Noname manuscript No.
(will be inserted by the editor)

A stochastic model of mortality rate with memory

A stochastic model of mortality rate with memory

Areli Ornelas · Francisco Delgado-Vences · Saul Diaz-Infante

Received: February 25, 2020/ Accepted: date

Abstract An accurate estimation of mortality rates is essential to make decisions. For example, the ensures companies, investment projects, among others, projects its operations according to estimations of these rates. However, the vast number of variables and its intricate relation implies a challenge for the estimation of these rates.

We assume the following hypothesis: the mortality rate has a strong relationship with its owns past. In this line, we propose a stochastic model with long-term memory that describes mortality. Then, using data from Italy, we provide statistical evidence via Hurst parameter estimation that does not reject our hypothesis. Further, we extract a subset of data to evaluate its forecasting performance, and we observe a good estimation.

To the best of our knowledge, our contribution is the first attempt that includes long term memory effects in the formulation of a model to describe mortality rates. Our results suggest that the hypothesis of an imperfect correlation intensity across generations would be more realistic and that sex is an important variable to consider in the next formulations.

Francisco Delgado-Vences
Conacyt-Universidad Autonomoa de Mexico.
Instituto de Matemáticas, Oaxaca, México
E-mail: delgado@im.unam.mx

Areli Ornelas
Conacyt-Instituto Politecnico Nacional-CICIMAR,
La Paz, México
E-mail: arelly.ornelas@conacyt.mx

Saul Diaz-Infante
Conacyt-Universidad de Sonora
Departamento de Matemáticas,
Hermosillo, Sonoran México E-mail: sdinfante@conacyt.mx

1 Introduction

Future planning in the demographic, economic, and actuarial areas is crucial. For instance, proper planning in social programs, government budgets, cost of insurance, and others depends on the forecast. However, constant changes in technology, lifestyle, migration, to name a few, make predicting a demanding task. These changes promote research in the field of risk management. Mortality impacts directly in cash costs and therefore need a reliable future projection.

Several models have been proposed to describe mortality. Pitacco et al. (2009) provides an interesting review of the historical background to early mortality tables. Some authors in their attempts to calibrate models to mortality rates make the assumption of perfect correlation across generations. However, one could think that correlation among close generations is high but not perfect. This hypothesis is our start point for this paper, we model the mortality hazard rates for the human population considering a model that include the high correlation among generations. In [10], they study a cohort-based model which uses the imperfect correlation of mortality intensity across generations. They implemented it on UK data for the period 1900-2008.

In this section we present a brief review of the Milevesky-Promislow model and discuss about long-range dependence. Milevesky-Promislow model the mortality hazard rate using the exponential of an Stochastic Differential Equation (see [18]).

Stochastic Differential Equations (SDEs) provide a powerful and flexible probabilistic structure for modeling phenomena in a multitude of disciplines: finance, biology, molecular dynamics, chemistry, survival analysis, epidemiology, just to name a few examples, the reader could see [22] or [11] for further information. The fractional Ornstein-Uhlenbeck process (FOU) process we are interested is an example of an SDE.

In [6], the authors present a generalization of the Milevesky-Promislow model that covers the case with diffusion coefficient as a function of t instead of a constant like in [18]. Also, they consider a type of Autoregressive model for the logarithm of the hazard rate[1]. It is known that all stationary invertible ARMA(p,q) processes are short memory processes[2]. In particular the authors in [6] used a modified AR(1) process, that is still a short memory process. In our model we consider the case with a long memory process. For another generalization of the Milevesky-Promislow model see also [26].

In the other hand we are interested in the long range dependence. Hurst [8] observed a phenomenon which is invariant to changes in scale, when he was

[1] see equations (13) or (16) in [6]

[2] see for instance page 739 in [1] or section 5.2 in [27]

studying the river water levels along the Nile. Another scale-invariant phenomenon was observed in studies of problems connected with traffic patterns of packet flows in high-speed data networks such as the Internet (see preface of [25]).

Long Range dependence (LRD) (also known as long memory, strong dependence or persistence) denotes the property of a stochastic process or a time series to exhibit persistent behavior across long time periods (see [25]). LRD is used in areas such as finance, econometrics, Internet, hydrology, climate studies, linguistics, geophysics or DNA sequencing among others.

One example of time-continuous long-range dependence model was introduced by Mandelbrot and van Ness (see [16]), they define the term fractional Brownian motion (fBm) for a Gaussian process with a specific covariance structure and studied its properties. This process is a generalization of classical Brownian motion also known as the Wiener process ([25]). The self-similarity and long-range dependence properties make the fractional Brownian motion suitable to model driving noises in different applications such as hydrology, short-term rainfall prediction, finance, etc..

The fBm is a particular case of the Fractional Gaussian noises (FGN), that are a type of stochastic processes. The FGN have the property of self-similarity and is used to model persistent dependency in time series. The autocovariance function of FGN is characterized by the Hurst exponent (H parameter). fBM and fOU are examples of fractional Gaussian noise.

We now discuss the model of Milevsky-Promislow Model (see [18] for the original paper or [6] for a recent generalization).

Set the survival probability of an individual aged x in the period $[t, T]$ as

$$S(t,T) := \mathbb{E}\Big[\exp\Big(-\int_t^T h_x(u)du\Big)\Big|\mathcal{F}_t\Big], \tag{1}$$

where $\{\mathcal{F}_t\}_{t\geq 0}$ is a filtration which represent the information until time t and $h_x(t)$ is the stochastic force of mortality or hazard rate. According with the Milevsky-Promislow Model $h_x(t)$ is given by

$$h(t) = h_0 \exp(\alpha_0 t + \alpha_1 Y_t), \tag{2}$$

where $h_0, \alpha_1, \alpha_2 > 0$. The process Y_t satisfies the SDE:

$$dY_t = -\lambda Y_t dt + \sigma dB_t, \tag{3}$$

where B_t is a Brownian motion, $Y_0 = 0$ and $\sigma, \lambda > 0$.

In this paper we assume, according with the Milevsky-Promislow Model, that h_t is given by (2). However Y_t satisfies an SDE whose solution is the fractional Ornstein-Uhlenbeck process. Indeed, we will assume that the stochastic process Y_t satisfies the following stochastic differential equation (SDE):

$$dY_t^H = -\lambda Y_t^H dt + \sigma dB_t^H, \tag{4}$$

where B_t^H is a fBM with Hurst parameter $1/2 \leq H < 1$, $Y_0^H = 0$, and $\sigma, \lambda > 0$. Y_t^H is called a fractional Ornstein-Uhlenbeck process (fOU) and has been studied exhaustively in the last decades.

A priori, we will assume that $\alpha_1 = 1$. Once we estimated the parameter H, we will adjust $\alpha_1 = T^{-H}$, which is devoted to control the variance of the process Y_t. We will use Y_t to denote Y_t^H.

Since stochastic mortality rate models take into account long time phenomena, we suggest a generalization of Milevesky-Promislow model, given by the equations (2) and (4), which cover the case of long range dependence of the data.

The main difference between this model and the original presented in [18] is that they consider the driving noise in the SDE as a standard Brownian motion instead of the fractional Brownian motion as in our model.

The paper is organized as follows. In section 2 we present a brief review of the fBM and fOU and we discuss some important properties of the fOU. Section 3 is dedicated to describe parameters estimation. In first place we discuss the estimation of parameter α_0. In second place we review several methods for estimating Hurst parameter H. Finally we present some results devoted to estimate the parameters σ, λ in the SDE. The results of the model are discussed in section 4.

2 On the fractional Gaussian noise

2.1 Fractional Brownian motion

We consider the Gaussian process$\{B_t^H, t \geq 0\}$, with $H \in (0,1)$, and with zero-mean and covariance function given by

$$R_H(t,s) := \mathbb{E}(B_s^H B_t^H) = \tfrac{1}{2}\big(t^{2H} + s^{2H} - |t-s|^{2H}\big). \tag{5}$$

This stochastic process is called a *fractional Brownian motion* (fBm) and was introduced by Kolmogorov [12] and studied by Mandelbrot and Van Ness in [16]. The parameter H is called Hurst index because of the statistical analysis developed by the climatologist Hurst [8]. The fBm is a generalization of

Brownian motion without independent increments, also it is a continuous-time Gaussian process.

