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

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

Received: April 15, 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 based on 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 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 next formulations.

Francisco Delgado-Vences
Conacyt-Universidad Nacional Autónoma de Ḿexico.
Instituto de Matemáticas, Oaxaca, México
E-mail: delgado@im.unam.mx
ORCID:

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

Saul Diaz-Infante
Conacyt-Universidad de Sonora
Departamento de Matemáticas,
Hermosillo, Sonoran México E-mail: sdinfante@conacyt.mx
ORCID: 0000-0001-9559-1293

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. Mortality impacts directly in cash and therefore need a reliable future projection.

Previous work has only focused on deterministic or stochastic models without memory effects. Pitacco et al. (2009) review the first mortality tables and models. Milevsky and Promislow (2001) report a linear SDE driven by Brownian motion that describes the mortality hazard rate. In Giacometti et al. (2011), the authors extend the Mikevesky model to an SDE with time-dependent diffusion and study a type of autoregressive model for the logarithm of the hazard rate. Jevtić et al. (2013) formulate a cohort-based model with imperfect correlation across generations and use data from UK to estimate parameters.

The present paper aims to show statistical evidence that mortality rates follow a stochastic process with long-range dependence (LRD), that is, a stochastic process with memory. According to Prakasa Rao (2010), a stochastic process or time series is LRD, if it has persistence behavior—below we give a formal definition.

We base our stochastic formulation in the fractional Brownian Motion (fBM) Mandelbrot and Van Ness (1968). Since fBM is a generalization of the standard Brownian Motion (BM) that still satisfies self-similarity and is LRD, fBM results to be a natural noise model to describe LRD behavior.

The main idea is to extend the stochastic model reported by Milevesky and Promislow to a model with LRD and verify its performance to fitting and forecasting with real data.

We obtain statistical evidence—via Hurst parameter estimation—that Italy mortality rate data is LRD. Our model captures women's mortality rate dynamics. But we observe that the Italian man has a higher mortality rate and variance. These results suggest that stratification by gender would be a direction for future formulations.

Our work exhorts to explore the long time memory effects of dynamics under uncertainty. Also these ideas could improve mathematical models, which considers mortality in its formulation, for example, models for infectious disease spread or population growth.

After this brief introduction, we aformulate our fractional model in Section 2. Section 3 reviews the fBM and presents the fractional Ornstein-Uhlenbeck (fOU) process. Section 4 outlines the applied method for parameter estimation. In Section 5, we implement our formulation to fitting data of mortality from Italy. Further, in this section, we also run forecasting to a subset of the data and evaluate its efficiency. Finally, we conclude in Section 7.

2 Model formulation

We now discuss the model of Milevsky-Promislow Model (see Milevsky and Promislow (2001) for the original paper or Giacometti et al. (2011) 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 $h_0, \alpha_1, \alpha_2 > 0$. 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 $\sigma, \lambda > 0$.

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 (3), which cover the case of long range dependence of the data.

The model we propose is a generalization in the following sense. As 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).

In this paper, we will assume that Y_t is an stochastic process that satisfies the SDE:

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

where B_t^H is a *fractional Brownian motion* (see subsection 3.1) with Hurst parameter $1/2 \leq H < 1$, $Y_0 = 0$, and $\sigma, \lambda > 0$. This SDE is the fractional Ornstein-Uhlenbeck process (see subsection 3.2 for a formal definition of this process).

The main difference between this model and the original presented in Milevsky and Promislow (2001) 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. It is well-known that the fractional Ornstein-Uhlenbeck process (fOU process) Y_t^H is a long memory process when $1/2 \leq H < 1$ (see subsection 3.2). Therefore, our proposed model is able to capture a long-range dependence of the data.

We interpret the SDE (4) as

$$Y_t^H = -\lambda \int_0^t Y_s^H ds + \sigma B_t^H. \tag{5}$$

The equation above lacks of a stochastic integral because we consider the case with additive noise. We refer the reader to Mishura (2008) for further reading on fractional stochastic calculus.

