.

(Day and U.S. Census Bureau 1996).

In another case, annual rates of mortality decline by age and sex were projected to converge to a target value for 2039 and were assumed to be identical afterwards (Gallop 2007). In still other examples, projected age-specific rates of mortality decline were kept different and, yet, taken from expert judgment rather than directly from past observations (Hollmann et al. 2000, Wilmoth 2005).

Despite apparent shortcomings, extrapolation remains to be a common method for official projections (e.g., Social Security Administration 1987, Day and U.S. Census Bureau 1996, Wilmoth 2005, Gallop 2007), which may also be attributed to the fact that extrapolation – though, wrongly in many instances – is often considered to be free of subjective judgment. Expert judgment, however, has proven its usefulness in demographic projections (e.g., Lutz et al. 1996, 1999); and it is reasonably argued that 'blind' extrapolation may not free us from necessity of substantive analysis. Yet, extrapolation based on statistical methods, perhaps, retains its central role in mortality projection, both as a tool as such and a framework for shaping the expert knowledge.

The purpose of this paper is to improve the performance and consistency of the most straightforward method of directly extrapolating mortality rates by age and sex. The paper focuses on the core engine of the method, i.e., on generating the central projection. Yet, it also presents preliminary version of the probabilistic model built on the proposed deterministic model. Using the models developed, extensive simulations were run in order to study estimation errors of model parameters.

Next two parts present the deterministic model and its illustrative applications, as well as study in forecast errors. Part 4 presents, preliminarily, probabilistic model based on the deterministic structure presented in the first part. Last parts of the paper present results of simulations-based studies of estimation errors of model parameters and concluding remarks.

2. Improved direct extrapolation method for the central scenario projection

Direct (linear) extrapolation of log-mortality rates may be formalized by the following model:

$$\ln[m(x,t)] = a_x + b_x t + \varepsilon_{x,t}, \quad (1)$$

here $m(x,t)$ is the central death rate for age x at time t; a_x, b_x - are the model parameters; and $\varepsilon_{x,t}$ is the error term[2]. Although the model (1) may formally be applied to each of the age-sex groups separately, it might be necessary to take into account cross-time and cross-age-sex autocorrelations of the error terms. We examined these autocorrelations elsewhere (Ediev 2007, 2008) in the context of comparing the direct extrapolation model (1) to the Lee and Carter model. Generally, conclusions we come to based on the study of autocorrelations are the following. *Firstly*, the LC model does not seem to be better fit to data compared to fitting separate trends. *Secondly*, fitting optimal trends to optimal data periods at different ages may significantly reduce extrapolation errors in the short-run. *Thirdly*, cross-time correlations, which usually disappear in about five years lag, suggest that such correlations may be ignored in estimating trend parameters, only if the data period covers not less than several decades. *Fourthly*, for the same reason as above, peculiarities of the most recent observation ('*jump-off*' year) might be important to take care of in the projecting to the nearest future (five to ten years); yet, for a longer-range projections their effect may be neglected[3]. *Fifthly*, cross-sex correlations may usually be neglected in parameters estimating. *Sixthly*, age-sex-specific trends, being *estimated* separately, may not be *projected* independently. Measures are to be taken in order to reconcile consistency in projections, which otherwise may turn implausible due to continuation of temporary divergent developments in data.

This analysis implies importance of developing a direct extrapolation method, which would meet the following conditions:

(i) Age-specific trends in mortality are to be examined and estimated separately.

(ii) The method should make use of the flexibility implied by separate consideration of age-sex-specific trends. In particular, when estimating trend parameters, it should be possible to detect and use data period of optimal length for each age-sex-specific death rate[4].

(iii) As the age-specific trends are estimated independently, the method should include some sort of reconciliation procedures, which will prevent forecasted age-sex profiles of death rates from turning implausible.

(iv) The method should be capable of taking into account additional information contained in the last observation (the projected trend should be consistent with the jump-off value).

Based on these considerations, we propose the following multi-stage method of direct extrapolation of death rates.

Step 1. Estimation of parameters of separate age-sex-specific linear trends.

Let us notate $\eta_{x,t} \overset{def}{=} \ln[m(x,t)]$. Model (1) may then be rewritten as follows:

[2] Hereinafter, we consider explicitly only the age variable, although the models considered may also include sex, race and other variables, which might be of interest in practice.

[3] Indeed, the cross-time correlations might still be important for a probabilistic projection, as they might affect typology of possible future paths. We do not focus, however, on probabilistic projections in this paper.

[4] We do not consider here possibility of non-linear trends of different type at different age-sex groups.

$$\eta_{x,t} = a_x + b_x t + \varepsilon_{x,t}\,. \tag{2}$$

Linear trends are not always relevant to the entire data set. Hence, only the recent period of data, which supports linearity assumption, is to be used in parameters' estimation. (Among other things, this will reduce forecasting errors in the short-run.) Formally, one should take into account possible autocorrelations of errors $\varepsilon_{x,t}$ to estimate parameters in (2). In a simpler approach proposed here, however, these correlations are ignored. (Yet, having them in mind, it is imposed that the data period may not be shorter than 20 years.) Standard error of the linear-regression forecast at time t is given by the formula:

$$\sigma_{\eta(x,t)} = \sqrt{E\left(\eta_{x,t} - \hat{\eta}_{x,t}\right)^2} = \sqrt{E\left(\eta_{x,t} - \left(\hat{a}_x + \hat{b}_x t\right)\right)^2} = \sigma_\varepsilon \sqrt{1 + \frac{1}{n} + \frac{\left(t - \bar{t}\right)^2}{n\left(\overline{t^2} - \bar{t}^2\right)}}\,, \tag{3}$$

hereinafter, cups denote estimates and forecasts, upper lines denote arithmetic averaging over the data period; n is the number of observations used in estimation, and σ_ε is standard deviation of the error term in (2), which may be estimated from the residuals. For estimating the longest possible recent data period with linear trend, the following relative deviations from the trend are analyzed for each possible beginning year t of the period:

$$\varphi_{x,t-d} = \frac{\left|\eta_{x,t-d} - \hat{\eta}_{x,t-d}\right|}{\hat{\sigma}_{\eta(x,t-d)}}\,, \tag{4}$$

where d is the time lag used in checking, whether the observation at year $t-d$ fits to the trend observed after t. At $d=1$, the procedure would provide check of fitness to the trend of the observation right preceding to the first observation included into the data period. Such choice, however, is not feasible, as for a short time lag, it is not possible to distinguish between trend change and random deviation from the trend. Although, the optimal choice for d could be a function of expected magnitude of trend change and of errors (3), we simplify the procedure by setting $d=10$. (However, $d = \min\left(10, t - t_{\min}\right)$ when t approaches to the very first year of observation to avoid extending beyond the available data.) The following simple rule is used, then, to detect the optimal starting year for the data period:

$$t_{start} = \min_{t \le t\,\max} \left\{ t : \varphi_{x,t-d} \ge \varphi_{\max} \right\} - 1\,, \tag{5}$$

where $\varphi_{\max} = 2$ is the threshold value, above which the deviation is considered to be significantly inconsistent with the trend, and $t_{\max}$ is the latest possible year for starting the data period. As it was mentioned above, $t_{\max} = t_1 - 20$ is 20 years prior the last observation (by t_1 we denote the last year in the data period, i.e. the year preceding the projection period). Additionally, we smooth estimates (5) by applying 5-year moving averaging to eliminate erratic variations. Some examples of such estimates are provided further down, in the bottom parts of the graphs presented in figure 2.

After detecting optimal data periods we estimate parameters of the separate age-sex-specific trends. Existence of cross-time autocorrelations—which is the case of countries with large population—may preclude from basing the estimation on short periods of data or doing it without including the autocorrelations into estimation procedures. One should also be aware that positive cross-time autocorrelations may significantly distort all the autocorrelation estimates, even when such estimates are based on relatively long periods of data (see, e.g., Box and Jenkins 1970, Bartlett 1946). In mortality data, duration of period when data exhibit more or less suitable linear trend is of several decades, which

is too short for the distorting effects of autocorrelations to be neglected. Therefore, it may not be considered reliable to estimate too much of correlation information from such data. For the reasons outlined, we propose here to neglect cross-age-sex correlations and obtain only a stylized correlation patterns for cross-time autocorrelations:

$$Correl\left(\varepsilon_{x,t}, \varepsilon_{x,t-k}\right) = \rho^k, \tag{6}$$

where ρ is the autocorrelation coefficient for one year lag. This coefficient may be estimated based on residuals of linear trends preliminarily fit to some period, when data are presumably showing linearity. Note that the correlation parameter ρ is an average over all age groups; therefore, data necessary for its calculation will be abundant even for a relatively short period; and period of one to three decades may usually fit the task. An advantage of having the same estimate for all age groups is that this estimate may be considered as virtually not correlated to data and estimations of the model parameters for each individual age group.

Under assumed autocorrelations (6), parameters in (2) may be estimated by the common formula of the Least Squares method (see details in Ediev 2008):

$$\hat{b} = \frac{\overline{yt} - \bar{y} \cdot \bar{t} + c\frac{n-1}{n}\overline{\Delta_y}}{\overline{t^2} - \bar{t}^2 + c\frac{n-1}{n}}, \tag{7}$$

$$\hat{a} = \bar{y} - \hat{b}\bar{t}, \tag{8}$$

where $c = \frac{\rho}{1-\rho^2}$ and $\Delta_{yt} = y_{t+1} - y_t$. (Relations (7), (8) turn into OLS formulas when autocorrelations are neglected, i.e., $c = 0$.)

Estimates (7), (8) are unbiased and covariances of parameters' errors are given by:

$$E\left[\begin{pmatrix}\hat{a}-a\\ \hat{b}-b\end{pmatrix}\begin{pmatrix}\hat{a}-a\\ \hat{b}-b\end{pmatrix}^T\right] = \frac{1}{n}\frac{\sigma^2}{\overline{t^2} - \bar{t}^2 + c\frac{n-1}{n}}\begin{pmatrix}\overline{t^2} + c\frac{n-1}{n} & -\bar{t}\\ -\bar{t} & 1\end{pmatrix}, \tag{9}$$

where index 'T' denotes transposition and σ^2 is the dispersion of the error term in (2), unbiased estimate for which is given by the following expression (see self-sufficient derivation in Ediev 2008):

$$\hat{\sigma}_\varepsilon{}^2 = \frac{\hat{E}^T C^{-1} \hat{E}}{n-2} = \frac{1}{(n-2)(1-\rho^2)}\left(\sum_{i=1}^{n} e_i{}^2 + \rho^2 \sum_{i=2}^{n-1} e_i{}^2 - 2\rho \sum_{i=1}^{n-1} e_i e_{i+1}\right), \tag{10}$$

where $e_i = y_i - \hat{y}_i$ - are the residuals of the model. From (9), (10), unbiased estimates for quadratic errors of the parameters' estimates are given by the following expressions:

$$\hat{\sigma}_a{}^2 = \frac{1}{n}\hat{\sigma}_\varepsilon{}^2\left(1 + \frac{\bar{t}^2}{\overline{t^2} - \bar{t}^2 + \frac{\rho}{1-\rho^2}\frac{n-1}{n}}\right), \tag{11}$$

$$\hat{\sigma}_b{}^2 = \frac{1}{n}\hat{\sigma}_\varepsilon{}^2\frac{1}{\overline{t^2} - \bar{t}^2 + \frac{\rho}{1-\rho^2}\frac{n-1}{n}}. \tag{12}$$

Step 2. Estimation of long-run plausible schedule of mortality decline rates b_x^*.

As it was described above, a drawback of extrapolative methods for projecting age-sex-specific mortality rates is that age-sex-specific trends, being extrapolated independently,

may produce implausible age-sex patterns of mortality. Therefore, we develop special procedures of correcting the estimates of the slopes in order to ensure consistency of the projection in the long-run. Such corrections may also improve the forecasting efficiency, as they utilize additional empirical knowledge about regularities in age structure of mortality, which is neglected when estimating age-specific mortality trends separately. To avoid reducing forecasting efficiency at individual ages, there should be no significant correction while projecting to the nearest future, while corrections should be more pronounced in the long-run, when changes of trends for individual ages are more likely.

The idea is to assume that age-specific slopes b_x gradually converge to eventual slopes b_x^*, the latter constructed in a way granting plausibility of long-term mortality dynamics. As mentioned in the introduction, such approach was already used in the literature, with eventual schedule b_x^* obtained from external, e.g., expert, considerations. Here, we propose deriving this schedule directly from estimated rates b_x.

As one may usually note from data, age-specific rates of mortality change (usually negative, as mortality was declining in the past) tend to increase with age and to be higher for males compared to females (Ediev 2008, see some examples in Fig. 2; see also a similar observation in Goss, et.al. 1998). At ages above 30 or so such pattern is, indeed, expectable theoretically, as violation of it would result, in the long run, in implausible mortality profile. Data, however, indicate that the same monotonic pattern may probably be extended to young ages as well. To study this tendency of age-specific rates of mortality change, we conducted extensive calculations based on data from different populations and on different periods of data. Typically, the age-sex-specific rates of mortality decline exhibit monotonic pattern, albeit with significant deviations from the overall monotonic tendency. The deviations from the overall tendency seem to be correlated between different populations. This is a natural reflection of spread of the medical and sanitary technologies. To illustrate the overall tendency of age-sex-specific rates of mortality decline, we present averages of estimates $\hat{b}_x$ obtained for different periods of time for six populations with relatively long history of mortality statistics available (see fig. 1). For each country, rates of mortality decline were estimated using the aforementioned procedures with different choice for the last year of observation. The last year of observation was varying from the earliest possible year until the latest available with step 10 years, i.e., for example, it was taking values 1780, 1790, 1800, …, 2005 for Sweden[5]. (For other countries, period covering the World War II was excluded from calculations; inclusion of those data does not alter our findings.) In each case, we estimated optimal durations of data periods by age and sex and obtained rates of mortality decline from (8). Graphs in fig. 1 present average values over estimates for all six countries based on all data periods analyzed. Curves of age-sex-specific rates of mortality decline are also supplemented by error intervals, which are obtained by increasing and decreasing the average values by one standard deviation of the estimates obtained. Evidently, results presented on fig. 1 support our hypothesis about monotonic increase by age of rates of mortality change (and also about those rates being higher for males compared to females[6]). Although, the monotonic pattern could be moderately violated at young ages and

[5] I acknowledge important conversation with Hans Lundström, who pointed that Swedish life tables prior to 1860 are significantly based on indirect reconstructions, which may undermine findings based on that period. Yet, restricting data to 1860+ does not alter our findings, so we keep using the entire data set from HMD.

[6] Empirical findings presented imply that one may also assume *similarity* of female and male rates of mortality decline in the long run. Such a choice might be valuable for populations with only a short data period available.

especially at age 0, these violations of the overall tendency are not significant in view of the variance of the estimated rates.