The fBm has the properties of Self-similarity and stationary increments. Moreover, its Sample-paths are almost nowhere differentiable. However, almost-all trajectories are Hölder continuous of any order strictly less than H: for each such trajectory, there exists a finite constant C such that

$$\mathbb{E}\big(|B_t^H - B_s^H|\big) \leq C|t-s|^{H-\epsilon},$$

for every $\epsilon > 0$.

For $H = \frac{1}{2}$ the covariance can be written as $R_{1/2}(t,s) = \min(s,t)$ and the process $B_t^{1/2}$ is the ordinary Brownian motion (Bm). The increments of this process, in disjoint intervals, are independent. However, the increments are *not* independent for $H \neq \frac{1}{2}$.

Set $X_n = B_n^H - B_{n-1}^H$, $n \geq 1$ a stochastic process. $\{X_n, n \geq 1\}$ is a Gaussian stationary sequence with unit variance and covariance function:

$$\begin{aligned}\rho_H(n) &= \frac{1}{2}\Big((n+1)^{2H} + (n-1)^{2H} - (2n)^{2H}\Big)\\ &\approx H(2H-1)n^{2H-2} \to 0, \quad \text{when } n \to \infty.\end{aligned}$$

Therefore,

- if $H > \frac{1}{2}$ then $\rho_H(n) > 0$ for n large enough and $\sum_{n=1}^{\infty} \rho_H(n) = \infty$. This is persistent process with positive correlation. In this case, we say that X_n has *long-Range dependence* property.
- If $H < \frac{1}{2}$, then $\rho_H(n) < 0$ for n large enough and $\sum_{n=1}^{\infty} \rho_H(n) < \infty$. This is an anti-persistent process with negative correlation.

For further information on fBM see [25], [21] or [19].

2.2 Fractional Ornstein-Uhlenbeck process (fOU)

The fOU is an SDE driven by a fractional Brownian motion. The same model was used in [18] or [6].

As we mentioned before, the survival probability $S(t,T)$ of an individual aged x in the period $[t,T]$, is given in the Equation (1) and $h_x(t)$ is the stochastic force of mortality or hazard rate given by the Equation (2). We will assume that Y_t is an stochastic process that satisfies the SDE:

$$dY_t^H = -\lambda Y_t^H dt + \sigma dB_t^H, \tag{6}$$

where B_t^H is a fBM with Hurst parameter $1/2 \leq H < 1$, $Y_0 = 0$, and $\sigma, \lambda > 0$. This SDE is the fractional Ornstein-Uhlenbeck process.

There are substantial differences in trying to solve Equation (6) with respect to the method use in [18] or [6]. We now discuss some of these differences. First we interpret the SDE (6) as

$$Y_t^H = -\lambda \int_0^t Y_s^H ds + \sigma B_t^H. \tag{7}$$

Notice that the equation above does not have a stochastic integral because we are considering the case with additive noise. Nevertheless, it is possible to consider the general case with multiplicative noise in which case it is necessary to define a stochastic integral with respect to fractional Brownian motion as a pathwise Riemann–Stieltjes integral (see, e.g., [28] for the original definition and [5] for advanced results).

Coming back to Equation (7), Cheridito et al in [3] have introduced the fractional Ornstein-Uhlenbeck process (fOU) and they have shown that the process

$$Y_t^H = \sigma \int_0^t e^{-\lambda(t-u)} dB_u^H, \tag{8}$$

is the unique almost surely continuous-path process which solves (7) (see also Theorem 1.24 in [25]). The integral in Equation (8) is a pathwise Riemann–Stieltjes integral. The fOU process is neither Markovian nor a semimartingale for $H \in (1/2, 1)$ (see [5]) but remains Gaussian and ergodic.

Moreover, when $H \in (1/2, 1)$, Y_t even presents the long-range dependence property (see Cheridito [3] or [25]).

The variance of the fOU process Y_t is given by the following expression (see[31]):

$$Var(Y_t) = \sigma^2 2H e^{-2\lambda t} \int_0^t s^{2H-1} e^{2\lambda s} ds. \tag{9}$$

Notice that when $H = 1/2$ we get

$$Var(Y_t) = \frac{\sigma^2}{2\lambda}\big(1 - e^{-2\lambda t}\big), \tag{10}$$

which is the variance of the standard Ornstein-Uhlenbeck process (see for instance [17] page 143).

If we consider the constant $\alpha_1 = T^{-H}$ and the Equation (9), the expression for the variance of $\alpha_1 Y_t$ is given by

$$\begin{aligned} Var(\alpha_1 Y_t) = \alpha_1^2 Var(Y_t) &= \alpha_1^2 \sigma^2 2H \int_0^t s^{2H-1} e^{-2\lambda(t-s)} ds \\ &\leq \alpha_1^2 \sigma^2 2H \int_0^t s^{2H-1} ds = \alpha_1^2 \sigma^2 2H \frac{s^{2H}}{2H}\Big|_{s=0}^t \\ &= \alpha_1^2 \sigma^2 t^{2H} = \sigma^2 (t/T)^{2H}, \end{aligned} \tag{11}$$

then $Var(\alpha_1 Y_t) \leq \sigma^2$ since $0 \leq t \leq T$, this implies that the variance of $\alpha_1 Y_t$ is bounded by a constant that does not depend on time. We will use α_1 to control the variance of the process Y_t.

3 Estimation of the parameters

In this section we will describe a methodology to estimate the parameters. We will use some R-libraries in order to estimate the parameters.

We need to estimate α_0, α_1 as well as σ, λ for the SDE model that we described in the previous section. Furthermore, The Hurst parameter (H) involved in the driven fractional Brownian motion will be also estimated, however, this estimation is highly complicated. To solve this problem we will use the empirical evidence that the Hurst value in the equation (6) is preserved, this means that the value of the Hurst parameters H, in the equation (6), for the fBm B_t^H and the one for the fractional Gaussian noise Y_t^H are the same. Observe that $\alpha_1 = T^{-H}$ will be calculated using the estimated Hurst parameter.

3.1 Estimation of the parameter α_0.

The model is given by the equation (2). In order to estimate α_0 we will assume that $\alpha_1 = 1$. Moreover, h_0 will be fixed to the initial value of the observed rates $h(t)$ Taking ln we obtain

$$\ln h(t) = \ln h_0 + \alpha_0 t + Y_t. \tag{12}$$

One simple method to estimate the parameter α_0 is by minimizing the sum of the square errors. Let S be given by

$$S := \sum_{t_{initial}}^{t_{final}} \Big(\ln h(t) - \ln h_0 - \widehat{\alpha_0} t \Big)^2.$$

Taking derivative of S with respect to $\widehat{\alpha_0}$ we get

$$\frac{\partial S}{\partial \widehat{\alpha_0}} = -2 \sum_{t_{initial}}^{t_{final}} \Big(\ln h(t) - \ln h_0 - \widehat{\alpha_0} t \Big) t = 0,$$

and from this equation we obtain $\widehat{\alpha_0}$:

$$\widehat{\alpha_0} = \frac{\sum_t t \ln h(t) - \ln h(0) \sum_t t}{\sum_t t^2}. \tag{13}$$

Once we have estimated α_0 we proceed to estimate the Hurst parameter, σ and λ.

3.2 Relation between the Hurst parameter and the H-index in the FOU

In this subsection we will discuss the procedure we have used to estimate the parameter H.

We will use the following empirical fact. Suppose that a fOU process is driving with a fBM with a given Hurst parameter H_0. Yerlikaya-Okzurt et al ([30]) have shown a relationship between the Hurst parameter H of the fractional Brownian motion and the Hurst parameter of the fractional Gaussian noise given by an SDE. In fact, they have found statistical evidence that the fOU should have the same value H_0 that the fBM (see table 1 Yerlikaya-Okzurt et al). Therefore, at least empirically, the value of H is the same. Then, it is possible to choose the same value of the parameter H for both processes. A formal proof of this fact, up to our knowledge, is missed.

The subsequent sections are devoted to present several methods to estimate the Hurst parameter for the fBM.

3.3 Estimation of the self-similarity index H for the fBM

The last subsection allows us to estimate the parameter H in one simple way. According to equation (12), the residuals are given by the expression

$$\hat{Y}_t = \ln h(t) - \ln h_0 - \widehat{\alpha}_0 t.$$

$\hat{Y}_t$ is a fractional Gaussian noise, so that we can use it to estimate H. Afterwards we will use $\hat{H}$ to approximate the Hurst parameters of the fractional Brownian motion B_t^H. For this purpose we will review some methods to estimate the parameter H.