3 Fractional Gaussian processes

3.1 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

$$R_H(t,s) := \mathbb{E}(B_s^H B_t^H) = \tfrac{1}{2}\big(t^{2H} + s^{2H} - |t-s|^{2H}\big). \tag{6}$$

This stochastic process is called a *fractional Brownian motion* (fBm) and was introduced by Kolmogoroff Kolmogoroff (1940) and studied by Mandelbrot and Van Ness in Mandelbrot and Van Ness (1968). The parameter H is called Hurst index because of the statistical analysis developed by the climatologist Hurst Hurst (1951). The fBm is a generalization of Brownian motion without independent increments, in fact it is a continuous-time Gaussian process.

fBm process has the properties of self-similarity and stationary increments. Its sample-paths are almost nowhere differentiable but almost-all Hölder continuous for any order strictly less than H. That is, for each trajectory, there exists a finite constant C such that for every $\epsilon > 0$

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

For $H = \frac{1}{2}$ the covariance of fBM can be written as $R_{1/2}(t,s) = \min(s,t)$, thus process $B_t^{1/2}$, is equivalent to the standard Brownian motion. However, if $H \neq \frac{1}{2}$, then this increments are *not* independent. Let $X_n = B_n^H - B_{n-1}^H$, $n \geq 1$. Then $\{X_n, n \geq 1\}$ is a Gaussian stationary sequence with unit variance and covariance function if

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

Therefore,

- if $H > \frac{1}{2}$, then $\rho_H(n) > 0$ for n large enough and $\sum_{n=1}^{\infty} \rho_H(n) = \infty$. In this case, we say that process X_n is persistent with positive correlation and that X_n has the *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 Prakasa Rao (2010); Nualart (2006); Mishura (2008).

3.2 Fractional Ornstein-Uhlenbeck process (fOU)

The fOU is an SDE driven by a fractional Brownian motion. Coming back to Equation (5), Cheridito et al. (2003) introduced the fractional Ornstein-Uhlenbeck process (fOU) and they showed that the process

$$Y_t^H = \sigma \int_0^t e^{-\lambda(t-u)} dB_u^H, \tag{7}$$

is the unique almost surely continuous-path process which solves (5) (see also Theorem 1.24 in Prakasa Rao (2010)). The integral in Equation (7) is a path-wise Riemann–Stieltjes integral. The fOU process is neither Markovian nor a semimartingale for $H \in (1/2, 1)$ (see Dudley and Norvaiša, 2011) but remains Gaussian and ergodic.

Moreover, when $H \in (1/2, 1)$, Y_t even presents the long-range dependence property Cheridito et al. (2003); Prakasa Rao (2010).

According to Zeng et al. (2012), the variance of the fOU process Y_t is given by the following expression:

$$\operatorname{Var}(Y_t) = \sigma^2 2H e^{-2\lambda t} \int_0^t s^{2H-1} e^{2\lambda s} ds. \tag{8}$$

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

$$\operatorname{Var}(Y_t) = \frac{\sigma^2}{2\lambda}\left(1 - e^{-2\lambda t}\right), \tag{9}$$

which is the variance of the standard Ornstein-Uhlenbeck process (see Mikosch, 1998, p. 143).

If we consider the constant $\alpha_1 = T^{-H}$ and Equation (8), then the variance of $\alpha_1 Y_t$ is given by

$$\begin{aligned}\operatorname{Var}(\alpha_1 Y_t) &= \alpha_1^2 \operatorname{Var}(Y_t) = \alpha_1^2 \sigma^2 2H \int_0^t s^{2H-1} e^{-2\lambda(t-s)} ds \\ &\leq \alpha_1^2 \sigma^2 2H \int_0^t s^{2H-1} ds = \alpha_1^2 \sigma^2 2H \frac{s^{2H}}{2H}\Big|_{s=0}^{t} \\ &= \alpha_1^2 \sigma^2 t^{2H} = \sigma^2 (t/T)^{2H},\end{aligned} \tag{10}$$

then $\operatorname{Var}(\alpha_1 Y_t) \leq \sigma^2$. Thus the variance of $\alpha_1 Y_t$ is bounded by a constant indepedent of time. We will use α_1 to control the variance of process Y_t.

4 Estimation of the parameters

In this section we describe a methodology to estimate the parameters. According to (2), we need to estimate α_0, α_1 and for SDE model (3) σ and λ. Furthermore, Hurst parameter H involved in the fractional Brownian motion is also estimated. To estimante this Hurst parameter we appeal to empirical evidence that H in equation (4) is invariant, that is, the value of the Hurst parameters H, in the equation (4), for the fBm B_t^H and the one for the fractional Gaussian noise Y_t^H are the same. Further, to estimate α_1, we assume that $\alpha_1 = T^{-H}$.

