Coherent mortality forecasting via the aggregate measure of life expectancy

Tiziana Torri*

1 Introduction

Mortality forecasts are a crucial aspect in the framework of population forecasting, affecting governments' interventions in pensions and health-care systems. It is undoubtedly more common to find in the literature mortality forecasts based on disaggregated data, such as age-specific death rates, rather than directly on life expectancy. If on one hand, modeling age-specific death rates provides clear insights into mortality dynamics over age and time, on the other hand, it does not guarantee the coherence of the forecast life expectancy. It is indeed the case, that forecasting mortality rates frequently leads to underestimation of future life expectancy. This is, for example, a well known result of the widely used Lee-Carter model (Lee and Carter, 1992), when results are investigated through the application of the cross-validation method. The model assumes a constant decline of the log-deaths rates, leading to a decelerating increase in life expectancy, as explained in Alho (1989). Such direct outcomes contradict the linear trend found by White (2002) for the series of the sex-combined life expectancies in 21 developed countries.

The previous considerations motivates our decision to forecast mortality through the direct forecast of life expectancy. We will apply extrapolation methods to project the historical mortality trends into the future, based on the assumption that the conditions which led to changing mortality in the past will continue to have a similar impact in the future. Hence, no limitations to the future length of life are imposed.

Early works on life expectancy's forecast are presented by De Beer and Alders (1999), that describe the development of life expectancy at birth in the Netherlands for men and women by a random walk with drift. Similarly, Keilman et al. (2001) use ARIMA models to predict the future mortality of

*Corresponding author: Max Planck Institute for Demographic Research, Konrad-Zuse-Strasse 1, 18057 Rostock, Germany; Tel. +49(0)381 2081 150, Fax. +49(0)381 2081 450; email: torri@demogr.mpg.de

the Norwegian population. A more complex model was constructed by Alders et al. (2007) for long-term stochastic population forecasts in 18 European countries. They estimated a GARCH-type time series models, based on a long series of life expectancy.

In the current work we apply two different models derived from the theory of the time series analysis. First, we apply the classic and already mentioned univariate ARIMA model, based on assumption of stability and stationarity of the process. Later on, we consider the Structural Time Series Models, framed in the more general structure of the state-space models. These models relax the assumption of stationarity, present in the ARIMA models, and assume time-varying parameters.

The next section presents the data we used for the analysis. Section 3 describes the ARIMA model and the Structural Time Series Model fitted on the data. An application of the models to the data is presented in Section 4, followed by a critical discussion of the method in Section 5.

2 Data

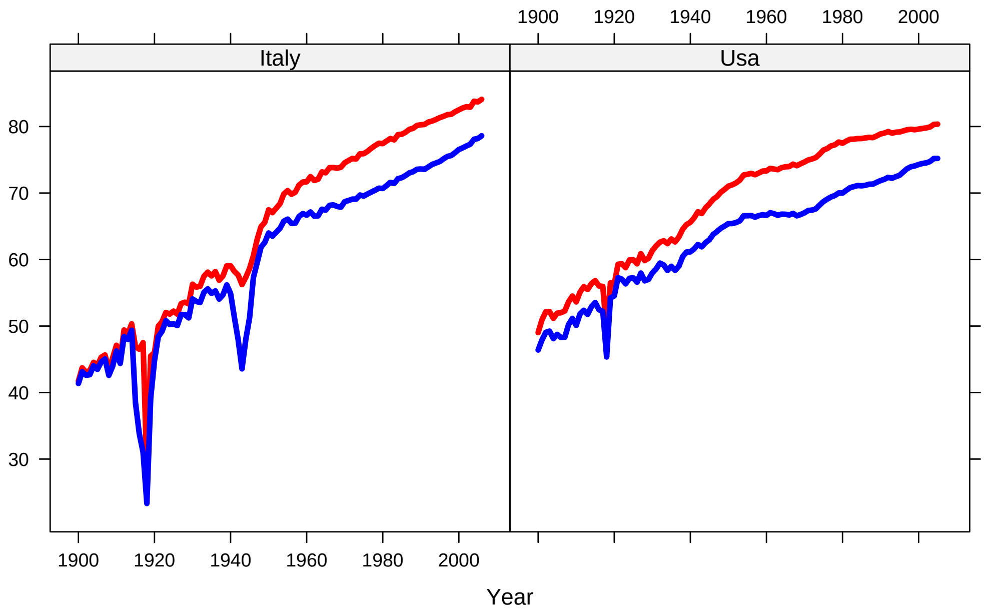
The analysis is conducted on Italian and U.S. data, taken from the Human Mortality Database (2008) (HMD) and the Berkeley Mortality Database (1995) (BMD). With regard to the United States, the HMD includes data only from 1933 to 2005. On the other hand, death rates by single year of age, derived from life tables prepared by the Office of the Chief Actuary in the (U.S.) Social Security Administration, run from 1900 to 1995. Italian mortality data run from 1872 to 2006.

We select values of life expectancies at birth from 1900 to 2006 for Italy, and from 1900 to 2005 for the United States. A plot of life expectancy at birth in Italy and the United States, for both sexes is provided in Figure 1. As may be observed also for most Western countries, life expectancies during the first and second half of the century are characterized by different trends, reflecting the shift that occurred in the primary causes of death.

The steady increase in life expectancy at birth observed in the first half of the century is an effect of the reduction in infant and child mortality. This trend is interrupted by obvious troughs during both the First and Second World Wars, and during the Spanish influenza epidemic. Exceptional events like these are excluded from our analysis, and their values are interpolated with the adjacent values. The slower upward trend starting to be observed in the late 1960s, reflects the reduction in old-age mortality (Torri and Vignoli, 2007).