3.3.1 R over S Analysis

Following [29]. The analysis begins dividing a time series $\{Z_i\}$ of length L into d subseries of length n and denote it by $\{Z_{i,m}\}$, $m = 1, \ldots, d$. Then, for each subseries $\{Z_{i,m}\}$, $m = 1, \ldots, d$:

1. Find the mean E_m and standard deviation S_m.
2. Normalize the data $Z_{i,m}$ by subtracting the sample mean $X_{i,m} = Z_{i,m} - E_m$ for $i = 1, \ldots, n$.
3. Create a cumulative time series $Y_{i,m} = \sum_{j=1}^{i} X_{j,m}$. for $i = 1, \ldots, n$
4. Find the range $R_m = \max\{Y_{1,m}, \ldots, Y_{n,m}\} - \min\{Y_{1,m}, \ldots, Y_{n,m}\}$;
5. Rescale the range R_m/S_m .
6. Calculate the mean value of the rescaled range for all subseries of length n

$$(R/S)_n = \frac{1}{d} \sum_{m=1}^{d} R_m/S_m.$$

It can be shown (see [29]) that the R/S statistic asymptotically follows the relation:

$$(R/S)_n \sim cn^H,$$

where c is a constant. Thus, the value of H can be obtained by running a simple linear regression over a sample of increasing time horizons

$$\log(R/S)_n = \log c + H \log n.$$

Equivalently, we can plot the $(R/S)_n$ statistics against n on a double-logarithmic paper. If the returns process is white noise then the plot is roughly a straight line with slope 0.5. If the process is persistent then the slope H is greater than 0.5; if it is anti-persistent then the slope H is less than 0.5. The "significance" level of the estimated parameter H is usually chosen to be one over the square root of sample length, i.e. the standard deviation of a Gaussian white noise.

A major drawback of the R/S analysis is that no asymptotic distribution theory has been derived for the Hurst parameter H . The only results known are for the rescaled (but not by standard deviation) range R_m itself, see [15].

3.3.2 Method of rescaled range analysis R/S

Following [25], chapter 9. This method was suggested by Hurst (1951). The series $\{X_j, 1 \leq j \leq N-2\}$ is divided into K nonoverlapping blocks such that each block contains M elements where M is the integer part of N/K. Let $t_i = M(i-1)$, where $t_i = M(i-1)$ is the starting point of the ith block for $i = 1, \ldots, K$. Define

$$R(t_i, r) = \max[W(t_1, 1), \ldots, W(t_i, r)] - \min[W(t_1, 1), \ldots, W(t_i, r)],$$

where r takes values in natural number whenever r satisfy the inequality $t_i + r \leq N$. Moreover, $W(t_i, k)$ is set as

$$W(t_i, k) = \sum_{j=0}^{k-1} X_{t_i+j} - k\left(\frac{1}{r}\sum_{j=0}^{r-1} X_{t_i+j}\right), \quad k = 1, \ldots, r.$$

Note that $R(t_i, r) \geq 0$ since $W(t_i, r) = 0$ and the quantity $R(t_i, r)$ can be computed only when $t_i + r \leq N$. Define

$$S^2(t_i, r) = \frac{1}{r}\sum_{j=0}^{r-1} X_{t_i+j}^2 - \left(\frac{1}{r}\sum_{j=0}^{r-1} X_{t_i+j}\right)^2.$$

The ratio $R(t_i, r)/S(t_i, r)$ is called the rescaled adjusted range. It is computed for a number of values of r that makes sense according to the definition. Observe that, for each value of r, we obtain a number of R/S samples. The number of samples decrease as r increases. However, the resulting samples are not independent. It is believed that the R/S-statistic is proportional to r^H as $r \to \infty$ for the fractional Gaussian noise. Assuming this property, it is possible to regress $log(R/S)$ against $log(r)$ to obtain an estimator for H.

3.3.3 FDWhittle Estimator

Following[23]. The Local Whittle Estimator (LWE) is a semiparametric Hurst parameter estimator based on the periodogram. It assumes that the spectral density $f(\omega)$ of the process can be approximated by the function

$$f_{c,H}(\omega) = c\omega^{1-2H}, \tag{14}$$

for frequencies ω in a neighborhood of the origin, c is a constant. The periodogram of a time series $\{X_t, 1 \geq t \geq N\}$ is defined by

$$I_N(\omega) = \frac{1}{2\pi N}\left|\sum_{t=1}^{N} X_t e^{i\omega t}\right|^2,$$

where $i = \sqrt{-1}$. Usually, it is evaluated at the Fourier Frequencies $\omega_{j,N} = \frac{2\pi j}{N}$, $0 \leq j \leq [N/2]$. Note that the periodogram is the norm of the Discrete Fourier transform of the time series (see Section 6.1.2 in [24] for instance).

The LWE of the Hurst parameter, $\hat{H}_{LWE}(m)$ is implicitly defined by minimizing

$$\sum_{j=1}^{m} log f_{c,H}(\omega j, N) + \frac{I_N(\omega j, N)}{f_{c,H}(\omega j, N)},$$

with respect to c and H, with $f_{c,H}$ defined in (14).

3.4 Estimation of σ and λ

There are several methods to estimate parameters σ and λ. For instance see [25] or the references in [20] or in [14]. In the following section we will do a brief review of some of these methods.

3.4.1 Estimation σ.

Brouste and Iacus [2] proposed some consistent and asymptotically Gaussian estimators for the parameters σ, λ and H of the discretely observed fractional Ornstein-Uhlenbeck process solution expressed in the stochastic differential equation. There is a restriction on the estimation of the drift λ: the results are valid only in the case when $1/2 < H < 3/4$.

The key point of this method of estimation is that the Hurst exponent H and the diffusion coefficient σ can be estimated without estimating λ. We will use this method to estimate the parameters σ and λ. Notice that H was already estimated.

Let $\boldsymbol{a} = (a_0, \ldots, a_K)$ be a discrete filter of order $L \geq 1$ and length $K+1$, $K \in \mathbb{N}$ and we require $L \leq K$, i.e.

$$\sum_{k=0}^{K} a_k k^j = 0 \quad \text{for } 0 \leq j \leq L-1 \quad \text{and} \quad \sum_{k=0}^{K} a_k k^L \neq 0.$$

Let it be normalized

$$\sum_{k=0}^{K} (-1)^{1-k} a_k = 1.$$

We will also consider a *dilated* filter $\boldsymbol{a}^{(2)}$ associated to $\boldsymbol{a}$. For $0 \leq k \leq K$ we define

$$a_k^{(2)} = \begin{cases} a_{k'}, & \text{if } k = 2k' \\ 0, & \text{otherwise.} \end{cases}$$

Since $\sum_{k=0}^{2K} a_k^2 k^j = 2^j \sum_{k=0}^{K} a_k k^j$ then the filter $\boldsymbol{a}^{(2)}$ has the same order than $\boldsymbol{a}$.

Let $Y^T = (Y_t : 0 \leq t \leq T)$ be the sample path of the solution of (8). A discretization of Y^T is

$$(X_n := Y_{n\Delta_N}, n = 0, \ldots, N), \qquad N \in \mathbb{N},$$

where $\Delta_N = T/N$ and N is the number of observations of Y_t. We denote by

$$V_{N,\boldsymbol{a}} := \sum_{i=0}^{N-K} \left(\sum_{k=0}^{K} a_k X_{i+k} \right)^2,$$

the *generalized quadratic variation* associated to the filter $\boldsymbol{a}$ (see for instance [9]). Then, define the following estimators for H and σ.

$$\hat{H}_N := \tfrac{1}{2} \log_2 \left(\frac{V_{N,\boldsymbol{a}^2}}{V_{N,\boldsymbol{a}}} \right), \tag{15}$$

$$\hat{\sigma}_N := \left(-2 \frac{V_{N,\boldsymbol{a}}}{\sum_{k,l} a_k a_l |k-l|^{2\hat{H}_N} \Delta_N^{2\hat{H}_N}} \right)^{1/2}. \tag{16}$$

Brouste and Iacus (see Th. 1 in [2]) have shown the next result.