Based on these empirical and formal considerations, we assume b_x^* to increase monotonically with age (and also to be higher or equal for males compared to females at similar age). A simplified Min-Max method for deriving such a monotonic function was proposed earlier (Ediev 2007). Here we use another method[7], based on obtaining consistent slopes $\hat{b}_x^*$ as a solution of the following optimization problem[8]:

$$Z\left[b_x^*\right] = \sum_{x=0}^{X} \frac{\left(\hat{b}_x - b_x^*\right)^2}{\hat{\sigma}_{bx}^2 + \breve{\sigma}_*^2} \to MIN \tag{13}$$

under constraints:

$$\begin{cases} \hat{b}_{x+1}^* \geq \hat{b}_x^*, \\ \hat{b}_{x,MALES}^* \geq \hat{b}_{x,FEMALES}^*, \end{cases} \tag{14}$$

here $\breve{\sigma}_*^2$ is a preliminary estimate of the dispersion of residuals $\delta_x = b_x - b_x^*$ (as a rough estimate, we use standard error of linear trend fit to $\hat{b}_x$) and standard errors of estimations $\hat{\sigma}_{bx}$ are obtained from (12). One may also use additional constraints: $b^{Min} \leq \hat{b}_x^* \leq b^{Max}$, with b^{Min} and b^{Max} selected from additional considerations (e.g., it might be plausible to set $b^{Max} = 0$ in many cases to prevent long-run mortality increase). Values of $\hat{b}_x^*$ obtained from (13) usually form a piece-wise constant function and may be smoothed by a moving averaging procedure:

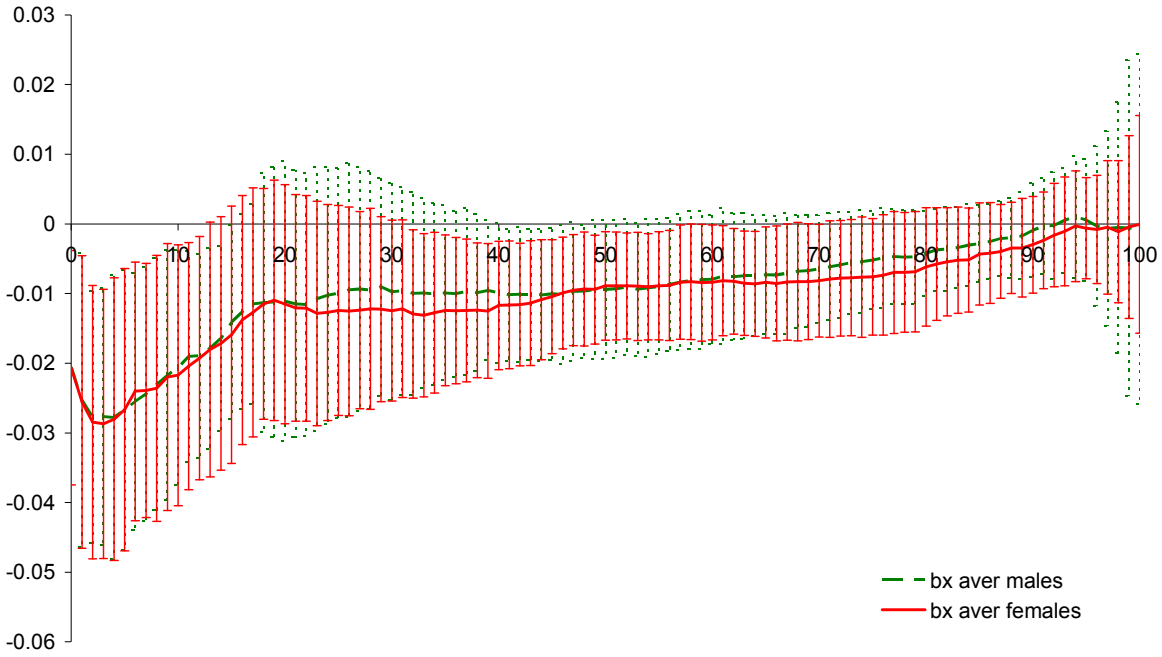
$$\tilde{b}_x^* = \frac{1}{2m+1} \sum_{y=x-m}^{x+m} \hat{b}_y^*, \tag{15}$$

where $2m+1$ is the length of the smoothing frame set at 11 years in the paper (at boundary age groups this frame is shortened accordingly). Typical examples are presented on figure 2, where slopes' estimates $\hat{b}_x$ as well as consistent slopes $\tilde{b}_x^*$ are presented for eight populations. Optimal age-sex specific starting years for the data periods are also shown in the bottom parts of the graphs presented in the figure.

[7] Extensive simulations conducted suggest, however, that both the method used here and the previously proposed simpler Min-Max methods are of similar performance.

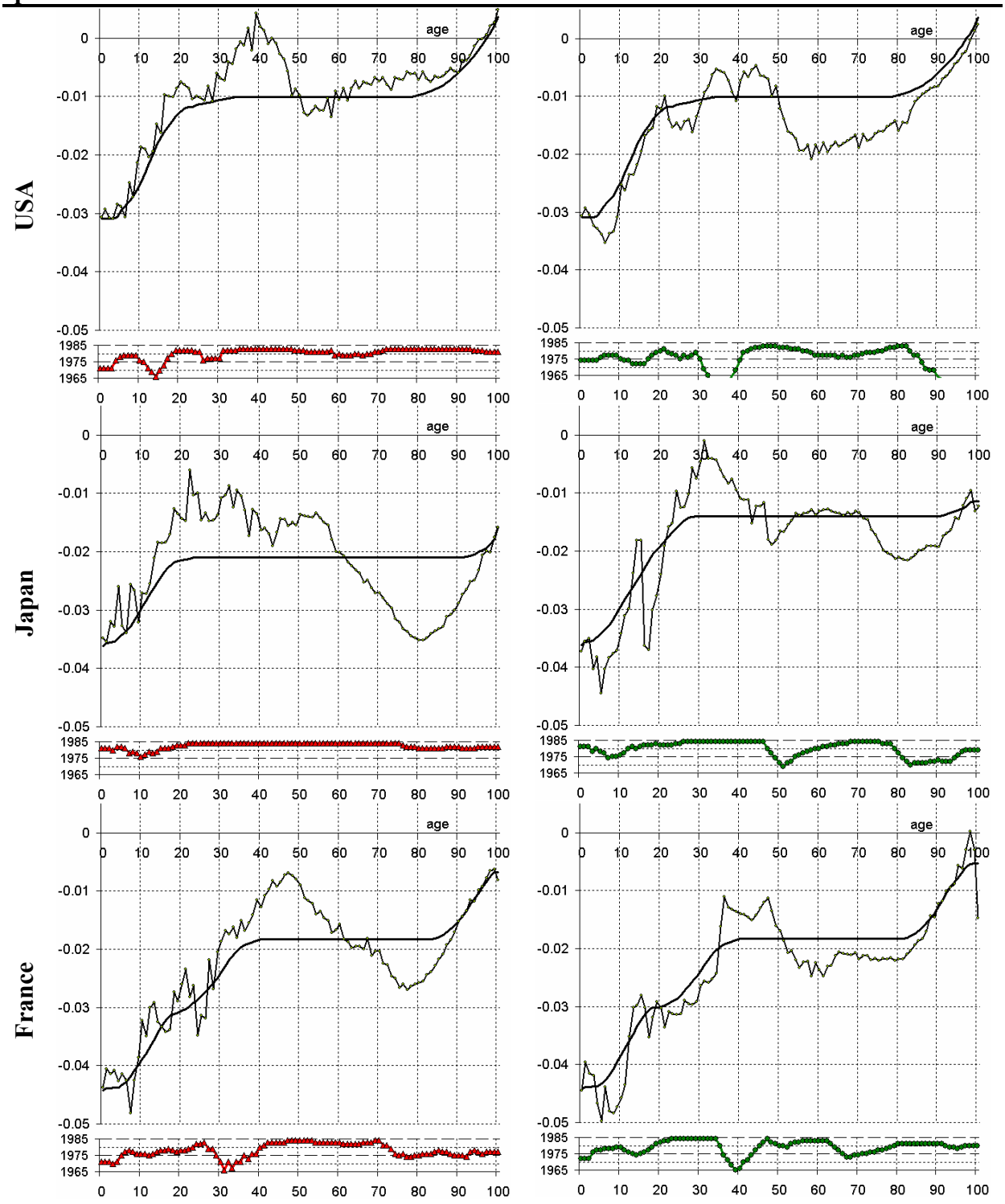
[8] The problem may conveniently be solved by the *dynamic programming* method.

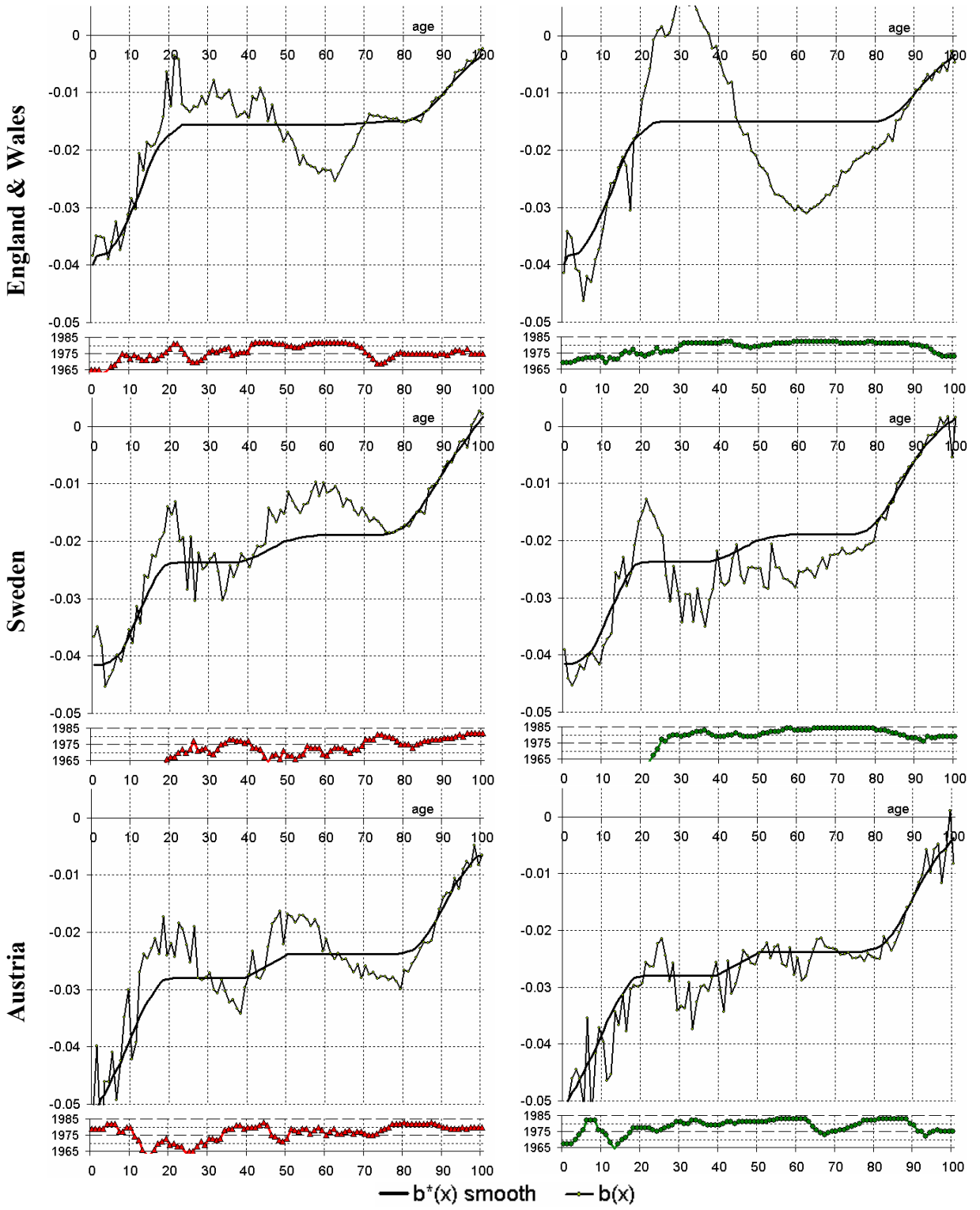
Figure 1. Averaged estimates and the standard errors for age-sex-specific rates of mortality decline for Denmark (1835-2005), England & Wales (1841-2005), Finland (1878-2005), Italy (1872-2003), Norway (1846-2005), and Sweden (1760-2005).



Notes: For each country, the end year of the data period was varied from the earliest possible to the latest available years by step five years. Error gaps (broken lines for males, solid lines for females) correspond to upward and downward deviations by one standard deviation as estimated from data.

Figure 2. Slopes of age-specific most recent optimal linear trends and corresponding consistent slopes obtained from (13): females (left column) and males (right column). In the bottom of each graph age-specific optimal beginning years of the trend fitting period are shown.





Step 3. Calculation of convergence parameters.

There are two sorts of convergence implied by the method. Firstly, the log-death rates are extrapolated as converging to their trends (from initial jump-off values). Secondly, slopes of the trends, in turn, converge to long-run rates of mortality change $\hat{b}_x^*$.

The speed of convergence to the trend from the jump-off value must depend on cross-time autocorrelations of residuals from linear trends in the past. Insignificant cross-

time autocorrelation may suggest ignoring the jump-off value and immediately starting the projection from the trend value. On the other hand, high cross-time autocorrelations may suggest gradual movement from the jump-off value to the trend. Under the assumed simple structure of these autocorrelations (6), one may use the following exponential model in deterministic setup, which corresponds to AR(1) process for residuals in stochastic setup:

$$\varepsilon_{xt} = \rho\varepsilon_{xt-1}. \quad (16)$$

The second kind of convergence—of trend slopes $\hat{b}_x$ to their long-run values $\hat{b}_x^*$—is modeled based on the estimated durations (5) of the most recent data periods consistent with linearity assumption. We model this convergence by assuming that age-sex-specific trends of log-death rates will be piece-wise linear in future, i.e., after some period of time they may change their slope (see similar construction in Sanderson and Scherbov 2004). Correspondingly, we consider the eventual slopes $\hat{b}_x^*$ as expected values for trend slopes in the long-run. We estimate the expected duration of period, during which trends remain linear, as

$$\lambda = \overline{t_1 - t_{start,x}}. \quad (17)$$

Note, that we obtain (17) by averaging the observed durations over all age and sex groups. Based on (17), we assume that probability of trend change at any given year is

$$\pi = 1/\lambda. \quad (18)$$

Correspondingly, in the deterministic setup, we assume that deviations of age-sex-specific slopes $\hat{b}_x$ from the eventual values $\hat{b}_x^*$ exponentially converge to zero following the model:

$$\hat{b}_{xt} - \hat{b}_x^* = (1-\pi)^{t-t1}\left(\hat{b}_{xt1} - \hat{b}_x^*\right). \quad (19)$$

The deterministic Direct Extrapolation Model: Summary

The direct extrapolation model for the central deterministic projection may be summarized as follows:

$$\hat{\eta}_{xt1} = \hat{a}_x + \hat{b}_x \cdot t_1,\ \hat{b}_{xt1} = \hat{b}_x,\ \varepsilon_{xt1} = \eta_{xt1} - \hat{\eta}_{xt1};$$

and, as $t = t_1 + 1, t_1 + 2, \ldots$:

$$\eta_{xt} = \hat{\eta}_{xt} + \varepsilon_{xt},$$

$$\hat{\eta}_{xt} = \hat{\eta}_{xt-1} + \hat{b}_{xt-1},$$

$$\varepsilon_{xt} = \rho\varepsilon_{xt-1},$$

$$\hat{b}_{xt} = \hat{b}_x^* + (1-\pi)\left(\hat{b}_{xt-1} - \hat{b}_x^*\right). \quad (20)$$

3. Empirical illustrations

We illustrate performance of the Direct Extrapolation (DE) model (20) by presenting projection results obtained both using the method itself and also by the Lee-Carter (LC) model. (We consider only the central scenario implied by the LC model.) The latter model is chosen for comparison, because of its wide usage for the mortality projection and also because of similarity of its results to those obtained by separate linear

extrapolation[9] (e.g., McNown 1992, Ediev 2008) of age-specific log-death rates without additional adjustments.

One of the main adjustments used in (20) is aimed to prevent implausible developments of the pattern of age-sex death rates. In order to be able to examine, weather and how often such implausible patterns appear in the projection, we present, firstly, results for long-range projections. Note, however, that dynamics leading eventually to implausible results may also produce less efficient projection in the short and medium run.