Unlike Italy, the United States experienced a slow decline in mortality

Figure 1: Life expectancy at birth for Italy and the U.S., 1900-2005. Females (red lines) and males (blue lines) data.



during the 1980s and 1990s, which was especially evident for females. Although many researchers have tried to explain such an unexpected deceleration in the improvement of life expectancy, no clear answer has been provided yet (Meslé and Vallin, 2005; Pampel, 2002; Vaupel et al., 2006).

3 Methods

Our aim is to model the described trend of the stochastic process of life expectancy at birth, and forecast it. We first consider the Autoregressive Integrated Moving Average models (ARIMA), derived from the classic time series analysis.

Follow the Structural Time Series (STS) models, a more flexible class of models, where the assumption of stationarity is relaxed. These model aim to present the characteristic components of a series, rather than to represent the underling data generating process. The statistical formulation of the components needs to be flexible enough to capture the general changes in the direction of the series. The components are regarded as being driven by a random component, and handled in the state-space framework, with the state of the system representing the various unobserved components of the model. Once in the state-space form, the Kalman filter allows to update the state as soon as new observations are available.

Common traits are shared by the two models. Assuming a linear structural time series models, with several disturbance terms, the different components can be combined returning a model with a single disturbance. This is a reduced form, corresponding to an ARIMA model characterized by restrictions on the parameter space. More insights on this aspect can be found in Harvey (1989, p. 67).

3.1 The ARIMA model

Classic time series analysis is based on the theory of stationary stochastic processes. A stochastic process is said to be stationary in a weak sense if the mean and the variance of the process do not change over time, and the covariance between values of the process at different time points depends only on the distance between the time points, and not on the time itself. Stationary stochastic processes are included in the Autoregressive Moving Average (ARMA) model. If at least one of the conditions for stationarity is not fulfilled, the process is non-stationary and standard parameter estimates no longer have their conventional asymptotic properties. Processes exhibiting non-stationary behaviour are included in the wider class of ARIMA models. The order of integration of time series is the number of times that the series must be differenced to make it stationary.

In a more general notation, we have an ARIMA(p, d, q) model, in which p is the order of the autoregressive process, d indicates the order of integration, and q is the order of the moving average process. Expressed as a formula, the ARIMA(p, d, q) model is equal to:

$$\nabla^d Y_t = \delta + \phi_1 \nabla^d Y_{t-1} + ... + \phi_p \nabla^d Y_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + ... + \theta_q \epsilon_{t-q} \tag{1}$$

where δ is the drift term, a constant parameter indicating the average change of the variable over time. ϕ_i, for $i = 1, ..., p$, are the parameters of the autoregressive part and θ_j, for $j = 1, ..., q$, are the parameters of the moving average part. The symbol ∇^d is the differentiating operator of order d, and ϵ_{t-j} is a sequence of $i.i.d.$ random variables with mean zero and variance σ^2.

The model selection strategies developed by Box and Jenkins (1976) help to select an adequate model. The model is first identified, choosing within the family of ARIMA models (e.g., AR(p), MA(q), ARMA(p,q), ARIMA(p,d,q)) and then estimated through maximum likelihood. A diagnostic checking of the model adequacy is also performed by looking at the fitted errors that should be a white noise. If model diagnostics show that the selected model is not adequate, the three-step procedure is repeated until the appropriate model is specified.

3.2 Structural Time Series Model

A univariate Structural Time Series (STS) model is specified by components that, though unobservable, have a direct interpretation. Such a model not only provides the basis for predicting future observations, it also describes the salient features of a time series. The classical decomposition model proposed by Harvey (1989), is characterized by a deterministic nature, and is expressed in the following way:

$$y_t = \mu_t + s_t + \varepsilon_t \, ,$$

where μ_t is a slowly changing function, the so-called level component, s_t is a function with period d, the seasonal component, and ε_t is the stationary random noise component.

In the specific case of life expectancy, we decompose the series into the level and irregular components. The statistical formulation of the level component needs to be flexible enough to respond to general changes in the direction of the series. The model used to represent the process is called the *local linear trend* model or *linear growth* model. The model is described as:

$$e_t^0 = \mu_t + \varepsilon_t \, , \tag{2}$$

where μ_t is the trend and ε_t is a white-noise disturbance term which is assumed to be uncorrelated with any stochastic elements in μ_t.

The trend may take a variety of forms, and here specifically, we consider the linear case. We have a deterministic linear trend when

$$\mu_t = \alpha + \beta \cdot t \, ,$$

that becomes a stochastic trend by letting α and β follow random walks. However this would lead to a discontinuous pattern of μ_t, and better results are obtained working directly with the current level μ_t (Harvey, 1989).

Given $\mu_0 = \alpha$, stochastic terms may be introduced with the following equations describing the behavior of the local level μ_t and the local trend β_t:

$$\begin{aligned} \mu_t &= \mu_{t-1} + \beta_{t-1} + \eta_{1t} \\ \beta_t &= \beta_{t-1} + \eta_{2t} \, , \end{aligned} \tag{3}$$

where η_{it} are mutually uncorrelated withe-noise disturbances with zero means and variances $\sigma^2_{\eta_i}$. The current level μ_t changes linearly over time, but the growth rate or trend β_t may also change. The effect of η_{1t} allows the level of the trend to shift up and down, while η_{2t} allows the slope to change.