4.1 Estimation of the parameter α_0.

According to equation (2), to estimate α_0 we assume $\alpha_1 = 1$. Moreover, h_0 is the initial value of the observed rates $h(t)$. Thus, from (2) we obtain

$$\ln h(t) = \ln h_0 + \alpha_0 t + Y_t. \tag{11}$$

To estimate parameter α_0 we minimize the sum of the square errors. Let

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

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 deduce that $\widehat{\alpha_0}$ satisfies:

$$\widehat{\alpha_0} = \frac{\sum_t t \ln h(t) - \ln h(0) \sum_t t}{\sum_t t^2}. \tag{12}$$

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

4.2 The relation between the Hurst parameter and the H-index in the FOU

In this subsection, we outline a method to estimate the parameter H. We use the following empirical fact. Suppose that an fOU process is driving with an fBM with a given Hurst parameter H_0. Yerlikaya-Özkurt et al. (2014) provide 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. They provide statistical evidence that the fOU should have the same value H0 that the fBM. Thus, under this empirical evidence, we assume the same value of the parameter H for both processes. A formal result of this fact, up to our knowledge, still is an open question.

The following sections present several methods to estimate the Hurst parameter for the fBM.

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

The last subsection allows us to estimate the parameter H as follows. According to equation (11), the residuals are

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

Note taht $\hat{Y}_t$ is a fractional Gaussian process that we can use to estimate H. Afterwards we use estimation $\hat{H}$ to approximate the Hurst parameters of the fractional Brownian motion B_t^H. To this end, we review some methods to estimate the parameter H, our main refernces are Weron (2002), .

4.3.1 R / S Analysis

Following Weron (2002), we coniser a finite time serie of lenght L indexed by i Z_i. Thus, let $\{Z_{i,m}\}$, $m = 1, \ldots, d$ be a finite sequence of d series of lenght n, such that $L = n \times d$. Next, 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

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

Accoridng to Mandelbrot (1974/75), the R/S statistic asymptotically follows the relation:

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

where c is a constant. Thus, we can estimate H by 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 scale. 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 the processes is anti-persistent, then the slope H is less than 0.5. The "significance" level of the estimated parameter H usually is chosen to be inverse proportional to the square root of sample length, that is, 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 Lo (1991).

4.3.2 Method of rescaled range analysis R/S

Here we follow Prakasa Rao (2010, Chap. 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 $i = 1, \dots, K$. Define

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

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

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

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.

4.3.3 FDWhittle Estimator

FollowingPark et al. (2011). The Local Whittle Estimator (LWE) is a semi-parametric Hurst parameter estimator based on the periodogram. LWE assumes that the spectral density $f(\omega)$ of the process can be approximated by the function

$$f_{c,H}(\omega) = c\omega^{1-2H}, \tag{13}$$

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, for example, Priestley, 1981, sect. 6.1.2).

The LWE of the Hurst parameter, $\hat{H}_{LWE}(m)$ is implicitly the result of 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 (13).

4.4 Estimation of σ and λ

There are several methods to estimate parameters σ and λ. For instance see Prakasa Rao (2010) or the references in Neuenkirch and Tindel (2014) or in Kubilius and Mishura (2012). In the following section we will do a brief review of some of these methods.

4.4.1 Estimation σ.

Brouste and Iacus (2013) 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 parameter λ —the results are valid only when $1/2 < H < 3/4$.

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

Let $\boldsymbol{a} = (a_0, \dots, a_K)$ 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 $Y^T = (Y_t : 0)$ be the sample path of the solution of (7). Consider the discretization of Y^T

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

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 $\boldsymbol{a}$ Istas and Lang (see, for example 1997). Then, define the following estimators for H and σ.

$$\hat{H}_N := \tfrac{1}{2} \log_2 \left(\frac{V_{N,\boldsymbol{a}^2}}{V_{N,\boldsymbol{a}}} \right), \tag{14}$$

$$\hat{\sigma}_N := \left(-2 \frac{V_{N,\boldsymbol{a}}}{\sum_{k,l} a_k a_l |k-l|^{2\hat{H}_N} \Delta_N^{2\hat{H}_N}} \right)^{1/2} . \tag{15}$$

Brouste and Iacus (2013) proved the following 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, that is,*

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

Moreover, for all $H \in (0,1)$, if $N \to \infty$, then

$$\sqrt{N}(\hat{H}_N - H) \stackrel{\mathcal{L}}{\to} N(0, \Gamma_1(\boldsymbol{a}, \sigma, H)),$$
$$\frac{\sqrt{N}}{\log N}(\hat{\sigma}_N - \sigma) \stackrel{\mathcal{L}}{\to} N(0, \Gamma_2(\boldsymbol{a}, \sigma, H))$$

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

With this result we obtain an estimator for parameterσ. Consier the following filters.