Firstly, we present results for would-be projections from 1900 until the contemporary period based on data from 19th century for three countries: Austria (Ediev and Gisser 2007), England and Wales, and Sweden (Human Mortality Database[10]), see figures 3-5. All three figures are constructed in a similar way. They contain four charts for projected and actual age-specific central death rates (per 1000, logarithmic scale) – two for females and two for males; and one chart representing actual and projected dynamics of life expectancies at birth for males as well as for females. Each chart containing graphs of death rates consists of four curves: of death rates estimated from the original data for the base year (1900) and for the latest year with data available (2005 for Austria and Sweden; 2003 for England and Wales) and curves of death rates projected to the last year by the DE method (20) and the LC method. Charts in the first row represent results obtained by using the entire data set available for 19th century, i.e., we do not use estimates (5) in this case. Being based on similar long-period data, both methods provide similar results, although, more detailed analysis reveals that the DE method produced lower infant mortality than the LC method. In both cases, the methods failed to predict reductions in mortality, which continued through the 20th century, although, in cases of England and Wales and of Sweden, substantial reductions in mortality at young ages were, indeed, projected. Overall accuracy of both methods improves when parameters are estimated based on the optimal starting years of the data period (5), as it is indicated by charts presented in the second row of each graph. (Since the LC method must be based on data with the same length at each age-sex group, we used the arithmetic average of age-sex-specific start years (5) as the start year for this method.) Despite overall similarity in projections produced by both methods, two dissimilarities are notable. Firstly, the LC method seems to be more likely to produce implausible profiles of mortality – concerning both the age structure of death rates and the sex differences in mortality. Secondly, even being based on data period of optimal length, the LC method produced much more moderately declining child mortality, which, in fact, was due to the monotonicity assumption (14), which was used in DE method and was not applied to the LC method. As a result, the LC method has produced dynamics of life expectancy at birth with slower growth and—in cases of Austria and Sweden—implausible crossovers, see the last charts on the figures.

The next projection exercise is focused on projections to the 21st century produced by both methods based on data from 20th century. For this exercise we have more countries with data available from the Human Mortality Database. Figures 6 to 13 contain graphs of projected age-sex-specific death rates and of dynamics of life expectancies at birth and at age 65 for Austria, England and Wales, Sweden, US, France, Japan, Italy, and Norway. (All calculations are based on the most recent optimal data periods, i.e., (5) applies to the DE method and arithmetic average of estimates (5) – to the LC method.) Since mortality decline in the late 20th century was even more dynamic than it was in late 19th century,

[9] Note, that we use a simplest version of the LC method, without additionally adjusting its parameters to the dynamics of life expectancy at birth or of the number of deaths. That is because similar adjustments could be possible to apply to the direct extrapolation method as well.

[10] Data used in the paper—unless otherwise stated—are taken from the *Human Mortality Database* sponsored by University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany), www.mortality.org or www.humanmortality.de.

problems associated with implausibly forecasted mortality dynamics in future are stronger. The LC method, which is applied here without any consistency adjustments, often produces implausible mortality profiles in the long-run. The cases of England and Wales, of United States, of Japan, and of Italy are especially notable in this respect. One may also note that the both methods produce extremely low mortality rates at young ages by the end of the 21st century (0.001 per 1000 and less). On the one hand, such a low levels may be considered unrealistic, and some age-sex-specific low limits to mortality may be applied, if there is enough biological or other evidence in support of such limits. On the other hand, such limits may also reflect our subjective biases due to particular experience of observing mortality in 20th century only. Also note, that aforementioned projections based on the 19th-century data also illustrate that the contemporary level of mortality would have been considered unrealistically low from the point of view of experience of the 19th century. For these reasons, we did not apply any lower-bound limits to projected mortality.

Last projection exercise is aimed at studying forecasting errors in different time horizons based on empirical data. Indeed, empirical studies of forecast errors must be taken with a deal of precaution, as we do not have a random sample of independent data series, neither each of the data series available represents a stationary process. Hence, results based on past forecast errors may only tentatively indicate possible performance of the models in the future. Nonetheless, such results do provide some additional insights. Table 1 contains some of the results obtained using different versions of the LC and DE models applied to past data (see descriptions below). For each modification of the methods, we present absolute percentage errors for death rates and absolute errors for life expectancies at 0 and 60 averaged over all possible base years of the projection, over both sexes, and (for mortality rates) over all age groups. Errors are presented for four different forecasting horizons: one, five, ten, and twenty years.

For the LC method, results are presented for three modifications: first two, based on extrapolating the k(t) function starting from the jump-off value (this would correspond to central path of the random walk with drift procedure and is reflected in 'rwd' notation) with single function k(t) estimated for both sexes ('uk') or with two sex-specific functions ('sk'), and the third modification based on ignoring the jump-off value for k(t) and following its regression line ('regr') with single k(t) function applied to both sexes. All three modifications are based on optimal data periods as explained above. (Results based on entire data sets were of good quality only in case of countries with short data periods.)

For the DE method, modifications to the main version (20) differ only with respect to the speed of convergence to the plausible rates of mortality decline: in basic version, it is kept as described above, see (17)-(19); an alternative version implies no convergence at all, i.e., it may, like the LC method, eventually produce implausible profiles of mortality; and the third version of the DE method implies instant convergence to long-run rates of mortality decline, i.e., rates obtained from (13)-(14) are used from the very first year of forecast.

Overall, the versions of the DE method outperform the versions of the LC method. (To facilitate reading the results presented in table 1, minimal values in each of the rows of the table are marked in bold.) For large populations with considerable autocorrelations between age-specific dynamics of death rates (US and, especially, Japan, for which country a considerable autocorrelations supportive to the LC framework were reported elsewhere, Ediev 2008), the LC method outperforms the DE method. This result, however, seems to reflect peculiarities of the post-war period, when mortality decline was accelerating[11]. Hence, the random walk with drift procedure, which assumes adjustment of

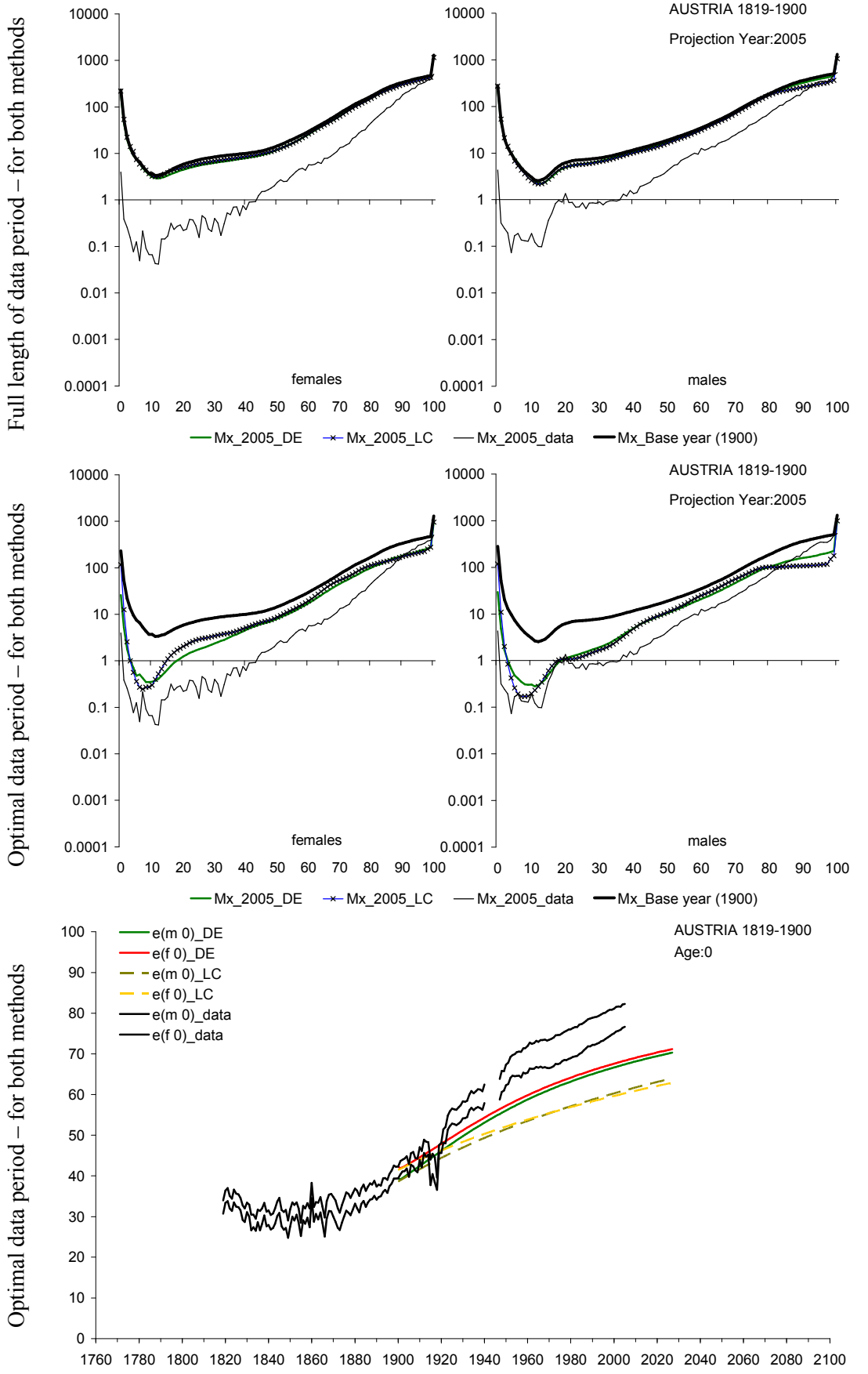
[11] Also note, that long-run forecast errors estimated for such a short period may provide only a biased picture of models' performance.

the future trend for the jump-off value of the last year of observation, gets closer to the (non-linear) trend. Such conclusion is supported, firstly, by better performance of the DE method compared to the LC method, when k(t) follows the regression line without adjustment for the jump-off value. Secondly, this conclusion is also supported by better performance of the DE-analog to the random walk with drift, when schedule of age-sex mortality in the base year is taken from observed jump-off values and not from intercepts of the regression lines (in short and medium-run, the method outperforms the LC method; results not shown in the table).

Next, it is notable from results presented in table 1 that DE method with convergence to plausible pattern of age-sex-specific rates of mortality decline performs better than the method without such convergence. Even more, in some cases, a version of the method with the monotonic profile of rates of mortality decline applied from the very beginning of the projection outperforms the main version based on (17)-(19). First of all, this is another argument in support of the assumption that long-run rates of mortality decline form a monotonic function, as it is proposed in the paper. Secondly, this may indicate that our procedure (13)-(14) aimed at estimating the long-run plausible schedule of rates of mortality decline provides, in fact, a more efficient estimate even for a short-run future. Seemingly, this might be a possibility in case of smaller-size populations, when high volatility of observed rates reduces efficiency of parameters' estimates. This question deserves more future research and accumulation of experience on model's performance.

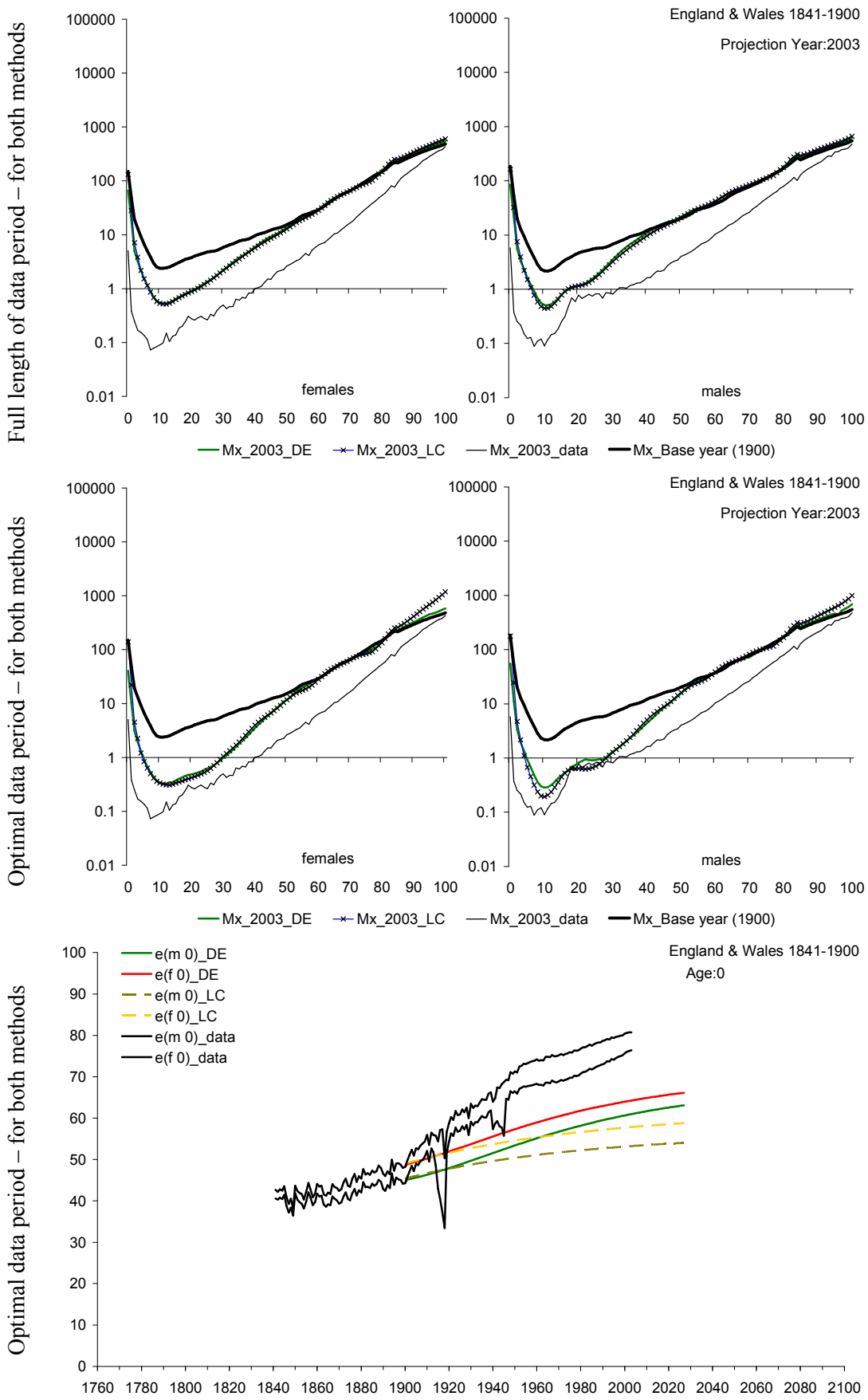
Concluding this section, it is notable that both the long-run projections and study of projection errors in the short- and medium-run support the main assumptions put into the DE model: using optimal duration of the data period and assumption about monotonic patterns of age-sex-specific rates of mortality decline do result in more efficient and plausible projections.

Figure 3. Projected age-sex specific death rates (per 1000, logarithmic scale) and life expectancy at birth since 1900 based on data prior to 1900. Austria.