Provided the decomposition of the time series of life expectancy into level, slope, and irregular components, we have to test for the deterministic or stochastic nature of each component. The parameters σ^2_ε and $\sigma^2_{\eta_i}$ are unknown, and are referred to as the hyper-parameters. If $\sigma^2_{\eta_2}$ is zero, equation (3) reduces to:

$$\mu_t = \mu_{t-1} + \beta \cdot t \, .$$

If both η_{1t} and η_{2t} have zero variance, equation (3) reduces to:

$$\mu_t = \alpha + \beta \cdot t \, .$$

In this case the trend of the process is deterministic and it is forecast with a straight line.

Under the assumption of normal distribution of the disturbances of ε_t and η_{it}, the estimation of the STS model is obtained within the state-space framework, where each component of the decomposition model is allowed to evolve randomly over time. The initial sample is augmented by one observation, each advancing year, and the estimating equations are updated together with the parameters estimation, by means of a Kalman filter (Durbin and Koopman, 2001; Harvey, 1989). Follows a description of the estimation procedure.

3.2.1 State-Space models

The representation of a state-space model for a multivariate time series of an observable variable y_t, with dimension $N \times 1$, consists of two equations. The first is called the *measurement equation* and identifies the linear relationship between the unobservable *state vector* α_t, with dimension $m \times 1$, and the vector of measurements, y_t. The equation is as follows:

$$y_t = Z_t \, \alpha_t + d_t + \varepsilon_t \qquad t = 1, \ldots, T \, , \tag{4}$$

where the regression matrix Z_t has dimension $N \times m$, d_t is a vector of dimension $N \times 1$, and ε_t is the vector of dimension $N \times 1$ of serially uncorrelated disturbances with mean zero and covariance matrix H_t.

The second equation is called the *transition equation* and describes the stochastic process of the *state vector* through a dynamic linear system characterized by errors, with zero mean and known variance. The process is generated by a first order Markov process, and described by the following equation:

$$\alpha_t = T_t\,\alpha_{t-1} + c_t + R_t\,\eta_t \qquad t = 1, \ldots, T\,, \tag{5}$$

where T_t is the transition matrix with dimension $m \times m$, c_t is an $m \times 1$ vector, R_t is an $m \times g$ matrix, and η_t is an $g \times 1$ vector of serially uncorrelated disturbances with mean zero and covariance matrix Q_t.

The specification of the state-space system is completed by assuming that the initial state vector α_0 has a mean equal to a_0 and a covariance matrix equal to P_0. Additionally, it is assumed that the disturbances ε_t and η_t are uncorrelated with each other in all the time periods and with the initial state α_0.

The system matrices Z_t, H_t, T_t, R_t, and Q_t - here assumed to be time invariant - may depend on a set of unknown parameters, referred as hyper-parameters, that should be estimated. The hyper-parameters determine the stochastic properties of a model, while the parameters c_t and d_t affect only the expected value of the state and observations in a deterministic way.

The estimation of the conditional distribution of the unobserved sequence of states α_t, given the observed data points y_t, is solved by applying the optimal recursive algorithm, known as the Kalman filter (Kalman, 1960).

3.2.2 The Kalman Filter

The Kalman filter is a recursive algorithm that consists of a set of equations which allow for the updating of the estimate when a new observation becomes available. It is called filter, because it practically filter the noise away from the observation, in an optimal way. It combines all available measurement data and prior knowledge about the system to produce an optimal estimate of the desired variables α_t. The recursive algorithm is characterized by two distinct phases:

- the *prediction* phase, in which, given the available information, an optimal and a prior estimator of the *state vector* is obtained;
- the *updating* phase, in which, new information arrives and a posterior estimation is obtained.

Under the assumption of normal distribution of the disturbances and initial state vector α_0, it is possible to calculate the distribution of α_t, conditional on the information set at time t. Let a_{t-1} denote the optimal estimator[1]

[1] The one minimizing the mean squared error

of α_{t-1}, based on the observation up to and including y_{t-1}. P_{t-1} is the covariance matrix of the estimation error which is equal to $E[(\alpha_{t-1} - a_{t-1})(\alpha_{t-1} - a_{t-1})']$. Given the parameters a_{t-1} and P_{t-1}, the optimal estimator of α_t is given by (see equation (5)):

$$a_{t|t-1} \;=\; T_t\, a_{t-1} + c_t \tag{6}$$

while the covariance matrix of the estimation error is

$$P_{t|t-1} \;=\; T_t\, P_{t-1}\, T_t' + R_t\, Q_t\, R_t' \,. \tag{7}$$

The two equations are known as the *prediction equations*. Once a new observation y_t is available, the estimator of α_t, $a_{t|t-1}$, can be updated trough the following *updating equations*:

$$\begin{aligned} a_t &= a_{t|t-1} + P_{t|t-1}\, Z_t'\, F_t^{-1}(y_t - Z_t\, a_{t|t-1} - d_t) \\ P_t &= P_{t|t-1} - P_{t|t-1} Z_t'\, F_t^{-1} Z_t\, P_{t|t-1}\,, \end{aligned} \tag{8}$$

where $F_t = Z_t\, P_{t|t-1}\, Z_t' + H_t\,, \quad t = 1, \ldots, T\,.$

These equations state that the optimal estimate at time t, a_t, is equal to the best prediction of its value before y_t is available, corrected by an optimal weighting value times the prediction errors[2].

The prediction equations together with the updating equations, construct the Kalman filter. Therefore, the Kalman filter is a continuous succession of the prediction phase, when a preliminary guess about the state of nature is formulated, and the updating phase, when the initial guess is corrected. The corrections are determined based on how well the guess has performed in predicting the next observation.