Theorem 1 *Let $\boldsymbol{a}$ be a filter of order $L \geq 2$. Then, both estimators $\hat{H}_N$ and $\hat{\sigma}_N$ are strongly consistent, i.e.*

$$(\hat{H}_N, \hat{\sigma}_N) \xrightarrow{\text{a.s.}} (H, \sigma) \quad \text{as} \quad N \to +\infty.$$

Moreover, we have asymptotic normality property: for all $H \in (0,1)$,

$$\sqrt{N}(\hat{H}_N - H) \xrightarrow{\mathcal{L}} N(0, \Gamma_1(\boldsymbol{a}, \sigma, H)) \quad \text{as} \quad N \to +\infty,$$
$$\frac{\sqrt{N}}{\log N}(\hat{\sigma}_N - \sigma) \xrightarrow{\mathcal{L}} N(0, \Gamma_2(\boldsymbol{a}, \sigma, H)) \quad \text{as} \quad N \to +\infty,$$

where Γ_1 and Γ_2 are symmetric positive definite matrices depending on σ, H and the filter $\boldsymbol{a}$. □

With this result is possible to obtain an estimator for the parameters σ.

We will use the following two filters:

- Classical filter. Let $K > 0$ and define
$$a_k := \frac{(-1)^{1-k}}{2^k}\binom{K}{k} = \frac{(-1)^{1-k}}{2^k}\frac{K!}{k!(K-k)!} \qquad \text{for } 0 \le k \le K.$$
- Daubechies filters (see [4] for the original definition). We set $K = 4$ and define the Daubechies filter by
$$\frac{1}{\sqrt{2}}\big(0.48296291314453, -0.8365163037378, 0.22414386804201, 0.12940952255126\big).$$

3.4.2 Estimation of the drift parameter λ when both H and σ are known

Hu and Nualart, [7] have shown that

$$\lim_{t\to\infty} Var(Y_t) = \lim_{t\to\infty} \frac{1}{t}\int_0^t Y_t^2 dt = \frac{\sigma^2 \Gamma(2H+1)}{2\lambda^{2H}} := \mu_2.$$

This equation gives a λ estimator, namely

$$\hat{\lambda}_N = \left(\frac{2\hat{\mu}_{2,N}}{\hat{\sigma}_N^2 \Gamma(2\hat{H}_N + 1)}\right)^{-\frac{1}{2\hat{H}_N}}, \tag{17}$$

where $\hat{\mu}_{2,N}$ is the empirical moment of order 2, i.e

$$\hat{\mu}_{2,N} = \tfrac{1}{N}\sum_{n=1}^{N} X_N^2.$$

Set $T_N = N\Delta_N$. We have the next result.

Theorem 2 *Let $H \in \left(\frac{1}{2}, \frac{3}{4}\right)$ and a mesh satisfying the condition $N\Delta_N^p \to 0$, $p > 1$, and $\Delta_N (logN)^2 \to 0$ as $N \to +\infty$. Then, as $N \to +\infty$,*

$$\hat{\lambda}_N \xrightarrow{\text{a.s.}} \lambda,$$

and

$$\sqrt{T_N}(\hat{\lambda}_N - \lambda) \xrightarrow{\mathcal{L}} N(0, \Gamma_3(\sigma, H)),$$

where $\Gamma_3(\sigma, H) = \lambda\left(\frac{\sigma_H}{2H}\right)^2$ *and*

$$\sigma_H^2 = (4H+1)\left(1 + \frac{\Gamma(1-4H)\Gamma(4H-1)}{\Gamma(2-2H)\Gamma(2H)}\right).$$

□

For the proof see Theorem 2 in [2].

4 Results with the fractional Ornstein-Uhlenbeck model

In this section we present the estimated mortality rates with the use of the model described in section 1. We obtain the data from the website of Human Mortality Database for the Italian population between 1950 to 2004.

In first place, we present the estimation of the H parameter. In second place, we present the results on simulated mortality rates using equations (16)-(17) to estimate the parameters σ, λ. The parameter α_0 has been fixed with the use of equation (13).

We run 10000 simulations of the mortality rates from ages 0 to age 90. To do that, we have simulate a fractional Brownian motion $B_t^{\hat{H}}$ and using equation (2) we have estimated the mortality rate. To run the fBm simulations we have used the function **fbm** which is includes in the R library **somebm**. We also include a 95.5% confidence interval.

We present the results for women and men in sections 4.2 and 4.3, respectively.

Estimations and predictions were performed using R Ver. 3.2.3 (R Core Team, 2015), and specialized packages Fractal (Time Series Modeling and Analysis Version 2.0-1, 2016), Pracma (Practical Numerical Math Functions 2.0.7, 2017) and somebm (some Brownian motions simulation function Version 0.1, 2016).

4.1 Hurst estimation

For the Hurst parameter estimate we have used three R routines: FDWhittle, RoverS and hurstexp. The two first routines are from the fractal library, while the latter is from the pracma library.

The former routine estimate the Hurst parameter by Whittle's method as was described in the subsection 3.3. RoverS routine estimate $\hat{H}$ by rescaled range (R/S) method. The hurstexp routine estimate $\hat{H}$ using R/S analysis.

Finally, figures 1 and 2 show the estimated Hurst parameter for women and men separately.

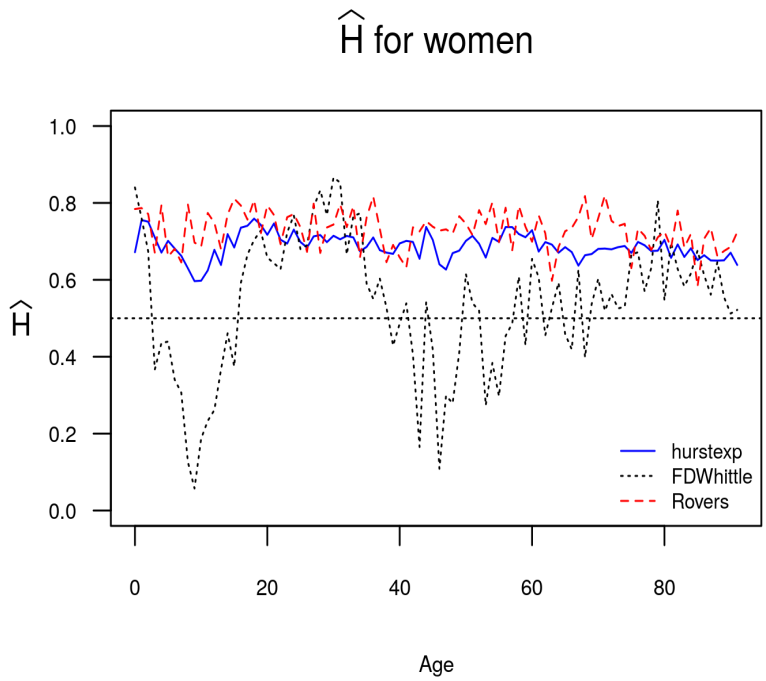

Fig. 1 Estimated Hurst parameter using R-routines.

With the rescaled range R/S and R/S methods we obtain a consistent estimator for the Hurst parameter in the sense that they do not present dramatic changes through the time. Moreover, the H estimated with these two methods take values in the interval (0.57,0.80) approximately. This tells us that the data has the long memory property as was mentioned in section 2. Same results are obtained for the men and women.

Notice that the estimated parameters for H using Whittle method have high variation through the time, in opposition to those obtained with the other two methods so that the estimated Hurst parameters using this method do not perform well in the simulations.

The high variation on the Hurst estimated values could be explained because Whittle method uses the periodogram to estimate H while the other two methods uses the raw data. The two approaches using Whittle likelihood and raw data are very different and hence they give very different estimates of

H. Since rescaled range R/S and R/S methods have estimated very similar H, we decide to use the Hurst coefficients obtained with the method of R/S to perform the mortality rates simulations.