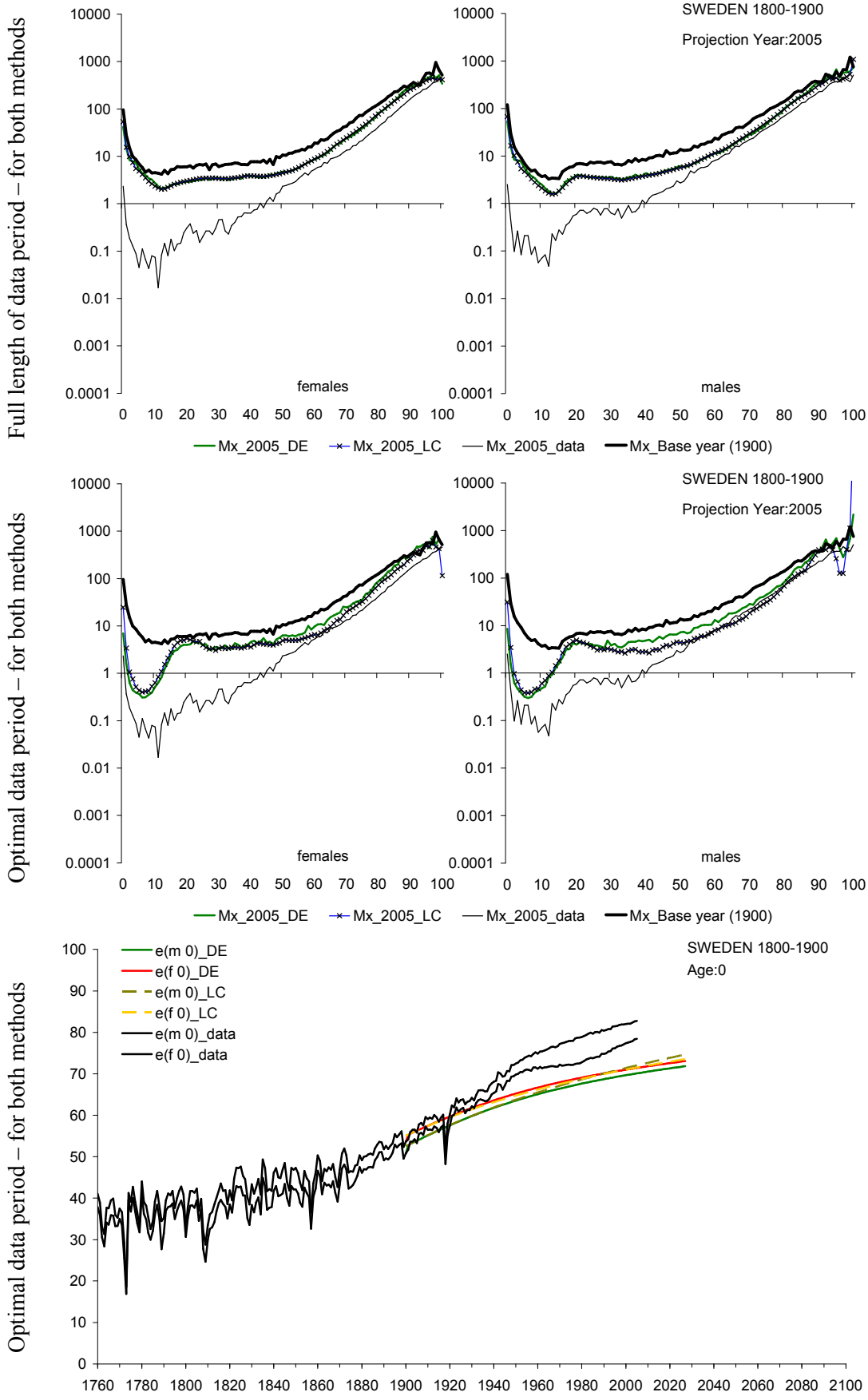
'DE' stands for the 'Direct Extrapolation' method (20); 'LC' stands for the 'Lee-Carter' method.

Figure 4. Projected age-sex specific death rates (per 1000, logarithmic scale) and life expectancy at birth since 1900 based on data prior to 1900. England and Wales.



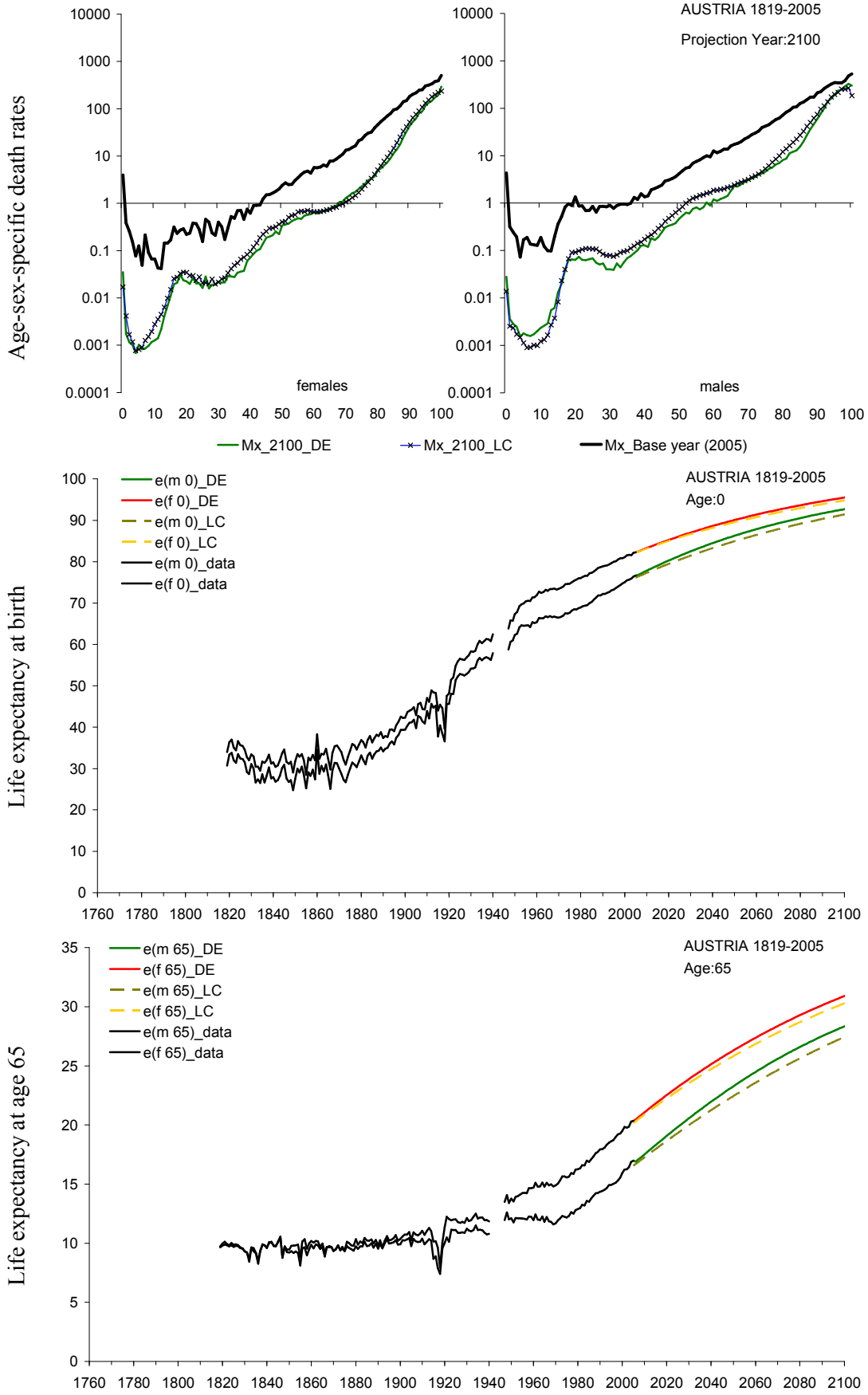
'DE' stands for the 'Direct Extrapolation' method (20); 'LC' stands for the 'Lee-Carter' method.

Figure 5. Projected age-sex specific death rates (per 1000, logarithmic scale) and life expectancy at birth since 1900 based on data prior to 1900. Sweden.



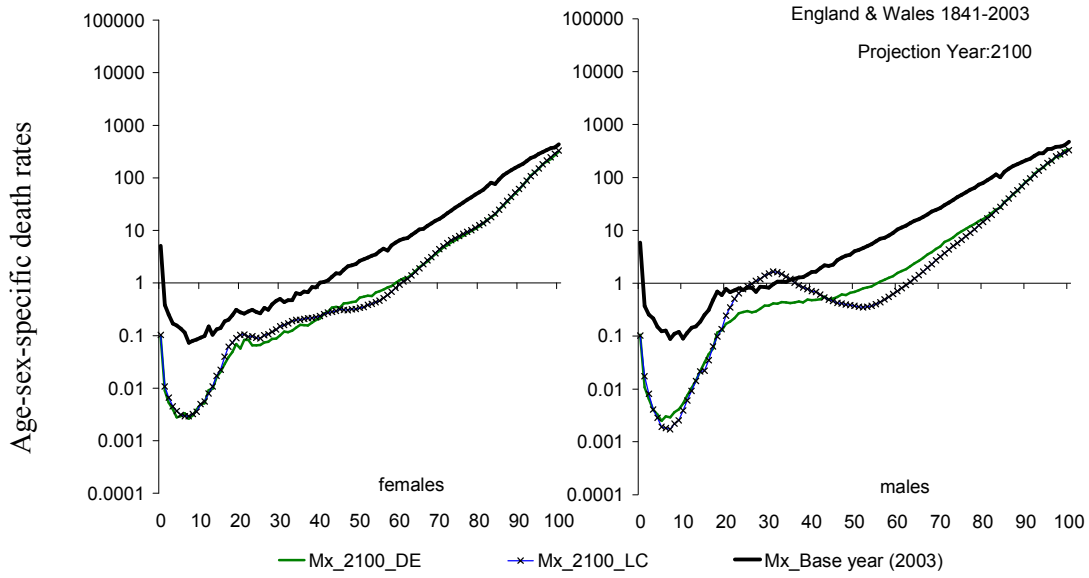
'DE' stands for the 'Direct Extrapolation' method (20); 'LC' stands for the 'Lee-Carter' method.

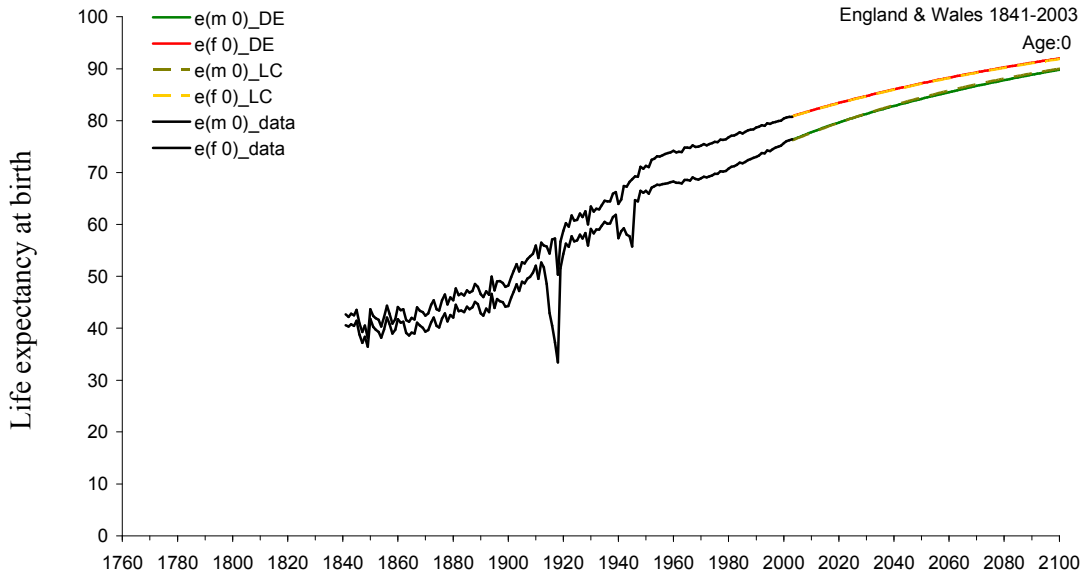
Figure 6. Projected age-sex specific death rates (per 1000, logarithmic scale) and life expectancies at ages 0 and 65 based on the most recent optimal data periods. Austria.

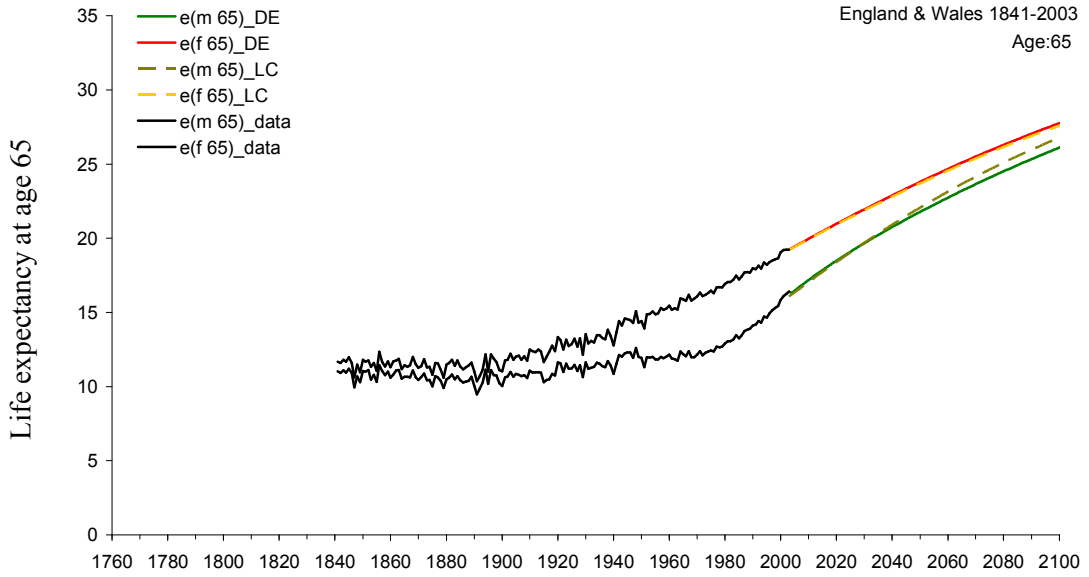


'DE' stands for the 'Direct Extrapolation' method (20); 'LC' stands for the 'Lee-Carter' method.

Figure 7. Projected age-sex-specific death rates (per 1000) and life expectancies at ages 0 and 65 based on the most recent optimal data periods. England and Wales.