The starting values of the Kalman filter may be specified in terms of $a_0 = 0$ and $P_0 = kI$, where I is the identity matrix and k a very large positive number, which may be considered a reasonable approximation of a prior state of ignorance. Based on a decision regarding the initial conditions, the Kalman filter provides the optimal estimator of the current state vector α_t, conditioned on the full information set $\{y_1, \ldots, y_t\}$.

3.3 Forecasting

After a satisfactory model is found for both methods, forecasts can be computed. Given the availability of data up to time T, the forecast will involve the observations y_T and the fitted residuals (i.e. the one-step-ahead forecast

[2]The prediction error is equal to the difference between y_t and the best prediction of its value.

errors) up to and including T. The l-steps ahead optimal predictor of the process, is the conditional expectation of y_{T+l} at time T. It is defined as an optimal predictor, since it minimized the mean square errors.

The estimation error, for any predictor, is equal to $y_{T+l} - \tilde{y}_{T+l|T} = [y_{T+l} - E(y_{T+l}|Y_T)] + [E(y_{T+l}|Y_T) - \tilde{y}_{T+l|T}]$, where Y_T denotes the information set $\{y_T, y_{T-1}, \ldots, \}$. Considering that the second term on the right hand side is fixed at time T, after squaring the whole expression and taking conditional expectation, we obtain the following expression for the $MSE(\hat{y}_{T+l|T}) = Var(y_{T+l}|Y_T) + [\hat{y}_{T+l} - E(y_{T+l}|Y_T)]^2$. The first term on the right hand side is not dependent on $\hat{y}_{T+l|T}$, and the second term is minimized when y_{T+l} equals the conditional mean $E(y_{T+l}|Y_T)$. This is written as:

$$\tilde{y}_{T+l|T} = E(y_{T+l}|Y_T) \tag{9}$$

Under the assumption of independence of the residuals the predictorss can be built up recursively by the *chain rule*.

4 Application

In the current section we apply the two models described in Section 3.1 and 3.2 to Italian life expectancy on the period from 1900 to 2006, and U.S. life expectancy from 1900 to 2005.

4.1 Forecasting life expectancy with an ARIMA model

When working with the ARIMA models, the first thing to do is to test for the stationarity of the series. Already from a visual inspection we can see that the series is characterized by an increasing trend over time and hence by a non-constant mean. It turned out indeed that the series are non-stationary and it is necessary to differentiate them in order to become stationary.

Applying the model selection strategies proposed by Box and Jenkins (1976), we select the most adequate ARIMA models for our data. Table 1 describes the order of the selected ARIMA models and the values of the estimated parameters, together with the corresponding confidence intervals. Bigger values of the drift δ are returned for females.

The same models are used to predict future life expectancy at birth until the year 2050. The estimated future values of the life expectancy at birth in the year 2050, together with the 80% and 95% prediction intervals are provided in Table 2 and plotted in Figure 2. We expect that in the year 2050 Italy will experience a period life expectancy at birth of 101 years for females and 94 for males, while for the U.S. we expect to observe a value equal to 94

Table 1: Order of the ARIMA model selected for Italy and the U.S. on the data periods 1900-2005, by sex.

	Female		Male	
	Italy	Usa	Italy	Usa
ARIMA(p,d,q)	(1,1,0)	(0,1,1)	(0,1,1)	(0,1,1)
δ	0.5765	0.2975	0.3479	0.2720
s.e.	(0.0628)	(0.0444)	(0.0542)	(0.0448)
ϕ_1	-0.4596	-	-	-
s.e.	(0.0956)	-	-	-
θ_1	-	-0.1739	-0.3092	-0.2125
s.e.	-	(0.1156)	(0.0886)	(0.1243)

for females and 87 for males. More information is provided by the prediction intervals.

Considering that the stability of the processes, is on the assumptions underlying the ARIMA model, we were curios to inspect the impact on the results of a changing data period. We thought it would be interesting to evaluate the future values of life expectancy in 2050 based on several data periods. For this purpose, we assumed that the ARIMA model estimated on the longest period 1900-2005 is also suitable for shorter periods. Estimations are recursively performed on a decreasing number of observations that move progressively from a longer (1900-2005) period to a conventionally chosen shorter data period (1980-2005). The smallest series included 26 observations because we deemed performing estimations on a smaller sample of data to be inappropriate. Figures 3 and 4 depict the estimated future life expectancy in the year 2050 and the corresponding parameters, whereby each of the values is the result of application of the model on a progressively smaller sample of data.

Predicted life expectancies shift from relatively constant values before 1950, to lower and increasing values for males and decreasing values for females after 1950. The different amplitudes of the prediction intervals are also remarkable. The changes observed around the year 1950 are expected, but the rapidity of the changes in the level of life expectancy in 2050, along with a change occurring in the trend, are surprising. The graphs suggest that the process may not be as stable as was initially assumed, and that changes occur exactly at that critical point. Further investigation of the assumption of stability of the process may be necessary to ensure that our estimations

Figure 2: Actual and forecast life expectancy at birth using an ARIMA models with associated 80% (dotted) and 95% (solid) prediction intervals. Italy, and the U.S., 1900-2050. Females (red lines), males (blue lines).