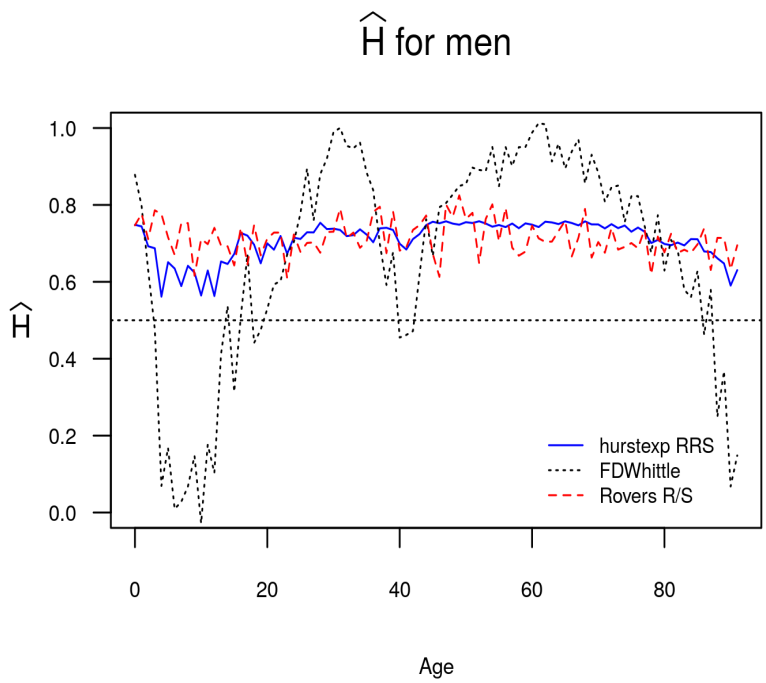

Fig. 2 Estimated Hurst parameter using R-routines.

4.2 Results for women

We present the results for 10000 simulations of the mortality rates for ages: $0, 5, 25, 50, 60, 70, 80, 90$. We graph the historical mortality rate, the mean of all simulations and the 95% confidence interval. See figures 3 and 4.

In general, for all ages, the model is well fitted, in particular, after the 80's. Nevertheless, there are some time periods where the model is not so good as we want to. For instance, the model underestimates the mortality rate for women age 0 during the period of 50s to 80s and for women age 50, 60, 80 and 90

during 60s to 80s approximately. Moreover, it also overestimates the mortality rate for women age 25 during the period of 50s to 80s.

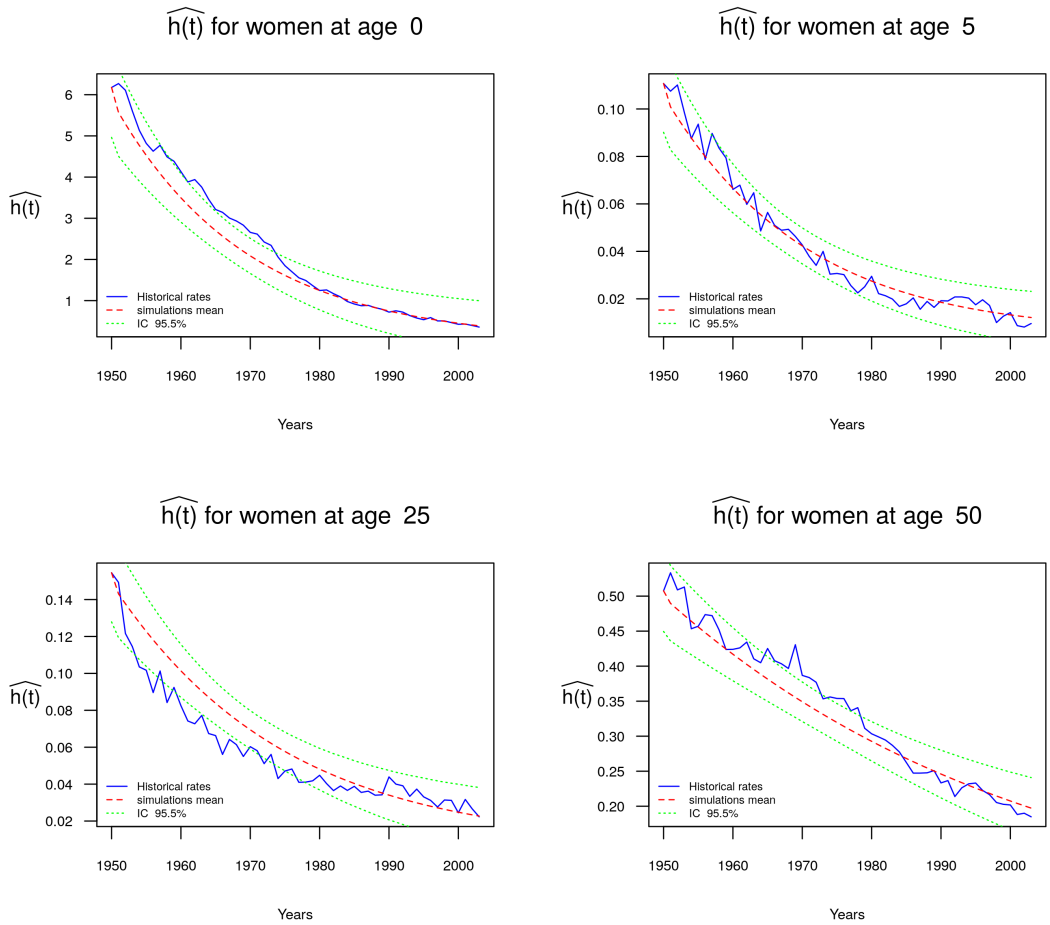

Fig. 3 Simulations for the rate mortality with the fOU model: ages $0, 5, 25, 50$ and $N = 10000$.

For older ages (see figure 4) we observe that for ages 60 and 70 the estimation is well fitted trough the years. We notice that predicted rates are not so far away and that the historical rates are inside the confidence interval. For the very oldest ages the estimation is not so good as for earlier ages. The main difference is in 50's when the absolute number of living persons arriving to those ages were small so that the variability of the estimates is larger.

All these suggest that a better model could include a short and a long-term memory process, so that the model could help us to control the short-term variations in a better way.

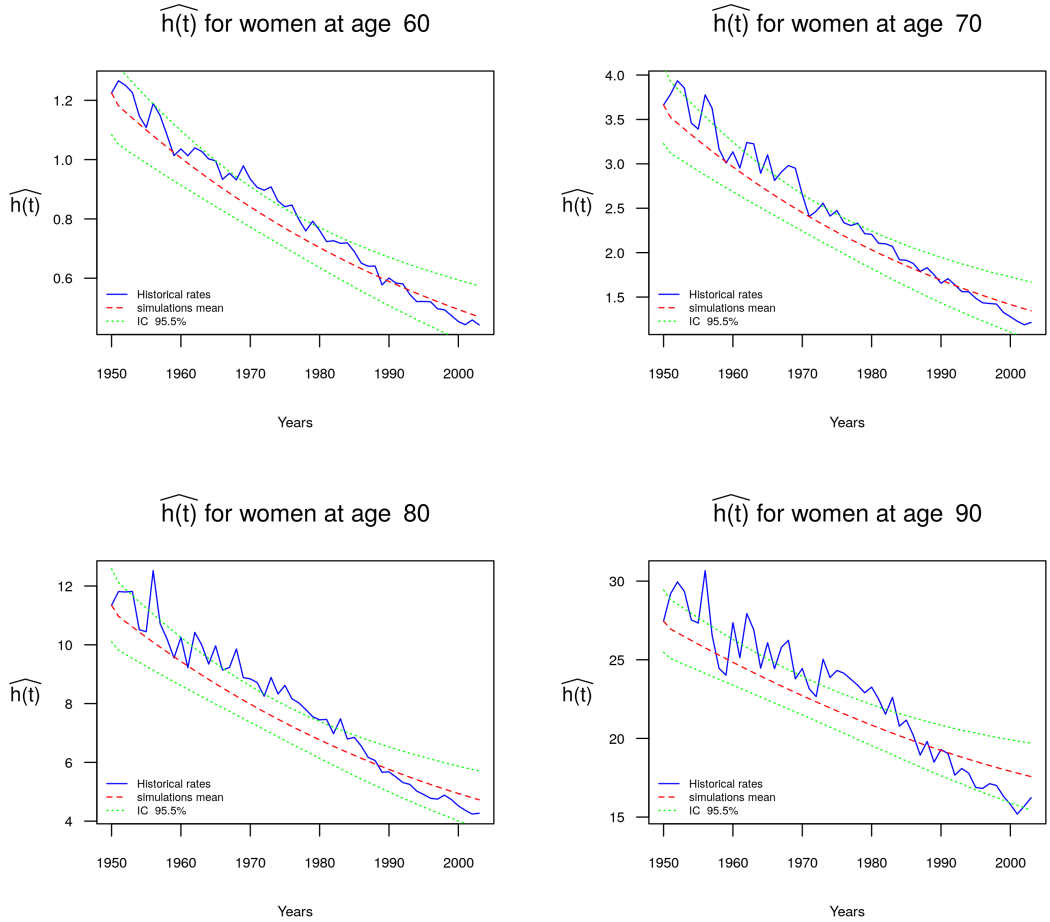

Fig. 4 Simulations for the rate mortality with the fOU model: ages $60, 70, 80, 90$ **and** $N = 10000$**.**

4.3 Results for men

As in the case for women, we present results for 10000 simulations of the mortality rates for ages: $0, 5, 25, 50, 60, 70, 80, 90$. We graph the historical rate

mortality, the mean of all simulations and the 95.5% confidence interval. See figures 5 and 6.