- Classical filter. Let $K > 0$ and define

$$a_k := \frac{(-1)^{1-k}}{2^k} \binom{K}{k} = \frac{(-1)^{1-k}}{2^k} \frac{K!}{k!(K-k)!} \qquad \text{for } 0 \leq k \leq K.$$

- Daubechies filters (see Daubechies, 1992, for original definition). Let $K = 4$ and define the Daubechies filter by

$$\frac{1}{\sqrt{2}} \big(0.482\,962\,913\,144\,53, -0.836\,516\,303\,737\,8,$$
$$0.224\,143\,868\,042\,01, 0.129\,409\,522\,551\,26 \big).$$

4.4.2 Estimation of the drift parameter λ H and σ are known

Hu and Nualart (2010) proved 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

$$\hat{\lambda}_N = \left(\frac{2\hat{\mu}_{2,N}}{\hat{\sigma}_N^2 \Gamma(2\hat{H}_N + 1)} \right)^{-\frac{1}{2\hat{H}_N}} \tag{16}$$

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

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

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 $N \to +\infty$,*

$$\hat{\lambda}_N {\to} \lambda, \ \ a.s.$$

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 (Brouste and Iacus, 2013, Thm. 2).

5 Results

In this section we present our estimated mortality rates according to obtain from Italy—the website of Human Mortality Database for the Italian population between 1950 to 2004.

We present the estimation of the H parameter. In second place, we present the results on simulated mortality rates using equations (15)-(16) to estimate the parameters σ, λ. The parameter α_0 has been fixed with the use of equation (12).

We sample 10 000 paths of the solution process to SDE (5) to estimate mortality rates of woman and man between 0 and 90 year old. To this end, we simulate paths of 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 5.2 and 5.3, 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).

5.1 Hurst parameter 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 fromthe pracma library.

The former routine estimate the Hurst parameter by Whittle's method as was described in the subsection 4.3. 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 5.1 show the estimated Hurst parameter for women and men separately.

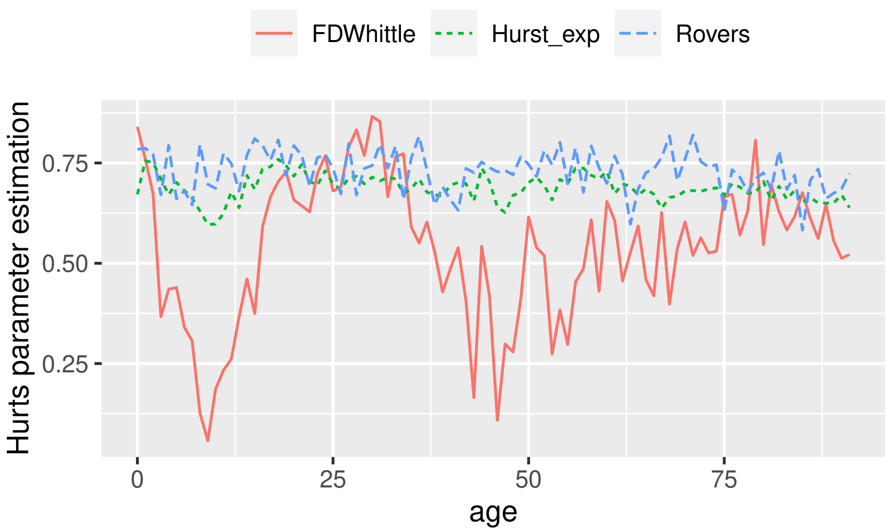

Fig. 1 Estimated Hurst parameter using R-routines.

With the rescaled range `R/S` and `hurst_exp` 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 3. 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.

5.2 Results for women

We present the results for 10 000 path 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 **??**.

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.