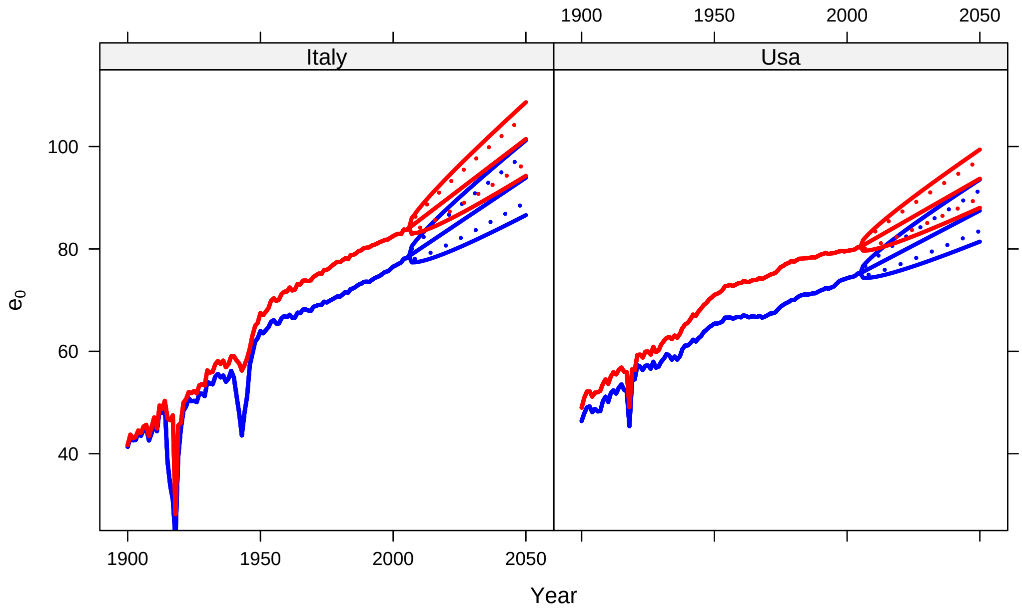


Table 2: Forecast life expectancy at birth in 2050 by ARIMA models with 80% and 95% prediction intervals. Italy, and the U.S.. Models evaluated on the data period 1900-2005.

	Italy	Usa
Female		
e_0	101.47	93.72
80% CI	(96.46–106.53)	(90.00–97.44)
95% CI	(93.79–109.21)	(87.47–100.00)
Male		
e_0	93.91	87.49
80% CI	(86.59–101.22)	(83.52–91.46)
95% CI	(86.59–101.22)	(80.68–94.36)

Table 3: Hyper-parameters from STS model on life expectancy. Italy and the U.S., 1900-2005, both sexes.

	Italy		Usa	
	F	M	F	M
σ^2_ϵ	0.2224	0.1992	0.0221	0.0474
$\sigma^2_{\eta_1}$	0.2863	0.3107	0.2665	0.2615
$\sigma^2_{\eta_2}$	0.0000	0.0000	0.0006	0.0000

are unbiased.

4.2 Forecasting life expectancy with a STS model

The results obtained in the former section motivate the decision to use a more flexible time series model like the STS model, where no assumption of stability and stationarity is made, and whose parameters are assumed to vary with time.

Combining the notion of the STS model presented in Section 3.2 and the state-space notation in sub-Section 3.2.1, we obtain the following local linear trend model to represent the behaviour of life expectancy at birth. The observable process e_0 is determined by the following *measurement equation*:

$$e^0_t = \left[\begin{smallmatrix} 1 & 0 \end{smallmatrix}\right] \alpha_t + \varepsilon_t \, , \tag{10}$$

and the *transition equations* linked to the *state vector*, α, can be written in the equivalent form:

$$\alpha_t = \left[\begin{smallmatrix} \mu_t \\ \beta_t \end{smallmatrix}\right] = \left[\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right] \left[\begin{smallmatrix} \mu_{t-1} \\ \beta_{t-1} \end{smallmatrix}\right] + \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right] \left[\begin{smallmatrix} \eta_{1t} \\ \eta_{2t} \end{smallmatrix}\right] \, , \tag{11}$$

where $Z = \left[\begin{smallmatrix} 1 & 0 \end{smallmatrix}\right]$, $T = \left[\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right]$, $H_t = \sigma^2_\varepsilon$, $R = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$ and $Q = \left[\begin{smallmatrix} \sigma^2_{\eta_1} & 0 \\ 0 & \sigma^2_{\eta_2} \end{smallmatrix}\right]$.

Applying the model to our data, we obtain the time varying values of the local level, local trend, and irregular components, plotted in Figure 5.

The estimated local levels show an increasing trend that is driven by the local trends and a stochastic component. The local trends are characterized by almost constant values. The presence of a deterministic slope is further confirmed by the values of $\sigma^2_{\eta_2}$, that are not significantly different from zero, with the only exception of U.S. females. The estimated values of the hyper-parameters are given in Table 3.

Figure 3: Estimation of the parameters of the ARIMA model for Italy and the U.S., by sex, evaluated on a data period beginning at different point in time T_0, and ending in 2005. Males (blue) and females (red) data.