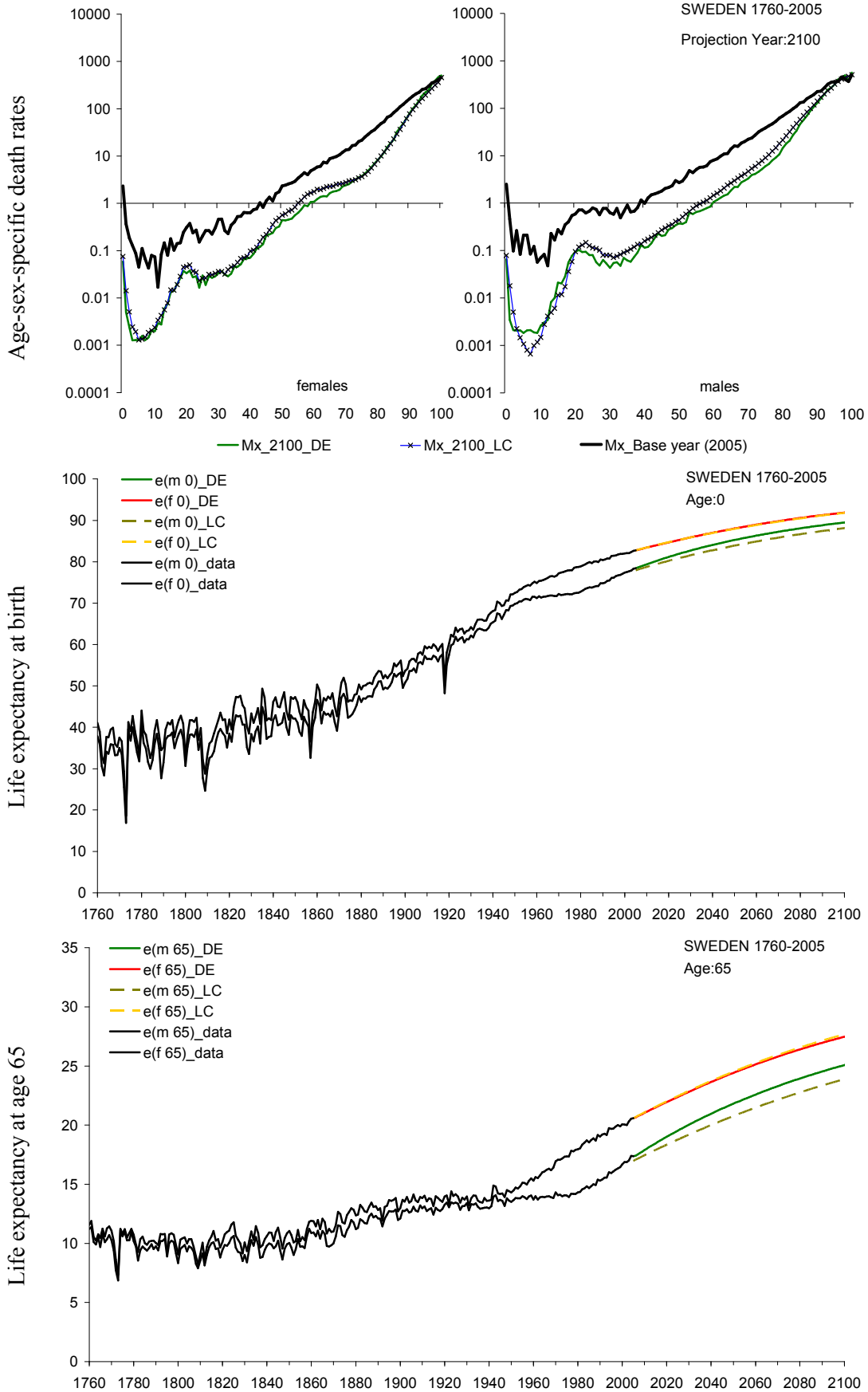






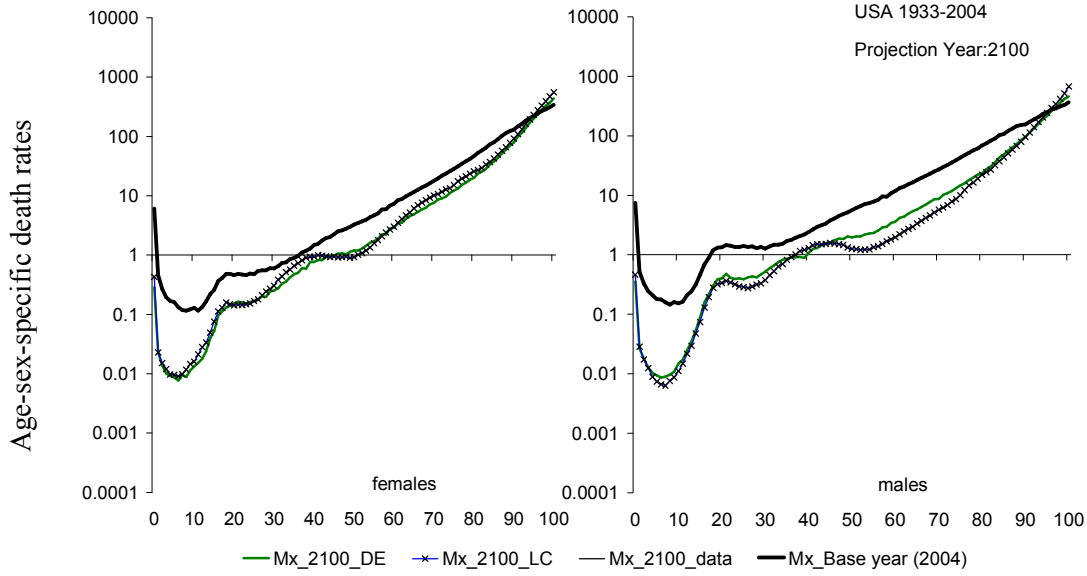
'DE' stands for the 'Direct Extrapolation' method (20); 'LC' stands for the 'Lee-Carter' method.

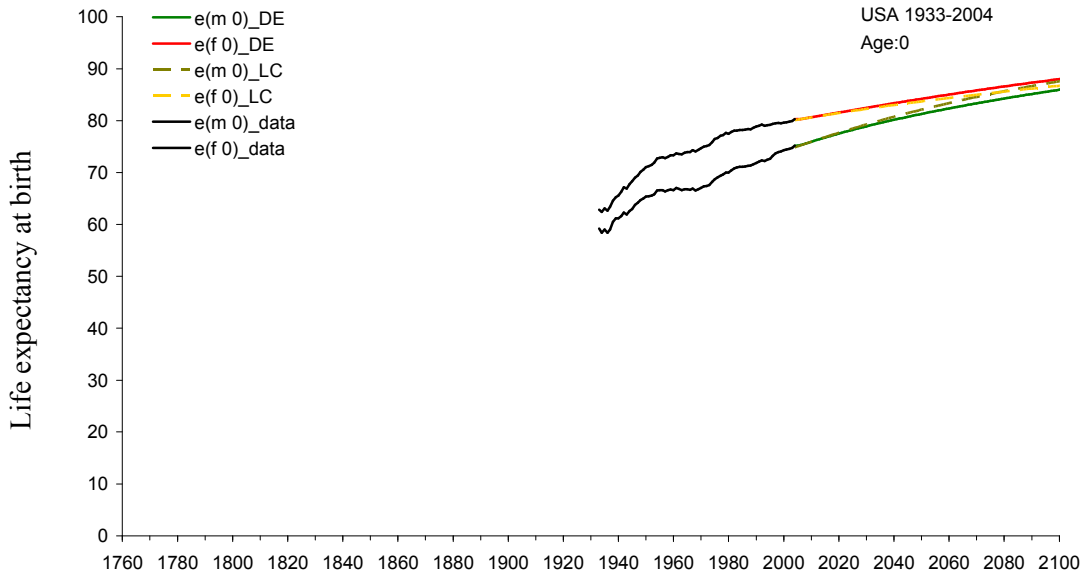
Figure 8. Projected age-sex specific death rates (per 1000, logarithmic scale) and life expectancies at ages 0 and 65 based on the most recent optimal data periods. Sweden.

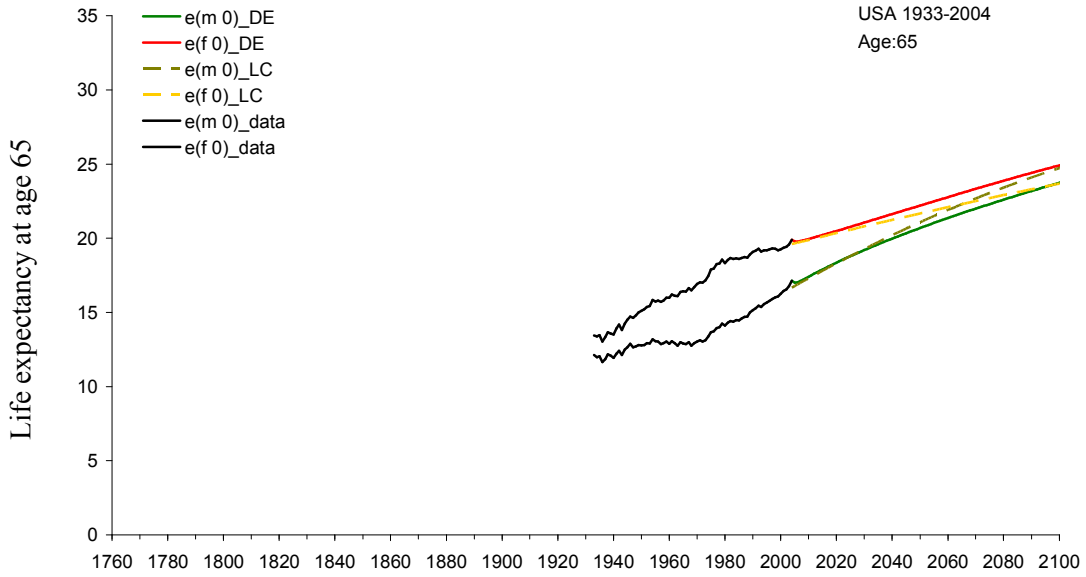


'DE' stands for the 'Direct Extrapolation' method (20); 'LC' stands for the 'Lee-Carter' method.

Figure 9. Projected age-sex specific death rates (per 1000, logarithmic scale) and life expectancies at ages 0 and 65 based on the most recent optimal data periods. US.

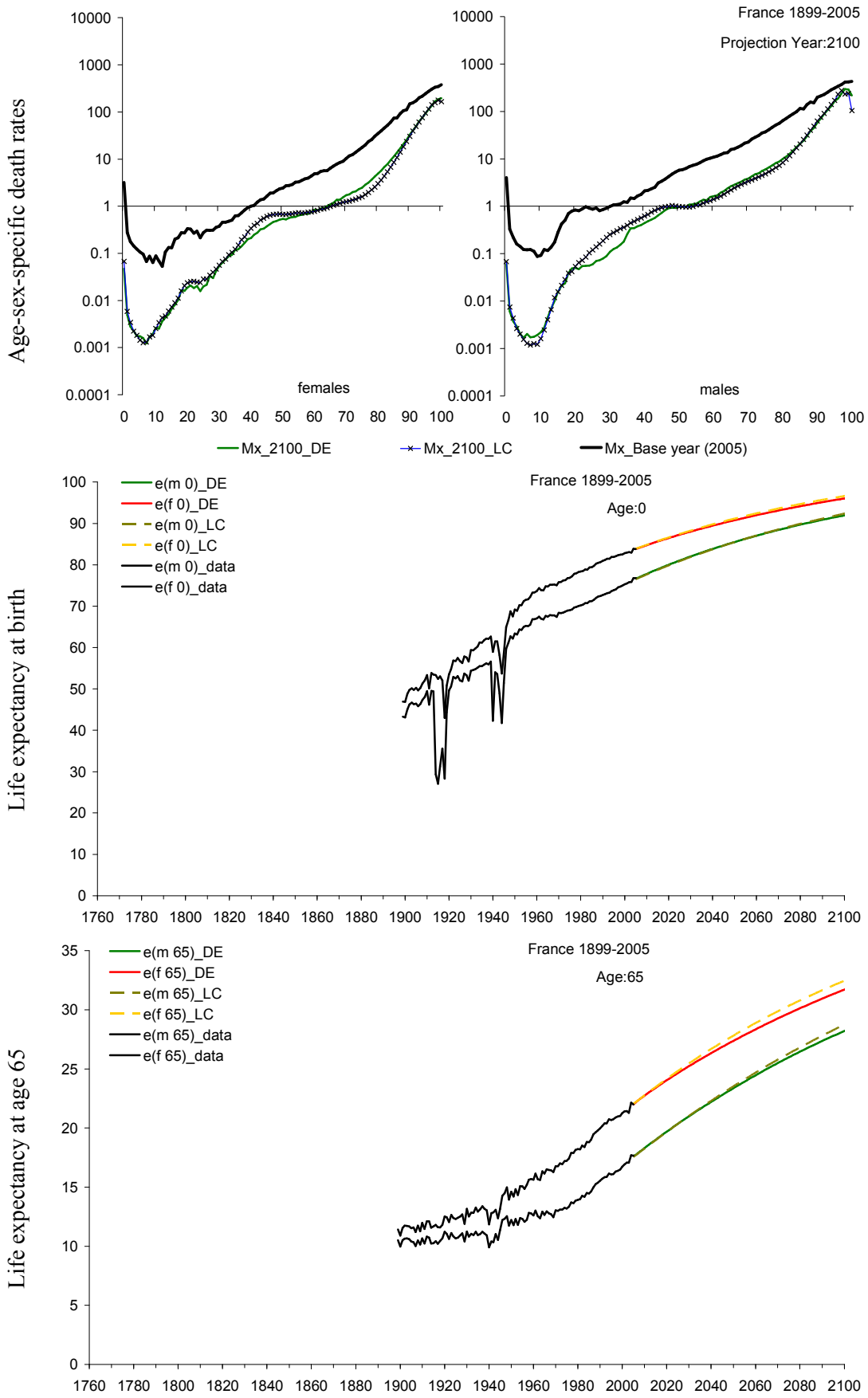






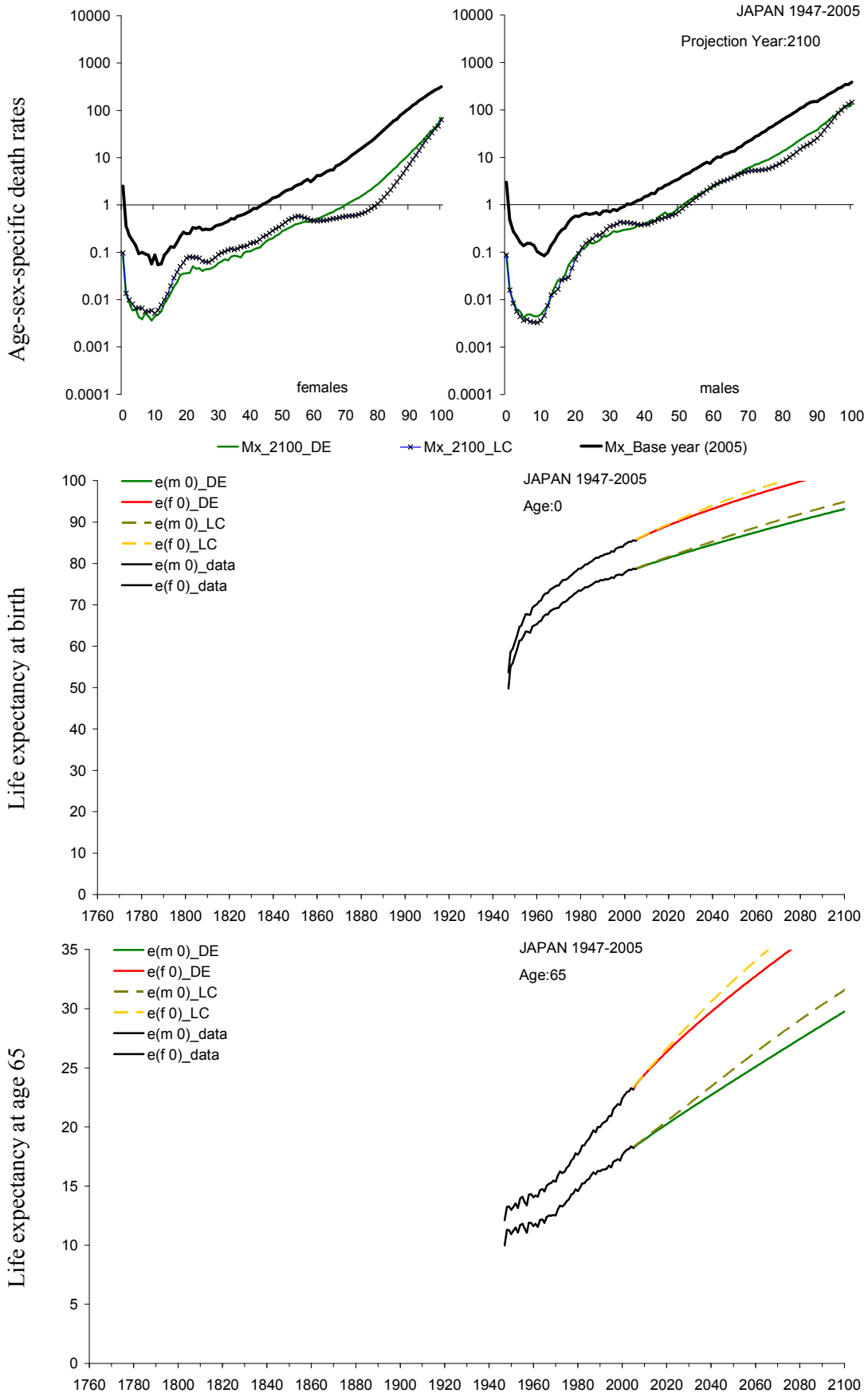
'DE' stands for the 'Direct Extrapolation' method (20); 'LC' stands for the 'Lee-Carter' method.

Figure 10. Projected age-sex specific death rates (per 1000, logarithmic scale) and life expectancies at ages 0 and 65 based on the most recent optimal data periods. France.