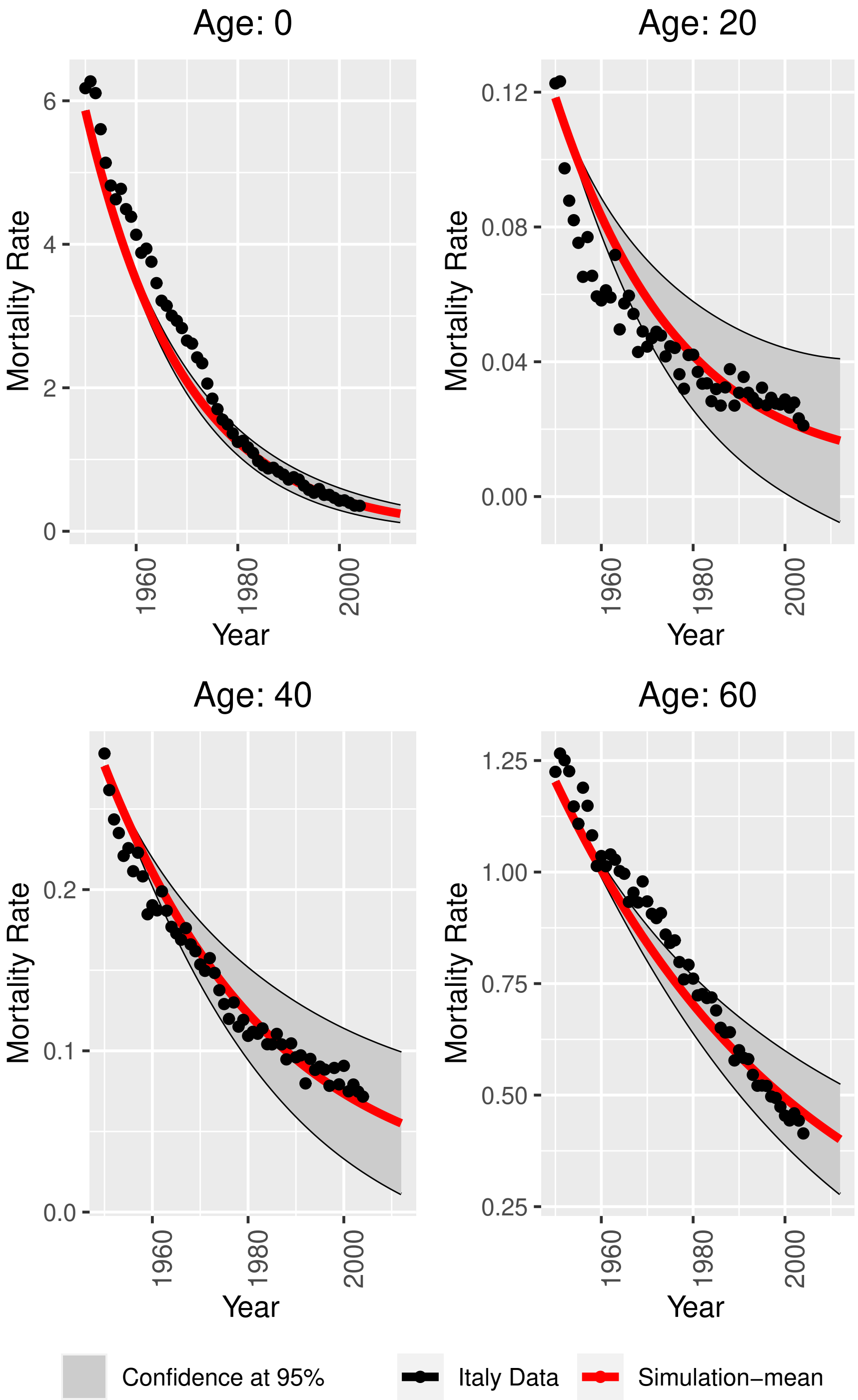

Fig. 3 Simulations for the rate mortality with the fOU model: ages 0, num20, 40, 60 and $N = 10\,000$ simulation paths.