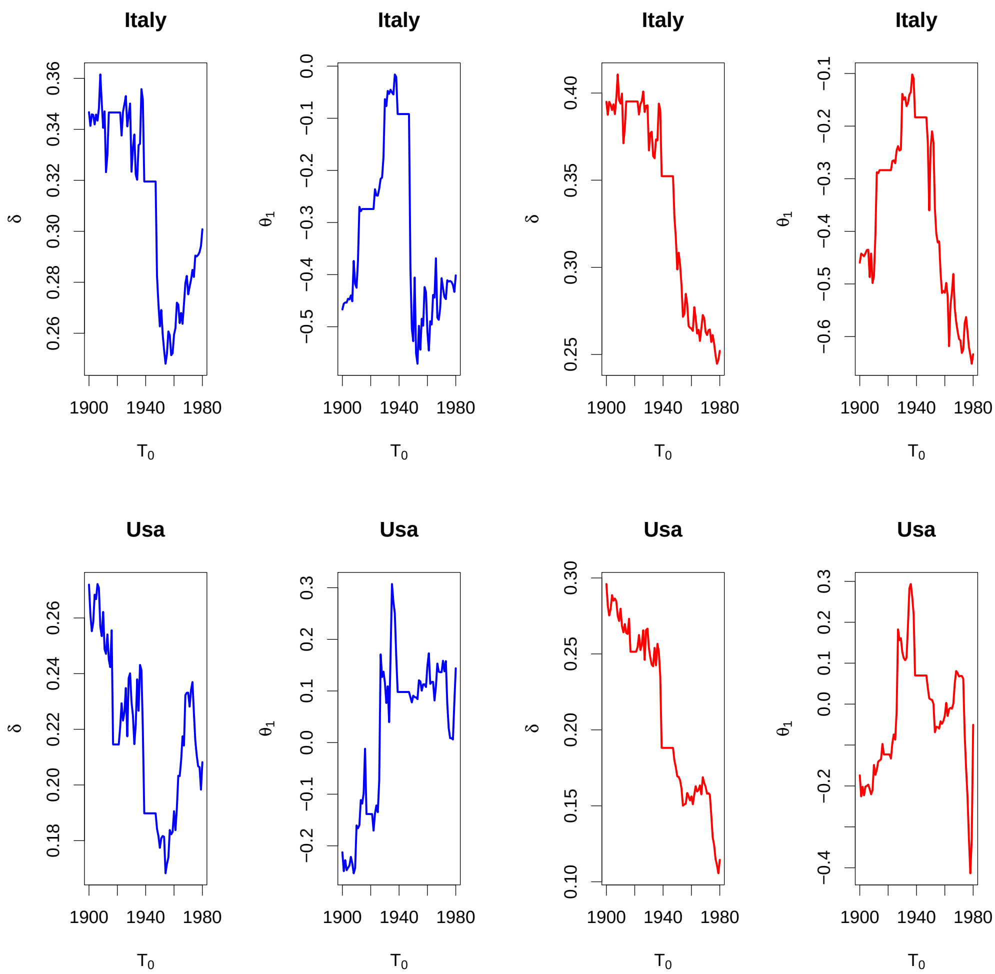


Figure 4: Forecasted life expectancy at birth for the year 2050, for Italy and the U.S., by sex evaluated on a data period beginning at different point in time and ending in 2005. Males (blue) and females (red) data.

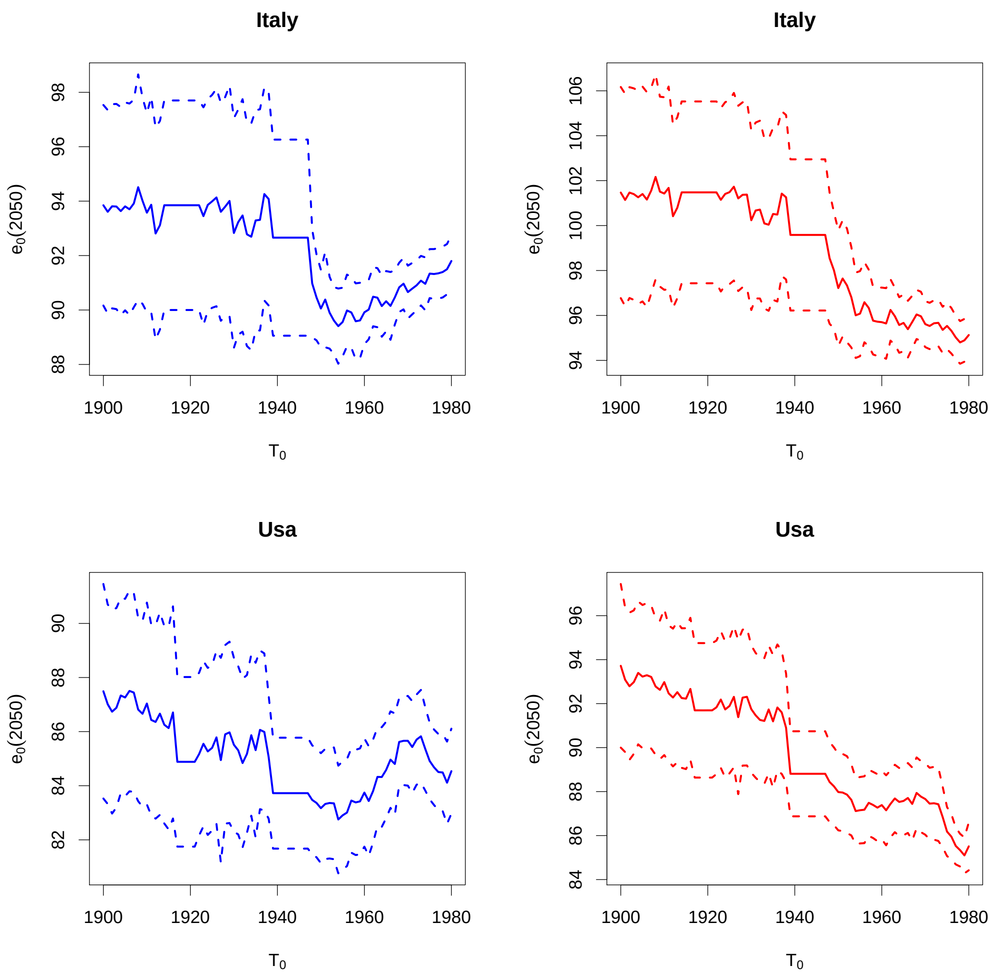


Figure 5: Local level, local trend, and irregular component. STS model applied on life expectancy at birth. Italy, and the U.S., 1900-2005. Females (red lines) and males (blue lines).