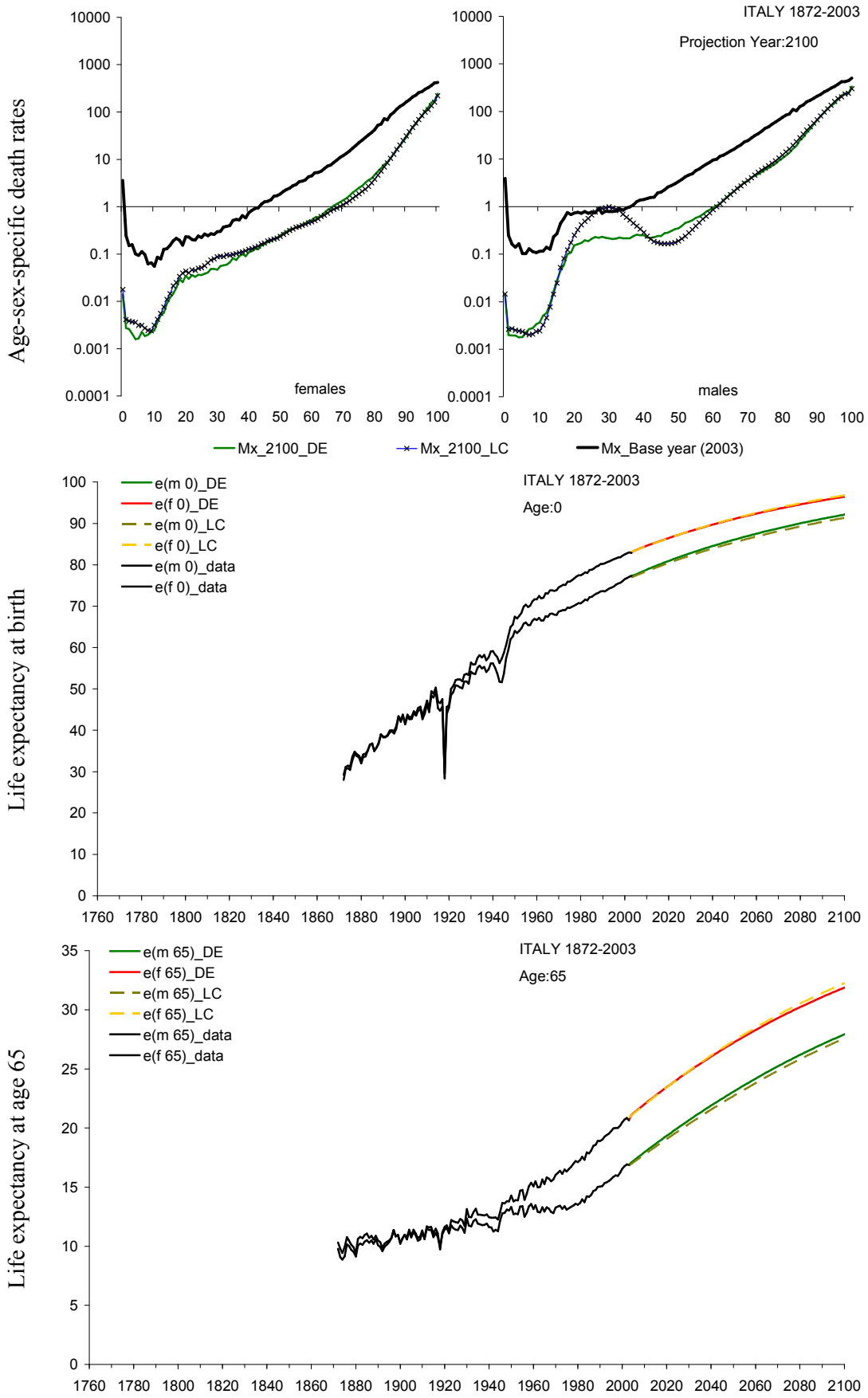
'DE' stands for the 'Direct Extrapolation' method (20); 'LC' stands for the 'Lee-Carter' method.

Figure 11. Projected age-sex specific death rates (per 1000, logarithmic scale) and life expectancies at ages 0 and 65 based on the most recent optimal data periods. Japan.



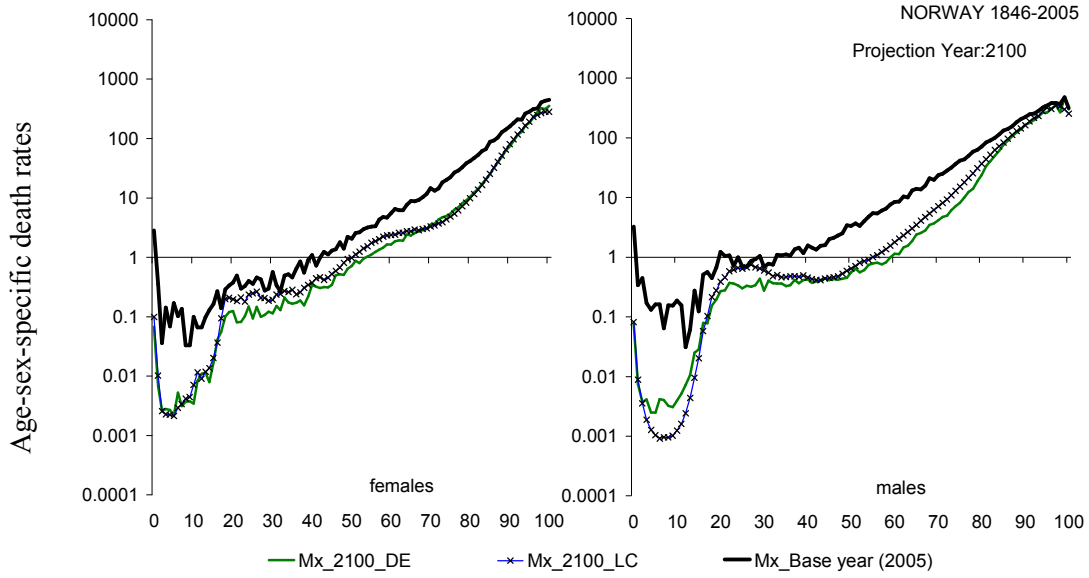
'DE' stands for the 'Direct Extrapolation' method (20); 'LC' stands for the 'Lee-Carter' method.

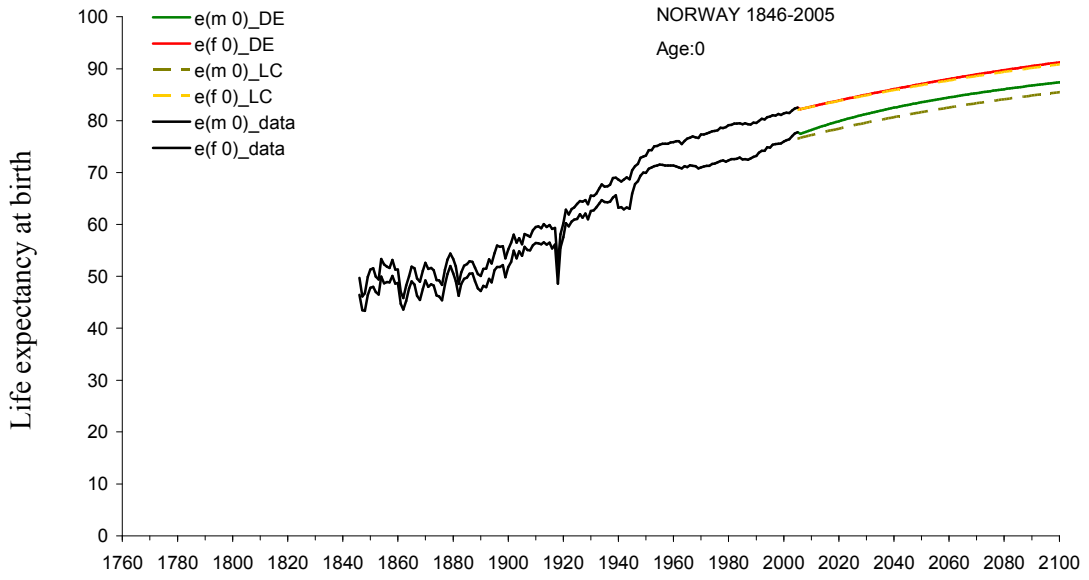
Figure 12. Projected age-sex specific death rates (per 1000, logarithmic scale) and life expectancies at ages 0 and 65 based on the most recent optimal data periods. Italy.

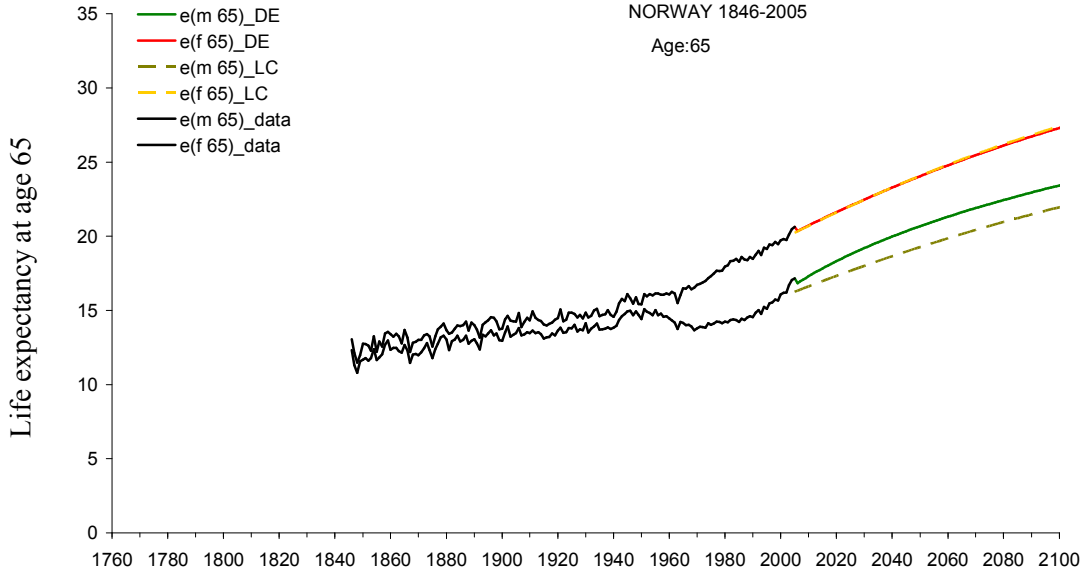


'DE' stands for the 'Direct Extrapolation' method (20); 'LC' stands for the 'Lee-Carter' method.

Figure 13. Projected age-sex specific death rates (per 1000, logarithmic scale) and life expectancies at ages 0 and 65 based on the most recent optimal data periods. Norway.







'DE' stands for the 'Direct Extrapolation' method (20); 'LC' stands for the 'Lee-Carter' method.

Table 1. Mean absolute percentage errors (MAPE) of death rates and mean absolute errors (MAE) of life expectancies at ages 0 and 60 averaged for males and females for all possible starting years of the forecast within the indicated data periods for different forecasting methods[†].

			DE (20) (Direct Extrapolation)			DE: no convergence to $b^*(x)$			DE: instant convergence to $b^*(x)$			LC: rwd_uk_od			LC: rwd_sk_od			LC: regr_uk_od		
country	**period**	**Lag, years**	**MAPE_Mx**	**MAE_e0**	**MAE_e60**	**MAPE_Mx**	**MAE_e0**	**MAE_e60**	**MAPE_Mx**	**MAE_e0**	**MAE_e60**	**MAPE_Mx**	**MAE_e0**	**MAE_e60**	**MAPE_Mx**	**MAE_e0**	**MAE_e60**	**MAPE_Mx**	**MAE_e0**	**MAE_e60**
Sweden	**1760-2005**	1	**13%**	**1.5**	**0.4**	**13%**	**1.5**	**0.4**	**13%**	**1.5**	**0.4**	16%	1.9	0.5	16%	1.8	0.5	14%	1.7	0.5
Sweden	**1760-2005**	5	17%	2.0	**0.5**	**17%**	**2.0**	**0.5**	**16%**	**1.9**	**0.5**	19%	2.3	0.6	19%	2.3	0.6	17%	2.1	0.6
Sweden	**1760-2005**	10	21%	2.2	**0.6**	21%	2.2	**0.6**	**20%**	**2.1**	**0.6**	22%	2.8	0.7	23%	2.7	0.7	20%	2.4	0.7
Sweden	**1760-2005**	20	32%	3.2	1.0	33%	3.3	1.0	30%	**3.0**	**0.9**	32%	3.9	1.0	33%	3.9	1.0	**29%**	3.4	1.1
Sweden	**1760-1913**	1	**13%**	**2.0**	**0.5**	**13%**	**2.0**	**0.5**	14%	2.3	0.6	15%	2.4	0.6	15%	2.4	0.6	15%	2.3	0.6
Sweden	**1760-1913**	5	**17%**	**2.6**	**0.7**	**17%**	**2.6**	**0.7**	**17%**	**2.6**	**0.7**	18%	2.9	0.8	18%	2.9	0.8	**17%**	2.7	**0.7**
Sweden	**1760-1913**	10	20%	**2.7**	**0.7**	20%	**2.7**	**0.7**	20%	**2.7**	**0.7**	20%	3.3	0.8	20%	3.3	0.8	**19%**	3.0	0.8
Sweden	**1760-1913**	20	28%	**3.7**	**1.0**	30%	3.8	**1.0**	28%	**3.7**	**1.0**	26%	4.4	1.1	26%	4.4	1.1	**24%**	4.0	1.2
England & Wales	**1841-1913**	1	**6%**	1.1	**0.3**	**6%**	1.1	**0.3**	**6%**	1.1	**0.3**	**6%**	**1.0**	0.4	**6%**	**1.0**	0.4	**6%**	1.1	0.4
England & Wales	**1841-1913**	5	**8%**	1.5	**0.4**	**8%**	1.5	**0.4**	**8%**	**1.3**	**0.4**	9%	1.5	**0.4**	9%	1.5	0.4	**8%**	1.4	**0.4**
England & Wales	**1841-1913**	10	**10%**	1.9	**0.6**	10%	2.0	**0.6**	**10%**	**1.6**	**0.6**	12%	2.3	**0.6**	12%	2.3	**0.6**	11%	1.8	**0.6**
England & Wales	**1841-1913**	20	**17%**	2.0	**0.8**	**17%**	2.7	0.9	**17%**	**1.1**	**0.8**	19%	3.2	**0.8**	19%	3.2	**0.8**	18%	2.3	0.9
England & Wales	**1946-2003**	1	**7%**	**0.3**	**0.2**	**7%**	**0.3**	**0.2**	**7%**	**0.3**	**0.2**	8%	0.4	**0.2**	8%	0.4	**0.2**	**7%**	**0.3**	**0.2**
England & Wales	**1946-2003**	5	**10%**	**0.5**	0.4	**10%**	**0.5**	**0.3**	**10%**	0.6	0.4	11%	0.7	0.4	11%	0.7	0.5	**10%**	0.6	0.4
England & Wales	**1946-2003**	10	**14%**	**0.9**	**0.6**	**14%**	**0.9**	**0.6**	**14%**	1.0	0.7	15%	1.1	0.7	15%	1.1	0.7	15%	1.0	0.7
England & Wales	**1946-2003**	20	26%	1.9	**1.4**	27%	1.9	**1.4**	**24%**	**1.9**	**1.4**	27%	2.0	1.5	27%	2.0	1.5	28%	1.9	1.5
USA	**1933-2004**	1	**5%**	**0.4**	**0.2**	**5%**	**0.4**	**0.2**	**5%**	**0.4**	**0.2**	**5%**	0.5	**0.2**	**5%**	0.5	0.3	6%	0.5	0.3
USA	**1933-2004**	5	11%	**0.9**	**0.4**	11%	**0.9**	**0.4**	**10%**	1.0	0.5	**10%**	**0.9**	**0.4**	**10%**	**0.9**	**0.4**	11%	0.9	0.5
USA	**1933-2004**	10	16%	1.4	0.7	16%	1.4	0.7	**15%**	1.5	0.7	**15%**	**1.3**	**0.6**	**15%**	**1.3**	**0.6**	**15%**	**1.3**	0.7
USA	**1933-2004**	20	**23%**	1.9	1.1	24%	1.8	1.1	**23%**	1.9	1.1	**23%**	**1.7**	**0.9**	23%	1.8	1.0	**23%**	**1.7**	1.1
JAPAN	**1947-2004**	1	7%	**0.2**	**0.2**	**7%**	**0.2**	**0.2**	**6%**	**0.2**	**0.2**	**7%**	0.4	**0.2**	**7%**	0.4	**0.2**	**7%**	0.3	0.3
JAPAN	**1947-2004**	5	11%	**0.4**	0.5	11%	**0.4**	0.5	11%	**0.4**	0.5	**10%**	0.5	**0.3**	**10%**	0.5	**0.3**	12%	**0.4**	0.5
JAPAN	**1947-2004**	10	18%	0.7	0.9	18%	0.7	0.9	18%	0.7	0.8	**16%**	**0.6**	**0.6**	**16%**	**0.6**	**0.6**	19%	0.7	0.9
JAPAN	**1947-2004**	20	33%	1.0	1.6	34%	1.1	1.6	33%	1.0	1.6	**30%**	0.8	1.1	**30%**	**0.7**	**1.0**	34%	1.1	1.7

[†] For DE method convergence type applies to convergence of age-sex specific rates of mortality decline to long-run plausible schedule of those rates; for LC method, 'rwd' stands for central scenario of the random walk with drift model of k(t), which implies starting the projection from the jump-off death rates, 'regr' implies following regression line for k(t), 'uk' and 'sk' correspond, respectively, to one common and to two sex-specific functions k(t) used in estimation and projection; 'od' stands for 'optimal data' period used to estimate model parameters.