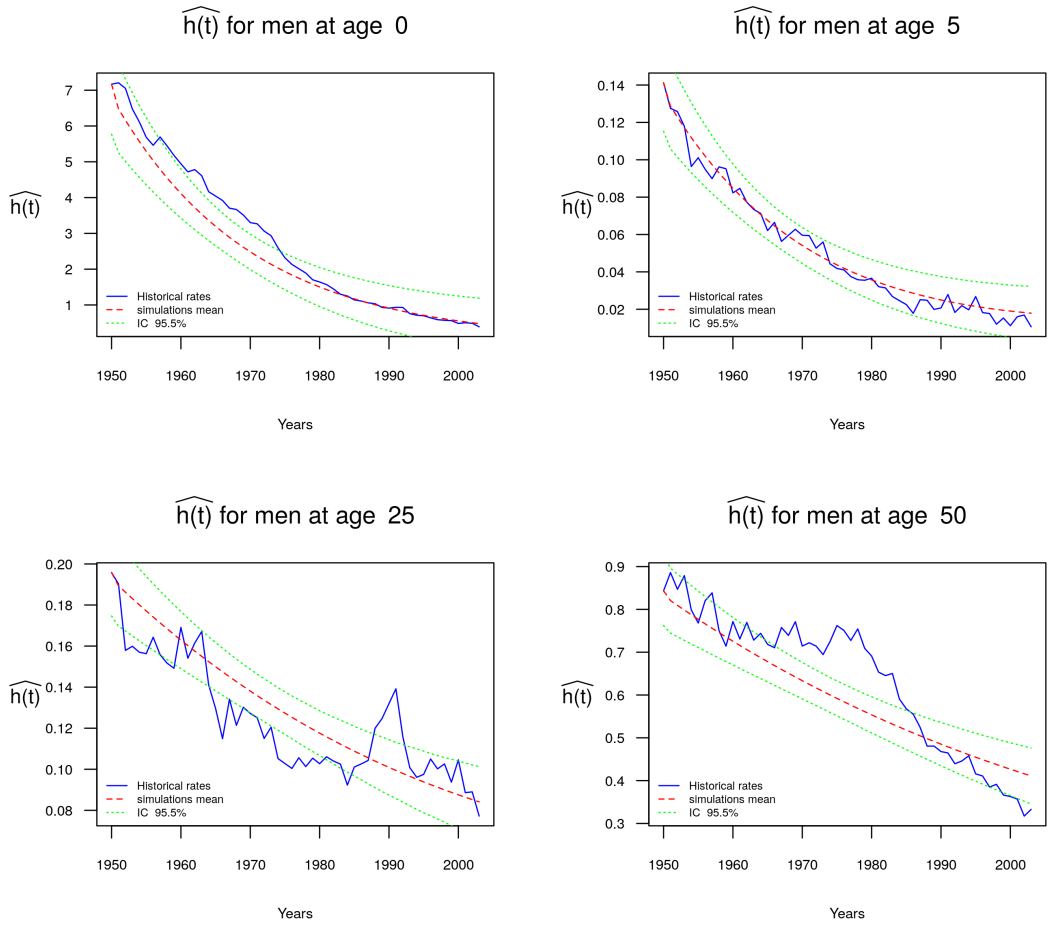

Fig. 5 Simulations for the rate mortality with the fOU model: ages $0, 5$ **and** $N = 10000$**.**

As in the case for women, the proposed model for men is well fitted. We observe an increase in the rates mortality at ages between 25 to 35, this caused a overestimation in the first 35 years and latter a underestimation of the mortality rates. As was mentioned before, if we include in the model a short-term process, we believe the model could be better fitted. The main example of a short-term process to try is a AR(p) with $p \leq 2$ or 3.

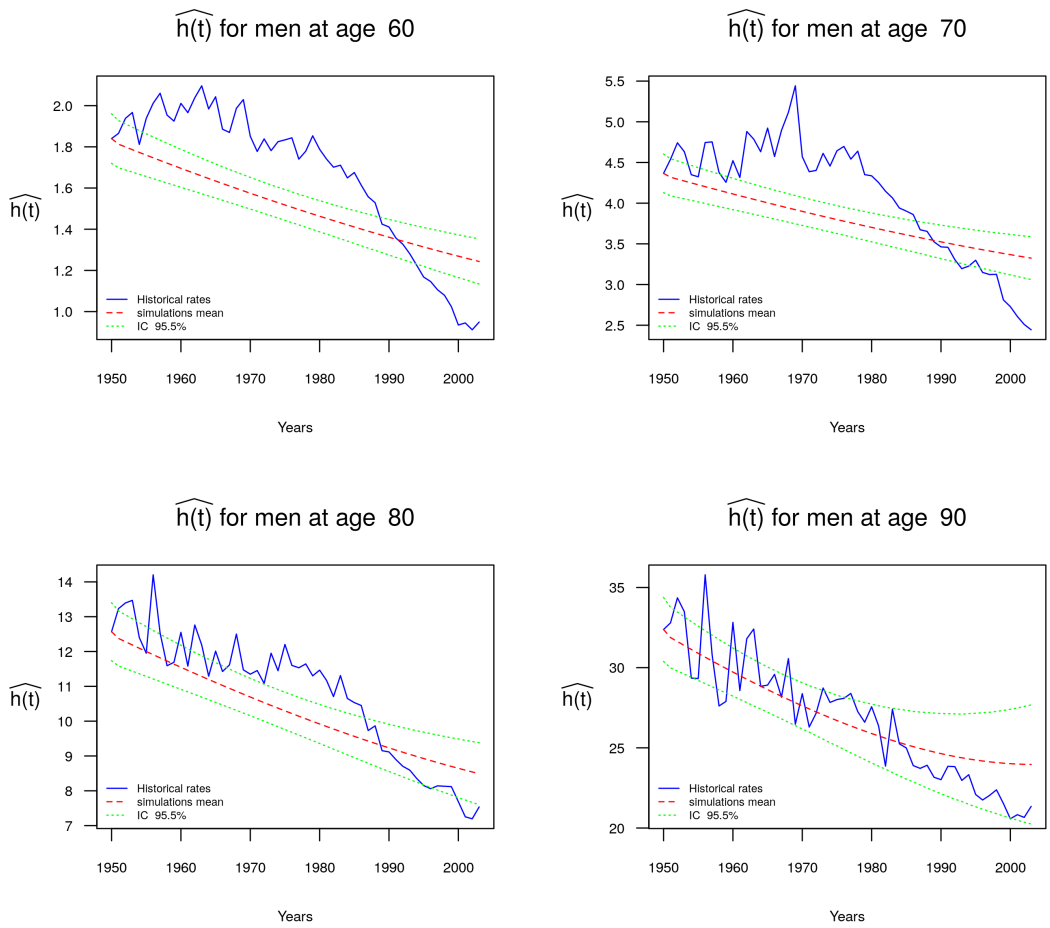

Fig. 6 Simulations for the rate mortality with the fOU model: ages $60, 70, 80, 90$ **and** $N = 10000$**.**

For older ages, we observed that the data is irregular, so it is necessary to use a more complex model to fit this data.

5 Forecast

When we use our model to forecast and compare with the real data between the years 2005 to 2014. In general, from our results we observe that the forecast for women are good for almost all ages we tested. For men the variability of the results is strong and the behavior of the forecast is in general not good as those for women; for instance for ages smaller than 10 the results are quite similar

than for women, even for ages between 10 and 45 it is possible to consider the results just good. However, for ages greater than 50, the results are not good, in fact, for older ages the results are bad: the model overstimate the mortality rates.

We present the results in the figures 7 for women at ages $0, 25, 50, 90$. For men we present more ages to ilustrate that the results are bad as the ages increases. see figures 8 and 9 at ages $0, 25, 50, 90$.