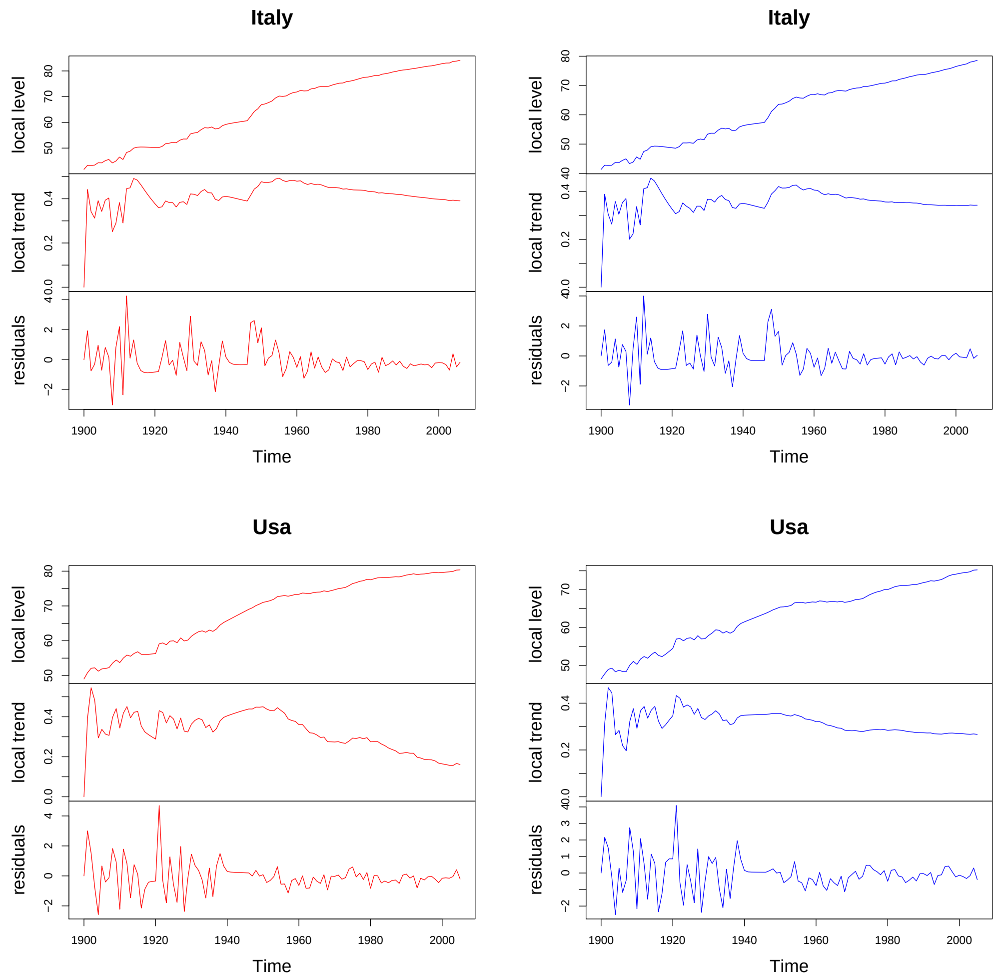


Table 4: Forecast life expectancy at birth in 2050 by STS model with 80% and 95% prediction intervals. Italy, and the U.S.. Models evaluated on the data period 1900-2005.

	Italy	Usa
Female		
e_0	101.32	87.62
80% CI	(95.86–106.78)	(77.91–97.34)
95% CI	(92.97–109.67)	(72.76–102.49)
Male		
e_0	93.67	87.52
80% CI	(87.99– 99.35)	(81.98–92.48)
95% CI	(84.99–102.35)	(79.21–95.26)

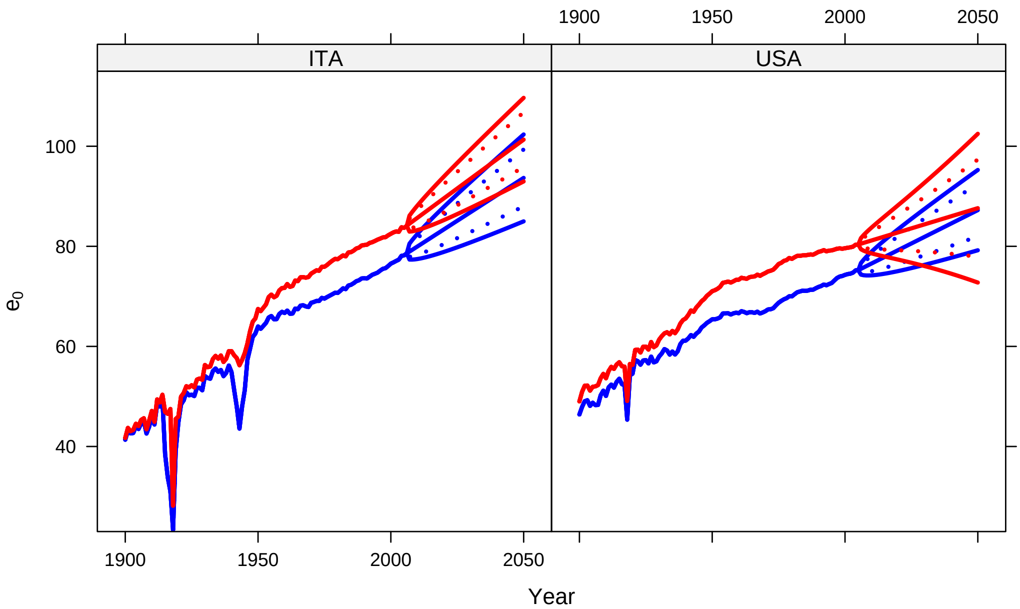
For those data returning a null variance of the local trend component, $\sigma^2_{\eta_2}$, the model used to describe the behaviour of life expectancy is reduced to the simpler *local level* model:

$$\begin{aligned} e_t^0 &= \mu_t + \varepsilon_t \\ \mu_t &= \mu_{t-1} + \beta + \eta_{1t} \, . \end{aligned}$$