4. Towards stochastic model: preliminary structure

The procedures presented above are developed in the deterministic framework in order to obtain the central projection scenario. Projections, however, are usually not limited to producing the main scenario; they should also provide measures of uncertainty of the future as it is seen from the contemporary experience and knowledge. Traditionally, this is also done in deterministic way, by producing alternative projection scenarios—high, low, and may be others—in addition to the central one. More recently, a growing attention has also been attracted to producing probabilistic projections, which present projection results in terms of distributions of the projected variables, i.e., attempt to assign explicitly or implicitly a probability to each of the possible trajectories of the future instead of pointing to the sensitivity of the projection by drawing several alternative scenarios.

Among other advantages, probabilistic projections provide internally more consistent way of describing the uncertainty at different levels of aggregation. Projected probabilistic distributions of individual age-sex-specific death rates, for example, are consistent with distribution of aggregate characteristics, such as life expectancy at birth. Unlike this, deterministic high and low scenarios may not consistently reflect uncertainty at different levels of aggregation, because they neglect independent variation of mortality at individual age-sex groups.

Partially, the same problem exists in some methods of probabilistic mortality projection available from the literature. In methods based on a parametric mortality model, for example, forecasted dynamics at individual age-sex groups would be derived from (stochastic) dynamics of model parameters, which, in turn, may be derived from dynamics of an aggregate mortality index (e.g., McNown and Rogers 1989; Lee and Carter 1992; Lutz et al. 1996, 1997, 1999; Keilman et al. 2002). In these models, information contained in the dynamic mortality schedules is compressed into a relatively concise set of parameters, which limits independent volatility of mortality rates at the level of individual age groups. In other models, more realistic representation of volatility of age-specific rates was obtained via increased number of parameters and more sophisticated model structure (e.g., Cairns et al. 2007; Hyndman and Ullah 2007; Hyndman and Booth 2008).

Here, we develop a different approach based on the model structure presented above. We propose method of probabilistic mortality projection, which allows direct extrapolation—in a probabilistic fashion—of mortality at individual age-sex groups. The main idea is to derive all necessary model parameters from the recently observed data, i.e., the method is aimed to be pure extrapolative – both in the central trend, and in the projected uncertainty. Namely, we propose the following stochastic variant of the deterministic model (20):

$$\hat{\eta}_{xt1} = \hat{a}_x + \hat{b}_x \cdot t_1,\ \hat{b}_{xt1} = \hat{b}_x,\ \varepsilon_{xt1} = \eta_{xt1} - \hat{\eta}_{xt1};$$

and, as $t = t_1 + 1, t_1 + 2, \ldots$:

$$\eta_{xt} = \hat{\eta}_{xt} + \varepsilon_{xt},$$
$$\hat{\eta}_{xt} = \hat{\eta}_{xt-1} + \hat{b}_{xt-1}, \qquad (21)$$

where residuals ε_{xt} follow AR(1) process

$$\varepsilon_{xt} = \rho\varepsilon_{xt-1} + \sqrt{1-\rho^2} \cdot \zeta_{xt}, \qquad (22)$$

ζ_{xt} are independent and normally distributed with zero mean and dispersion (10), and the slopes $\hat{b}_{xt}$ are modeled in a more complicated fashion, as a combination of the estimated underlying plausible schedule $\hat{b}_x^*$ and of estimated deviations $\hat{\delta}_x$ from that schedule:

$$\hat{b}_{xt} = \hat{b}_{xt}^{*} + \hat{\delta}_{xt}, \tag{23}$$

$$\hat{\Delta}_t = \left(\hat{\delta}_{0t}^{M}, \hat{\delta}_{1t}^{M}, ..., \hat{\delta}_{100t}^{M}, \hat{\delta}_{0t}^{F}, \hat{\delta}_{1t}^{F}, ..., \hat{\delta}_{100t}^{F}\right)^{T} = \hat{C}_{\Delta}^{1/2}\xi_t, \tag{24}$$

where $\hat{C}_{\Delta}^{1/2}$ is the square root of the matrix of correlation coefficients, and ξ_t is a vector of 202 independent normal variables with null expected values and dispersions estimated from the most recent schedule of rates of mortality decline:

$$\hat{\sigma}_{\Delta}^{2} = \theta \frac{\hat{\Delta}_{t1}^{T} \hat{C}_{\Delta}^{-1} \hat{\Delta}_{t1}}{2N}, \tag{25}$$

here $\theta \approx 1.03$ is the correction factor derived from simulations (see further down; for individual populations presented in the paper, this factor varies in between 1.02 and 1.04). At each year of projection, the vector ξ_t is generated taking into account assumed probability of trend change (18), which is estimated from the optimal durations of data periods as was discussed above:

$$\xi_{t1} = \hat{C}_{\Delta}^{-1/2} \hat{\Delta}_{t1}, \tag{26}$$

and, as $t = t_1 + 1, t_1 + 2, ...$:

$$\xi_t = \begin{cases} \xi_{t-1}, with \text{ Probab. } (1-\pi), \\ \xi, with \text{ Probab. } \pi. \end{cases} \tag{27}$$

Here ξ is a newly generated vector of normally distributed independent variables with null expected value and dispersion (25). Such model will, on one hand, keep cross-age-sex correlations at modeled level and, at the same time, will imply trend changes with probability (18).

Proper development (including model for $\hat{b}_{xt}^{*}$) and application of this model is beyond the scope of this paper. Following, we use basic structure of this model to conduct simulations in order to study the properties of estimates of the model parameters, focusing on estimates of the plausible rates of mortality decline.

5. Simulations-based study of estimations of model parameters

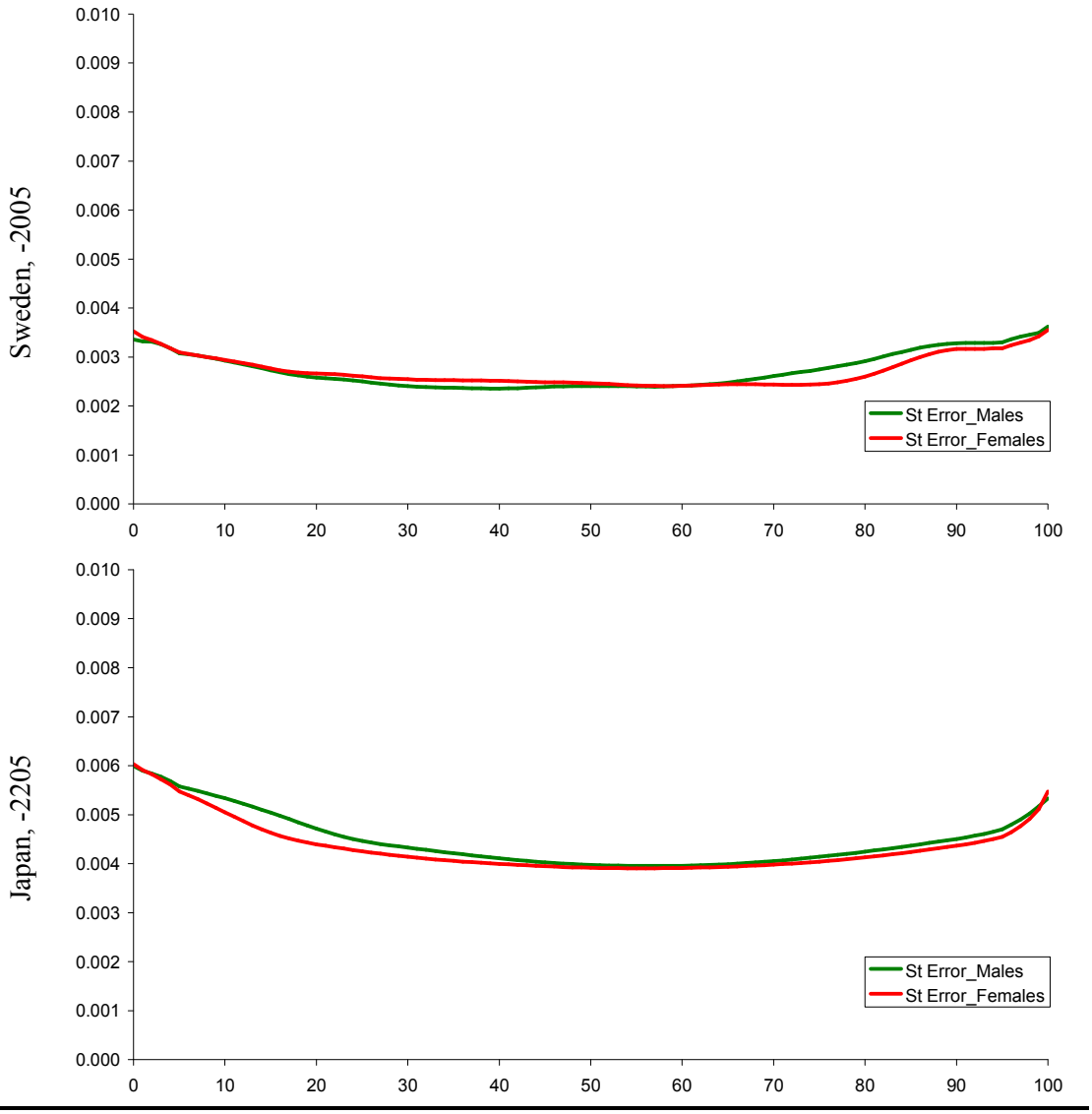
We used the model (23)-(24) above to study the estimation errors of the initially obtained consistent rates of mortality decline $\hat{b}_{xt}^{*}$ in intensive numerical simulations. In the simulation model, we use the estimated profiles $\hat{b}_{x}^{*}$ (for males and females) of consistent rates of mortality decline for a given country and simulate "observed" rates $\hat{b}_{x}$ according (23), i.e., by adding the error term $\hat{\delta}_{x}$, which is generated randomly following (24), where correlation matrix was estimated from the original data. (We assume correlation coefficients being a function of difference in age/sex only.)

For each set of simulated rates $\hat{b}_{x}$ monotonic rates of mortality decline are re-estimated following the same procedures (14), (15) as were used in estimating $\hat{b}_{x}^{*}$ from the original data. (We studied both methods proposed earlier: the simpler Min-Max method proposed in Ediev (2007) and also the presented above more sophisticated method (14) based on minimizing the sum of squares of the residuals (Ediev 2008).) Comparing the rates $\hat{b}_{x}^{*}$ re-estimated from simulated data to the original estimates allows studying both

the biasness and standard errors of the estimation procedures. For each country, data for which are used in simulations, we have conducted 10000 simulation rounds.

Results of simulations show that – despite complex and non-linear structure of the model – estimates $\hat{b}_x^*$ of consistent profile of mortality decline rates obtained by method (14) are exhibiting only minor biases on average. Also, quite unexpectedly, the simulations indicate that – at least, in practical cases analyzed – the simple Min-Max method proposed earlier (Ediev 2007) has the same statistical efficiency as the more elaborate least-squares method (14) proposed later on (Ediev 2008).

Figure 14. Standard estimation errors of consistent rates of mortality decline $\hat{b}_x^*$ obtained from simulations based on parameters derived from the recent mortality dynamics for Sweden and Japan



Simulations allow studying distribution of (unobserved) estimation errors $\Omega_x = b_x^* - \hat{b}_x^*$ of consistent slopes' estimates $\hat{b}_x^*$. Although distribution characteristics of the errors are related to the distribution of δ_x, the relation is not straightforward, as, for example, autocorrelation patterns are distorted due to the monotonicity constraints used to

derive $\hat{b}_x^*$. It is reasonable to expect, however – and it is supported by the simulations' results – that for age groups, which are distant from each other, autocorrelations should not be significantly different from those of δ_x, as the monotonicity conditions affect more severely adjacent ages. At the same time, cross-sex correlations for Ω_x at same age groups are significantly higher compared to the cross-sex correlations of deviations δ_x.

Dispersions of errors Ω_x form a U-shaped function, being highest for the youngest and oldest age classes (see examples in fig. 14; note that estimation errors are lower than the rates of mortality decline by an order of magnitude). For populations studied in the paper, standard errors of $\hat{b}_x^*$ at youngest and oldest ages are approximately 75% of standard deviations of δ_x. The average standard error of $\hat{b}_x^*$ at ages 49-51 was about 14% lower compared to the value for the youngest and oldest age groups. Such a moderate difference may be neglected in deterministic variant scenarios. Hence, in deterministic setup (in order to produce high and low variants), standard errors of $\hat{b}_x^*$ estimates may be taken equal at all age-sex groups (according to the simulations, they might be taken as 70% of the estimated standard deviations of δ_x). Also notable is that the estimation errors do not show considerable differences by sex.

Based on simulations results, we have also considered biasness and efficiency of estimates for the standard deviation of δ_x and of autocorrelation patterns for δ_x. Calculations were conducted for three methods of estimation: one, which neglects autocorrelations between the deviations δ_x:

$$\hat{\sigma}_\delta^{\,2} = \frac{\hat{\Delta}^T \hat{\Delta}}{2N}, \tag{28}$$

and the other two methods assuming autocorrelations:

$$\hat{\sigma}_\delta^{\,2} = \frac{\hat{\Delta}^T \hat{C}_\delta^{\,-1} \hat{\Delta}}{2N}, \tag{29}$$

where $\hat{\Delta} = \left(\hat{\delta}_0^M, \hat{\delta}_1^M, ..., \hat{\delta}_{100}^M, \hat{\delta}_0^F, \hat{\delta}_1^F, ..., \hat{\delta}_{100}^F\right)^T$ - vector combining the estimated age-specific deviations for both males and females; $\hat{C}_\Delta$ is their correlation matrix; and $N = 101$ is the number of age classes. The correlation matrix is assumed to have same structure as estimated from data (assuming that correlations between residuals in two age-sex groups form a function of difference in age/sex), or, alternatively, as following the exponential model:

$$Correl\left(\delta_x, \delta_{x+k}\right) = \rho_*^{\,k},\ k = 0,1,2...,$$
$$Correl\left(\delta_x^{\,Males}, \delta_{x+k}^{\,Females}\right) = \rho_{S*}\rho_*^{\,k},\ k = 0,1,2..., \tag{30}$$

here correlation coefficients ρ_* and ρ_{S*} are estimated from observed (simulated) deviations $\hat{\delta}_x = \hat{b}_x - \hat{b}_x^*$:

$$\hat{\rho}_{*}=\frac{\overline{\left(\hat{\delta}_x-\overline{\hat{\delta}_x}\right)\left(\hat{\delta}_{x+1}-\overline{\hat{\delta}_{x+1}}\right)}}{\sqrt{\overline{\left(\hat{\delta}_x-\overline{\hat{\delta}_x}\right)^2}\cdot\overline{\left(\hat{\delta}_{x+1}-\overline{\hat{\delta}_{x+1}}\right)^2}}},$$

$$\hat{\rho}_{S*}=\frac{\overline{\left(\hat{\delta}_x^{M}-\overline{\hat{\delta}_x^{M}}\right)\left(\hat{\delta}_x^{F}-\overline{\hat{\delta}_x^{F}}\right)}}{\sqrt{\overline{\left(\hat{\delta}_x^{M}-\overline{\hat{\delta}_x^{M}}\right)^2}\cdot\overline{\left(\hat{\delta}_x^{F}-\overline{\hat{\delta}_x^{F}}\right)^2}}}. \qquad (31)$$

Characteristic patterns of simulated distributions of estimated standard deviations of δ_x together with the true value originally put into the simulations (the simulation parameters correspond to the Japanese case) are presented in figure 15. Estimated values show notable deviations from the original values, the bias being strongest for the estimates based on neglecting the autocorrelations. Estimates based on the implied correlation structure or on the exponential model (30) perform considerably better. Based on the simulations' results, we have estimated the correction coefficient in (25) for the estimation bias in (29).