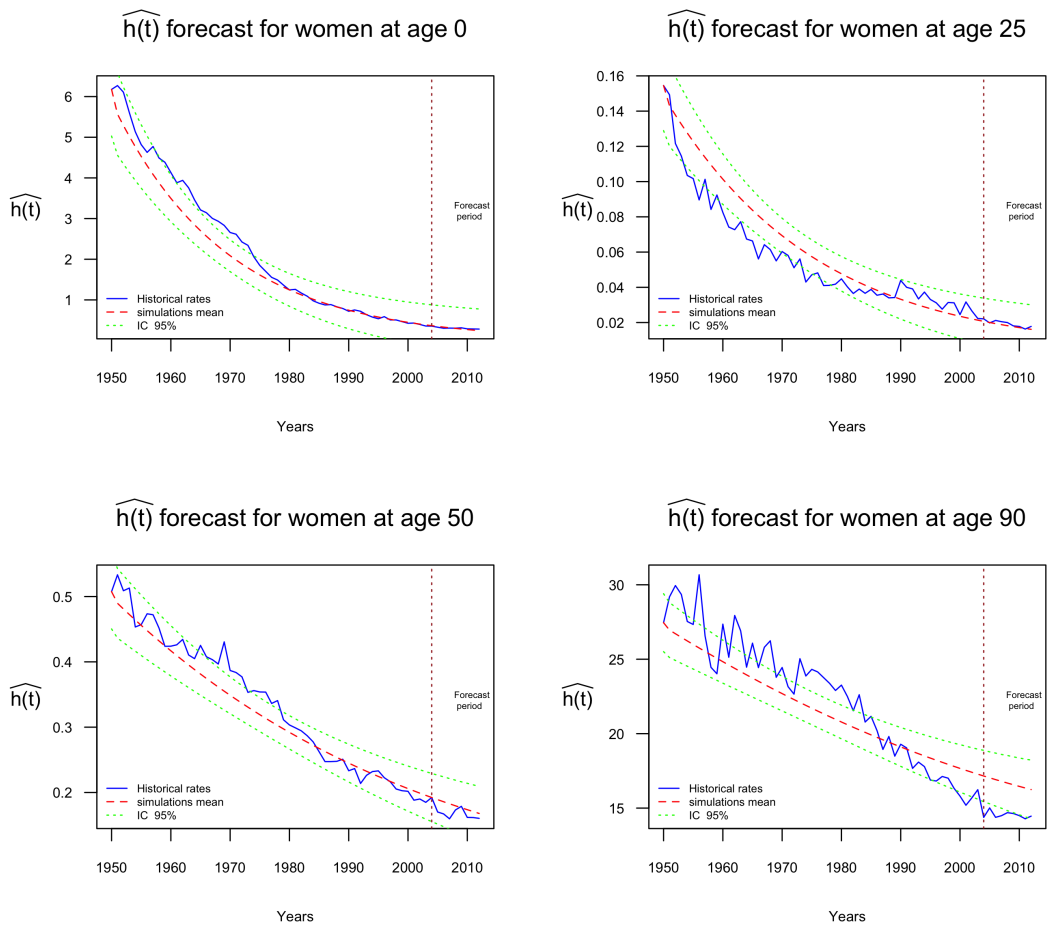

Fig. 7 Forecast for the rate mortality with the fOU model. Women at ages $0, 25, 50, 90$.

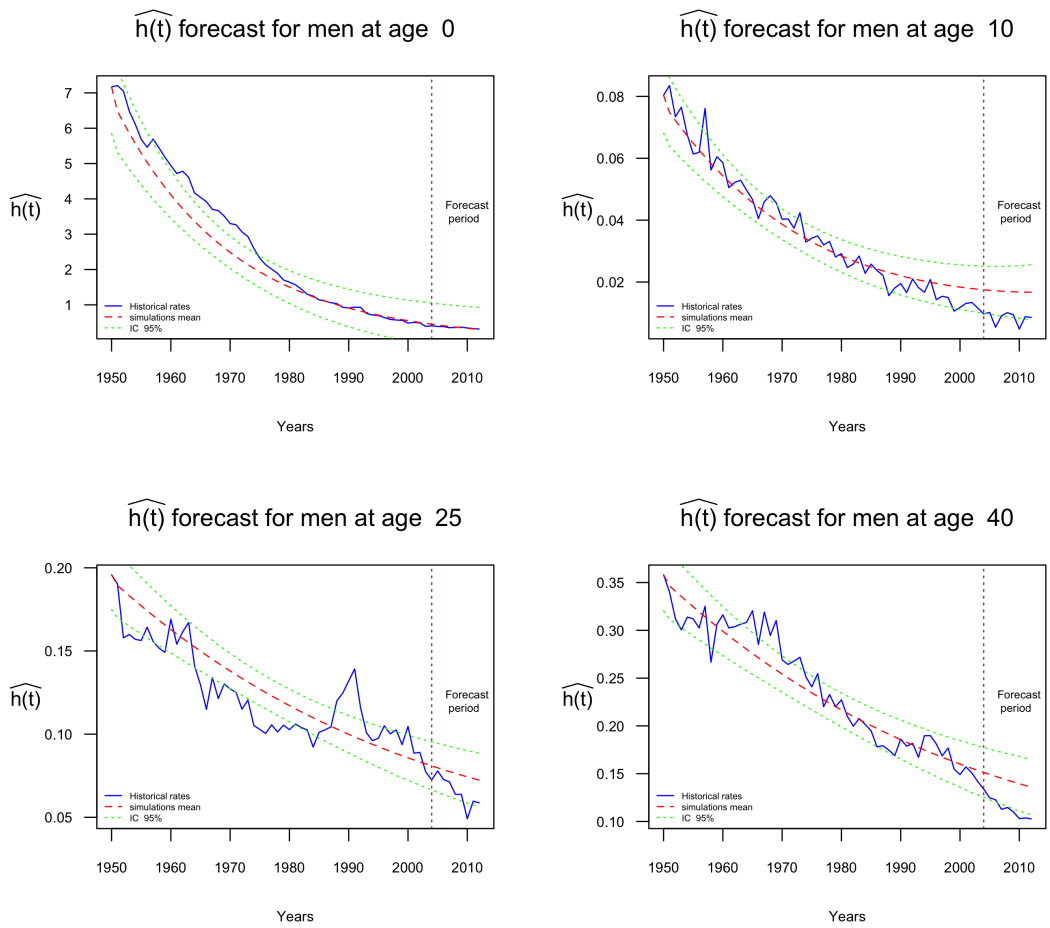

Fig. 8 Forecast for the rate mortality with the fOU model. Men at ages $0, 10, 25, 40$.