Future values of life expectancy are plotted in Figure 6, and the corresponding values in the year 2050 are given in Table 4. The prediction suggests big advancements for the Italian female life expectancy, characterized by a steep increase. Predicted female life expectancy for the year 2050 reaches the value of 101 years for Italy, and 94 for the U.S.. The initial gap of 3 years observed between the U.S. and Italy in 2005 is expected to rise by 2050 to almost 8 years. The results for men show a similar trend. Future values of men life expectancy in the year 2050 are equal to 94 for Italy, and 87 for the U.S.. The initial gap in life expectancy between U.S. and Italy was three years in 2005, and is expected to widen by 2050 to 7 years.

A visual inspection suggests that the future values of female U.S. life expectancy follow a smooth path. It is questionable, however, whether the gap between male and female life expectancy will decline to such an extent. We can believe that if no actions are undertaken to improve U.S. female mortality, which has worsened in recent years, this scenario is likely.

Figure 6: Actual and forecast life expectancy at birth using an STS model with associated 80% (dotted) and 95% (solid) prediction intervals. Italy, and the U.S., 1900-2050. Females (red lines), males (blue lines).



5 Conclusions

The urge to compute mortality forecasts, utilized by governments when planning the allocation of their resources between the pension and health-care systems, stimulate a vivid interest in the topic.

In the current work we directly performed forecast of the stochastic process of life expectancy at birth, aiming to model the linear behaviour of the series observed in the past. With this respect, two different extrapolation methods are applied. We began applying the widely used ARIMA model to Italian and U.S. data, and found out that the assumed stability of the process is not confirmed by the data. Namely, the parameters of the ARIMA model that we assumed generated the process, are not constant over time, but changing.

Alternatively, we proposed a more flexible model, that relaxed the assumptions of stability and stationarity of the process, previously made. We used the so-called Structural Time Series models, generally applied in disciplines other than demography, such as engineering. Written in a state-space framework, such model allow to update the estimated future trends as the sample period changes.

The two distinct models used to forecast life expectancy produced quite similar results of the median value of life expectancy in the year 2050. The only exception is the U.S. female life expectancy, that shows a gap of approximately 6 years in the forecast value. The exception in U.S. females was already observed in the estimated value of the hyper-parameter of the local trend component. A value of the variance different from zero, captured the stochastic behaviour of the local trend component. Slightly different estimation of the prediction intervals are returned. Using a more flexible model we have been able to capture and forecast the slow decline in American female mortality, already observed from the 1980s. It is questionable whether these trend will persist in the future. Corrective actions may be undertaken by the government to improve recent trends in mortality.

References

Alders, M., N. Keilman, and H. Cruijsen (2007). Assumptions for long-term stochastic population forecasts in 18 European countries. *European Journal of Population 23*, 33–69.

Alho, J. M. (1989). Relating Changes in Life Expectancy to Changes in Mortality. *Demography 26*, 705–709.

Andrews, D. W. K. (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica, 61*, 821–856.

Berkeley Mortality Database (1995). University of California, Berkeley (USA). Available at `www.demog.berkeley.edu/ bmd/states.html`.

Box, G. E. and G. M. Jenkins (1976). *Time Series Analysis: Forecasting and Control*. Holden-Day, San Francisco.

De Beer, J. and M. Alders (1999). Probabilistic Population and Household Forecasts. Technical report, Joint ECE-EUROSTAT work session on Demographic Projections.

Durbin, J. and S. J. Koopman (2001). *Time Series Analysis by State Space Methods*. Oxford University Press.

Hansen, B. (1992). Tests for parameter instability in regressions with I(1) processes. *Journal of Business & Economic Statistics 10*, 321–335.

Harvey, A. C. (1989). *Forecasting, structural time series models and the Kalman filter*. New York: Cambridge University Press.

Human Mortality Database (2008). University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at `www.mortality.org` (Data downloaded on September 2008).

Kalman, R. E. (1960). A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering 82*, 34–45.

Keilman, N., D. Q. Pham, and A. Hetland (2001). Norway's Uncertain Demographic Future. Technical report, Social and Economic Studies 105. Oslo: Statistics Norway.

Kuan, C. M. and K. Hornik (1995). The generalized fluctuation test: A unifying view. *Econometric Reviews 14*, 135–161.

Lee, R. and L. Carter (1992). Modelling and Forecasting the U.S. Mortality. *Journal of the American Statistical Association 87*, 659–671.

Meslé, F. and J. Vallin (2005). Why do american trends in female old age mortality diverge from those of other advanced industrialized countries? PAA Meeting.

Pampel, F. (2002). Cigarette Use and the Narrowing Sex. *Population and Development Review 28*, 77–104.

Torri, T. and D. Vignoli (2007). Forecasting the Italian population, 2005-2055: a stochastic approach. *Genus 63*(1/2), 93–118.

Vaupel, J. W., R. Rau, C. G. Camarda, and K. G. von Kistowsky (2006). Can heterogeneity of populations explain differences in mortality? Working paper, Chestnut Hill, MA, Center for Retirement Research.

White, K. M. (2002). Longevity advances in high-income countries, 1955-96. *Population and Development Review 28*, 59–67.