Unfortunately, simulations also indicate that the estimates of the correlation coefficients may be of poor efficiency. In figure 16, we present simulated distributions for the estimates of the autocorrelation coefficients (28). Estimates are clearly biased and of low efficiency.

Summarizing the simulation results for deviations of rates of mortality decline from the plausible schedule and for estimation errors of the plausible rates, one may note necessity to take autocorrelations in deviations into account, albeit for those autocorrelations one may not provide a sound estimates. Nonetheless, also note that even if wrong estimates for the autocorrelations (28) lead to biased estimates for the standard deviations of the residuals δ_x and, consequently for the estimation errors of the underlying plausible and projected rates of mortality decline, the impact of these deficiencies on accuracy of the projection will be limited, as bias in correlations will limit effect of the estimation bias of standard deviations of residuals. Apparently, this problem needs further investigation beyond the current study. One direction of research important in view of the simulation results is to develop some model structures of autocorrelations based on cross-country and long-run studies.

Figure 15. Simulated distributions of estimates for standard deviations of rates of mortality decline $\hat{b}_x$ from the plausible schedule of those rates $\hat{b}_x^*$

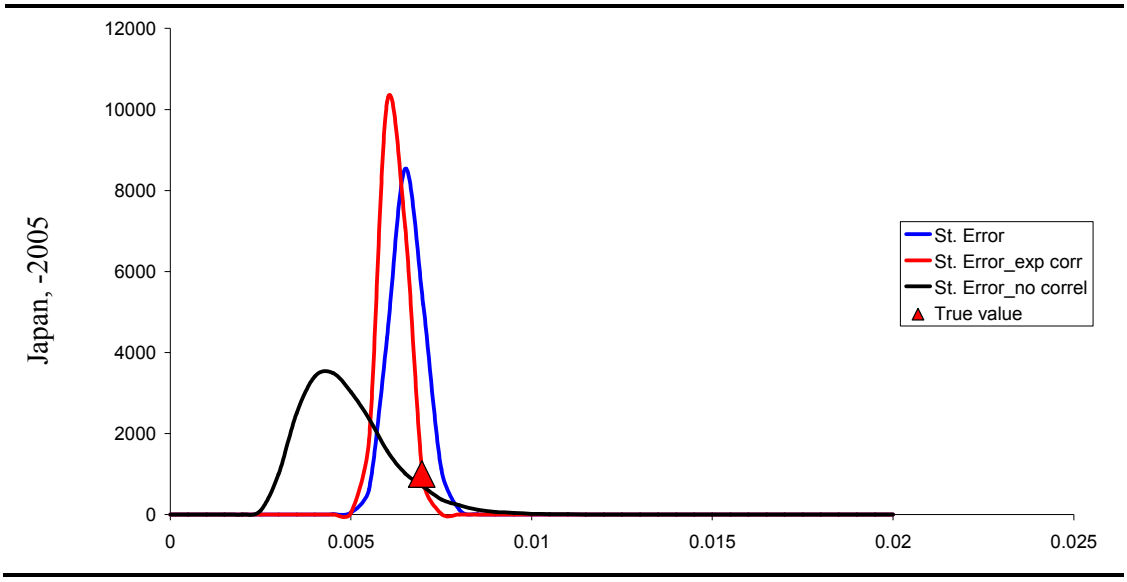
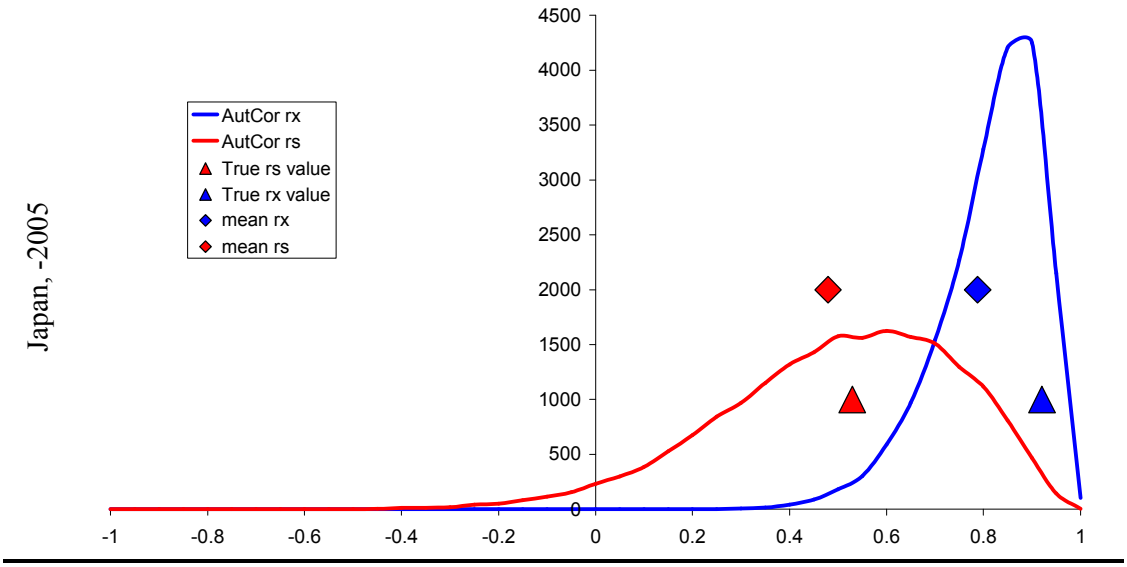


Figure 16. Simulated distributions of estimates for autocorrelation coefficients (28)



6. Concluding remarks. Further improvements and developments of the method

We start our study from reiterating that extrapolative approach to mortality forecasting—though remaining one of the most important practical methods—possesses some drawbacks. In particular, extrapolation of age-sex specific death rates may result in implausible age-sex patterns of the projected mortality.

Therefore, we supplement the direct extrapolation with additional procedures aimed at improving its short-term projecting efficiency and long-term consistency. In particular, we estimate optimal age-sex-specific durations of most recent data periods, when linearity of dynamics of log-mortality rates can be assumed. These most recent linear trends are extrapolated in to the future. To avoid implausible projected mortality patterns in the long-run, however, we also estimate long-run plausible schedules of age-sex-specific rates of mortality decline, to which mortality dynamics is assumed to gradually converge. Key

assumption used to derive these schedules is that long-run rates of mortality decline form a pattern monotonically decreasing with age, rates of decline for females being higher than for males. Such an assumption is derived from the general tendency of mortality declining rates to decrease with age found in empirical observations. It is also supported by theoretical considerations and by improved performance of the projection method shown on past data.

Estimates and forecasting results presented also point to possibility to simplify and improve efficiency of the method: either by assuming similarity of male and female rates of mortality decline in the long run and/or by abandoning the convergence model (17)-(19) for the rates of mortality decline and, instead, applying the monotonic profile (13)-(14) of rates from the very first year of projection. Other assumptions used in the paper (e.g., minimal duration of the data period, criteria used to detect the optimal duration of the data period, etc.) may also be further improved.

The model structure allows introducing the uncertainty into the forecast. This may be done either deterministically, by developing variant scenarios of future dynamics, or probabilistically, by developing a stochastic model for mortality dynamics as outlined above. Both methods may be based on objective measures of uncertainty, which may be derived within the framework proposed in the paper. Namely, standard errors of estimates of parameters of age-sex-specific trends, as well as of long-run schedules of rates of mortality decline together with estimates of autocorrelations between model residuals and between deviations of observed rates of mortality decline form long-run estimates may be used for that purpose. Illustrative projections undertaken in this way show prominence and convenience of this approach. At the same time, results of simulations conducted in order to study properties of estimates of model parameters reveal certain problems. In particular, efficiency of estimates of dynamics of age-sex-specific rates of mortality decline strongly depend on estimates of autocorrelations, which, however, are shown to be strongly biased and inefficient. Although, these problems may have a minor effect of forecast properties, an additional study is to be undertaken on the problem. Another necessary development to the probabilistic model concerns modeling changes in the underlying plausible schedule b_x^*. In its current form, the model assumes this schedule to be fixed, although a more realistic assumption would be to allow for its changes in the long-run. Analysis of data (not shown here) suggests that these changes may efficiently be modeled based on the past variations of rates of mortality decline.

Data and estimates for different populations point to possibility of improving the projection efficiency—especially in the long-run—by developing the method in the multi-regional framework. Prominence of this approach was demonstrated by Sanderson and Scherbov (2004); it was also successfully implemented by Li and Lee (2005) in the context of the LC method. For the direct extrapolation method proposed here, a multi-regional approach may be based on assuming all populations to have similar long-run rates of mortality decline (so, that ever-increasing divergence of populations' mortality schedules will be precluded) and, possibly, on assuming that asymptotically, all the populations should converge to the same trajectory of mortality dynamics (which, in addition to long-run rates of decline will also affect estimates of parameters of convergence to the long-run trend).

Acknowledgment
The work is partially based on results obtained within the project "Analysis, evaluation and adjustment of extensive models for mortality projections" carried out by the Vienna Institute of Demography for Statistik Austria.

References

Bartlett, M. S. 1946. On the theoretical specification of sampling properties of autocorrelated time series. *J. Royal Stat. Soc*. **B8**, 27.

Bongaarts, J. 2005. Long-Range Trends in Adult Mortality: Models and Projection Methods. *Demography*, 42(1): 23-49.

Bongaarts, J. and G. Feeney. 2003. Estimating mean lifetime. *Proceedings of the National Academy of Sciences USA*, 100(23): 13127-13133.

Booth, H. 2006. Demographic Forecasting: 1980 to 2005. *Int. J. of Forecasting*, 3(18): 547-581.

Box, G. E. P., G. M. Jenkins. 1970. *Time series analysis: forecasting and control*. Holden-Day: San-Francisco. (More recent editions are also available.)

Cairns, A.J.G., D. Blake, K. Dowd, G.D. Coughlan, D. Epstein, and M. Khalaf Allah. 2008. Mortality Density Forecasts: An Analysis of Six Stochastic Mortality Models. Discussion paper PI-0801. http://www.ma.hw.ac.uk/~andrewc/papers/ajgc53.pdf

Continuous Mortality Investigation. 2007. *Stochastic projection methodologies: Lee-Carter model features, example results and implications.* Life Office Mortality Committee. Revised Working Paper 25. http://www.actuaries.org.uk/__data/assets/pdf_file/0011/38846/wp25revised.pdf

Day, J. Ch., U.S. Census Bureau. 1996. *Current Population Reports*, P25-1130, U.S. Government Printing Office, Washington, District of Columbia.

De Grey, A.D.N.J. 2006. Extrapolaholics Anonymous: Why Demographers' Rejections of a Huge Rise in Cohort Life Expectancy in This Century are Overconfident. *Ann. N.Y. Acad. Sci.* 1067, 83-93.

De Jong, P. and C. Marshall. 2007. Mortality Projection Based on the Wang Transform. *Astin Bulletin* 37(1): 149-161.

Ediev, D. and R. Gisser. 2007. Reconstruction of the historical series of life tables and of age-sex structures for the Austrian population in 19th-first half of 20th centuries. *Vienna Yearbook of Population Research*: 327-355.

Ediev, D.M. 2007. An approach to improve the consistency of mortality projections obtained by the Lee-Carter method // Eurostat methodologies and working papers. Work session on demographic projections. Bucharest, 10-12 October 2007. Luxembourg: Office for Official Publications of the European Communities. P. 101-115.

Ediev, D.M. 2008. Extrapolative Projections of Mortality: Towards a More Consistent Method. Part I: The Central Scenario. Vienna, Vienna Institute of Demography of Austrian Academy of Sciences. Working paper WP 03/2008. 50 pp. http://www.oeaw.ac.at/vid/download/WP2008_03.pdf

Gallop, A. 2007. Methods used in drawing up mortality projections. Mortality Projections in the United Kingdom. Fifteenth International Conference of Social Security Actuaries and Statisticians. Helsinki, Finland, 23-25 May 2007. Available online: http://www.issa.int/pdf/helsinki07/2gallop.pdf

Gómez de León, C. 1990. *Empirical EDA Models to Fit and Project Time Series of age-specific Mortality Rates*. Discussion Paper. Oslo, Central Bureau of Statistics.

Goss, S.C., A. Wade, F. Bell, and B. Dussault. 1998. Historical and projected mortality for Mexico, Canada, and the United States. North American Actuarial Journal 2(4): 108-126.

Hollmann, F.W., T.J. Mulder, and J.E. Kallan. 2000. Methodology and Assumptions for the Population Projections of the United States: 1999 to 2100. *Population Division Working Paper No.* 38. U.S. Census Bureau, Washington, D.C. Available online: http://www.census.gov/population/www/documentation/twps0038.pdf

Human Mortality Database. Online database sponsored by University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany), www.mortality.org or www.humanmortality.de.

Hyndman, R.J. and H. Booth. 2008. Stochastic population forecasts using functional data models for mortality, fertility and migration. International Journal of Forecasting, 24, 323-342.

Hyndman, R.J. and S. Ullah. 2007. Robust forecasting of mortality and fertility rates: a functional data approach. Computational Statistics & Data Analysis, 51, 4942-4956.

Keilman, N., D.Q. Pham, & A. Hetland. 2002. Why population forecasts should be probabilistic - illustrated by the case of Norway. *Demographic Research* 6(15): 409-454, www.demographic-research.org/Volumes/Vol6/15/.

Keyfitz, N. 1985. *Applied mathematical demography*. N. Y., etc.: Springer.

Lee, R.D. and L.R. Carter, 1992. Modeling and Forecasting U.S. Mortality. *J. of American Statistical Association*, 87 (419): 659-671.

Li, N. and R. Lee, 2005. Coherent Mortality Forecasts for a Group of Populations: an Extension of the Lee-Carter Method. *Demography*, 42 (14): 575-594.

Lutz, W., W. Sanderson, and S. Scherbov. 1996. Probabilistic population projections based on expert opinion. Pages 397-428 in W. Lutz, ed. The Future Population of the World. What Can We Assume Today? London: Earthscan, Revised Edition.

Lutz, W., W. Sanderson, and S. Scherbov. 1997. Doubling of world population unlikely. *Nature* 387, 803-805.

Lutz, W., W. Sanderson, and S. Scherbov. 1999. Expert-based probabilistic population projections. Pages 139-155 in Wolfgang Lutz, James W. Vaupel, and Dennis A. Ahlburg, eds. Frontiers of Population Forecasting. Supplement to Population and Development Review, Vol. 24, 1998. New York: Population Council.

McNown, R. 1992. Modeling and Forecasting U.S. Mortality: Comment. *J. of American Statistical Association*, 87 (419): 671-672.

McNown, R. and A. Rogers. 1989. Forecasting Mortality: A Parameterized Time Series Approach. *Demography*, 76 (4): 645-660.

Oeppen, J. and J.W. Vaupel. 2002. Broken Limits to Life Expectancy. *Science* 10 (May): Vol. 296. no. 5570, pp. 1029 – 1031. DOI: 10.1126/science.1069675.

Olshansky, S.J., Carnes, B.A., Désesquelles, A. 2001. Prospects for Human Longevity. Science 291 (5508):1491-1492.

Sanderson, W. and S. Scherbov. 2004. Putting Oeppen and Vaupel to Work: On the Road to New Stochastic Mortality Forecasts. IIASA Interim Report IR-04-049. 16 pp. http://www.iiasa.ac.at/Admin/PUB/Documents/IR-04-049.pdf

Social Security Administration. 1987. Social Security Area Population Projections, 1987, Actuarial Study No. 99. Washington, D.C.: U.S. Office of the Actuary.

Wilmoth, J.R. 2005. Overview and Discussion of the Social Security Mortality Projections. Working Paper for the 2003 Technical Panel on Assumptions and Methods, Social Security Advisory Board, Washington, D.C.