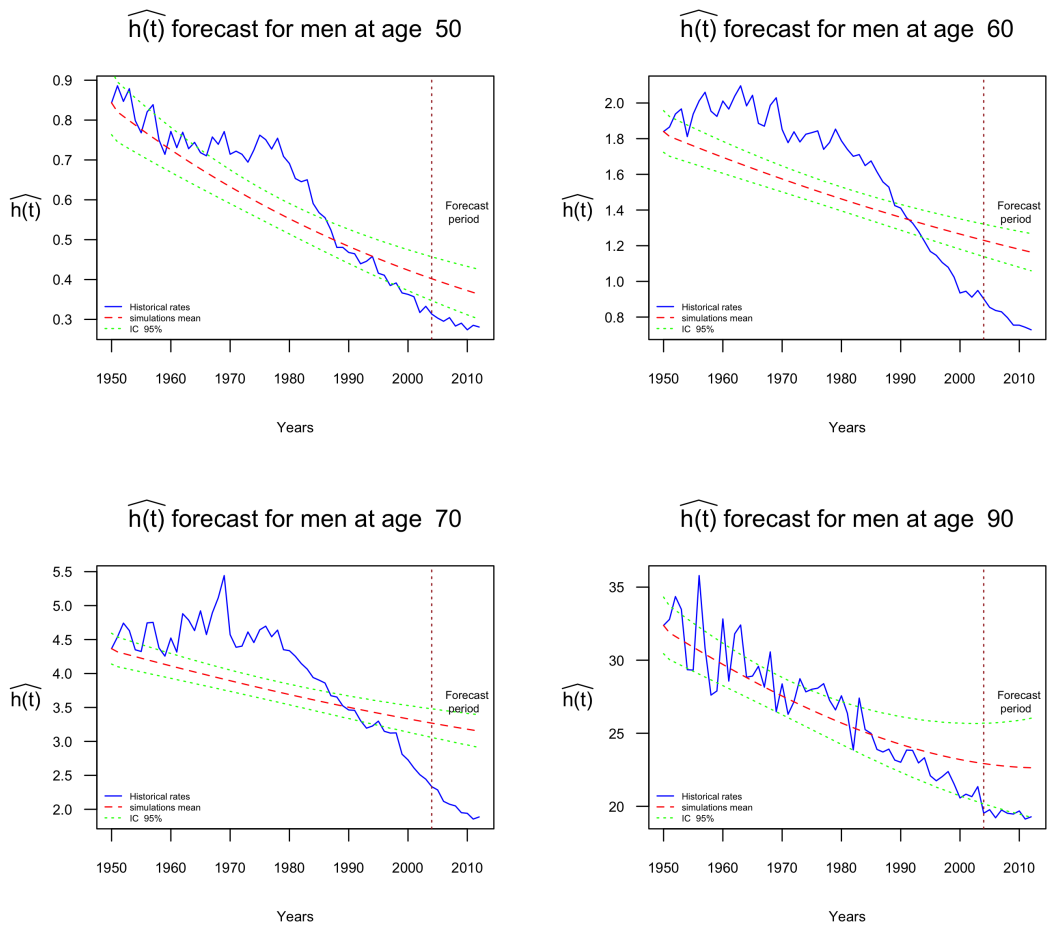

Fig. 9 Forecast for the rate mortality with the fOU model. Men at ages $50, 60, 70, 90$.

6 Conclusions

We have applied our proposed model to the Italian mortality rates with a geometric-type fractional Ornstein-Uhlenbeck process. Our main hypothesis was that, for a fixed age, the mortality rates changes through the time slowly, so that a stochastic differential equations that captures the long-range dependence could be a good model. We have used a stochastic differential equation

with a fractional Brownian motion as a driven noise with $H \in (0.5, 1)$ in order to satisfy the long-range dependence property. With the data we have fixed the Hurst coefficient and we have confirmed our hypothesis since we have found that the estimated Hurst is in $(0.58, 0.8)$.

Notice that we have consider a more general model that the one used in [6]. This is because we have included the possibility that the Hurst parameter could be equal to $1/2$, which is the case when the fractional Brownian motion becomes a standard Brownian motion. Therefore, when $H = 1/2$ we recover the Giacometti, Ortobelli and Bertocchi model.

The model is, specially for women, well behaved. For men at some ages we found some shortcomings that suggest the use of more terms in order to improve the model. The long-range dependence model proposed in this paper is good enough to reproduce the mortality rates. If we add some extra terms to make it more flexible to reproduce the cases where the mortality rates have more variations then it will generate a more accurate model. We are starting to work on this extension of the model. Moreover, a multiplicative noise model will be the subject of a future research.

References

1. J. Beran, Y. Feng, S. Ghosh, R. Kulik: Long-Memory Processes. Probabilistic Properties and Statistical Methods. Springer. (2013)
2. A. Brouste, S. M. Iacus: *Parameter estimation for the discretely observed fractional Ornstein-Uhlenbeck process and the Yuima R package.* Computational Statistics August 2013, Volume 28, Issue 4, pp 1529-1547. (2013).
3. P. Cheridito, H. Kawaguchi, M. Maejima *Fractional Ornstein-Uhlenbeck processes.* Electron J Probab 8(3):1-14, (2003).
4. I. Daubechies: Ten Lectures on Wavelets, SIAM (1992)
5. R.M. Dudley, R. Norvaiša: Concrete functional calculus, in: Springer Monographs in Mathematics, Springer, New York, 2011.
6. R. Giacometti, S. Ortobelli, M. Bertocchi:*A Stochastic Model for Mortality Rate on Italian Data* Journal of Optimization Theory and Applications 149 (1), 216-228. (2011).
7. Y. Hu, D. Nualart : *Parameter estimation for fractional Ornstein–Uhlenbeck processes.* Stat Probab Lett 80(11–12):1030–1038. (2010).
8. H. E. Hurst: *Long-term storage capacity in reservoirs.* Trans. Amer. Soc. Civil Eng. 116, 400-410. (1951).
9. J. Istas, G. Lang : *Quadratic variations and estimation of the local Hölder index of a Gaussian process.* Annales de l'Institut Henri Poincaré 23(4):407–436. (1997).
10. P. Jevtić, E. Luciano, E. Vigna: *Mortality surface by means of continuous time cohort models.* Insurance:Mathematics and Economics. Volume 53, Issue 1, July 2013, Pages 122–133. (2013).
11. P. E. Kloeden, E. Platen: Numerical Solution of Stochastic Differential Equations. Stochastic Modelling and Applied Probability, Vol. 23, Springer (1992)
12. A. N. Kolmogorov: *Wienersche Spiralen und einige andere interessante Kurven im Hilbertschen Raum.* C. R. (Doklady) Acad. URSS (N.S.) 26, 115-118. (1940).
13. D. Koutsoyiannis: *The Hurst phenomenon and fractional Gaussian noise made easy.* Hydrological Sciences Journal des Sciences Hydrologiques, 47(4). August (2002)

14. K. Kubilius, Y. S. Mishura: *The rate of convergence of Hurst index estimate for the stochastic differential equation* Stochastic Processes and their Applications. Vol. 122, Issue 11, 3718–3739. (2012)
15. A. W. Lo: *Long-Term Memory in Stock Market Prices* Econometrica Vol. 59, No. 5 (Sep., 1991), pp. 1279-1313. (1991).
16. B. B. Mandelbrot, J. W. Van Ness: *Fractional Brownian motions, fractional noises and applications.* SIAM Review 10, 422-437. (1968).
17. T. Mikosch : Elementary Stochastic Calculus With Finance in View. Advanced Series on Statistical Science and Applied Probability, Vol 6. World Scientific Publishing (1999)
18. M.A. Milevsky, S.D. Promislow: *Mortality derivatives and the option to annuitise.* Insur. Math. Econ. 29, 299–318 (2001)
19. Y. S. Mishura: Stochastic Calculus for Fractional Brownian Motion and Related Processes. Springer, (2008).
20. A. Neuenkirch, S. Tindel: *A least square-type procedure for parameter estimation in stochastic differential equations with additive fractional noise.* Stat Inference Stoch Process, vol. 17, 99–120, (2014).
21. D. Nualart: *Fractional Brownian motion: stochastic calculus and applications* Proceedings of the International Congress of Mathematicians, Madrid, Spain, (2006)
22. B.K. Öksendal: Stochastic Differential Equations: An Introduction with Applications. Berlin: Springer. (2010).
23. C. Park, F. Hernandez-Campos, L. Long, J. Marron, J. Park, V. Pipiras, F. Smith, R. Smith, M. Trovero, Z. Zhu: *Long range dependence analysis of internet traffic.* J. Appl. Stat. 38(7), 1407-1433 (2011)
24. M. Priestley Spectral Analysis and Time Series. Academic Press (1982)
25. B. L. S. Prakasa Rao: Statistical Inference for Fractional Diffusion Processes. Wiley Series in Probability and Statistics. (2010)
26. A. Rossa, L. Socha: *Proposition of a Hybrid Stochastic Lee-Carter Mortality Model.* Advances in Methodology & Statistics/Metodoloski zvezki 10.1 (2013).
27. R. H. Shumway, D. S. Stoffer: *Time Series Analysis and Its Applications: With R Examples.* Springer Texts in Statistics. 3rd. ed. (2011).
28. L.C. Young: *An inequality of the Hölder type, connected with Stieltjes integration*, Acta Mathematica 67 251–282. (1936).
29. R. Weron:*Estimating long-range dependence: finite sample properties and confidence intervals.* Physica A 312 (2002) 285 - 299
30. F. Yerlikaya-Ozkurt, C. Vardar-Acar, Y. Yolcu-Okur, G.-W. Weber: *Estimation of the Hurst parameter for fractional Brownian motion using the CMARS method.* J. Comput. Appl. Math., 259, pp. 843-850. (2014).
31. C: Zeng, Y. Q. Chen, Q. Yang: *The fBm-driven Ornstein-Uhlenbeck process: Probability density function and anomalous diffusion.* Fractional Calculus and Applied Analysis September 2012, Volume 15, Issue 3, pp 479-492. (2012).