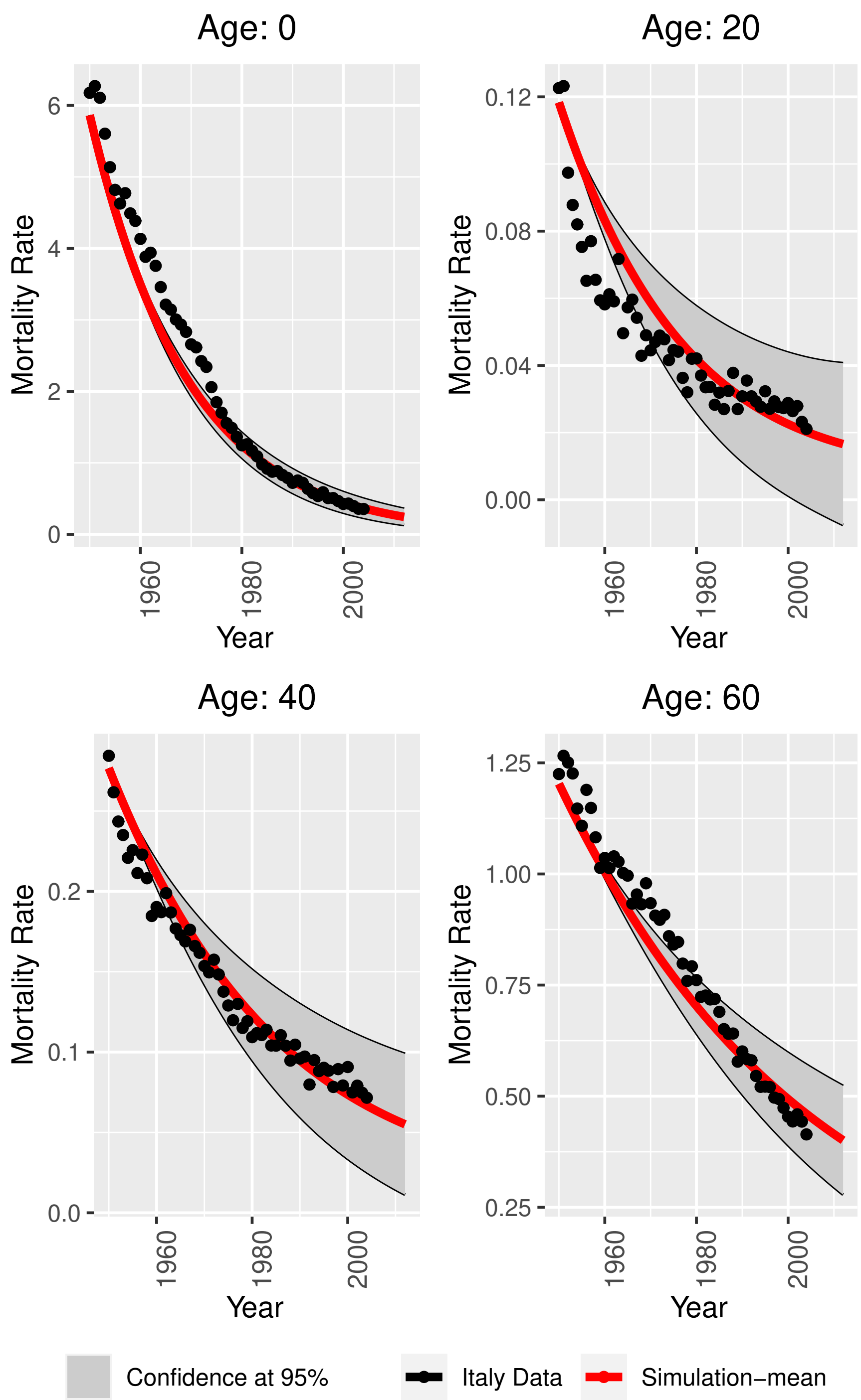

Fig. 4 Simulations for the rate mortality with the fOU model: ages 0, num20, 40, 60 and $N = 10\,000$ simulation paths.

For older ages (see figure **??**) 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.

5.3 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 **??** and **??**.

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 $\mathrm{AR}(p)$ with $p \leq 2$ or 3.

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

6 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 **??** 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 **??** and **??** at ages $0, 25, 50, 90$.

7 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, for all ages, is in $(0.58, 0.8)$.

Notice that we have consider a more general model that the one used in Giacometti et al. (2011). 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. From this results, and in opposition of the European normative on insurance that does not discriminate by gender, we conclude from our model, applied to the Italian case, that modelling the mortality by gender could improve the risk management of the insurance companies.

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. A. Brouste and S. M. Iacus. Parameter estimation for the discretely observed fractional Ornstein-Uhlenbeck process and the Yuima R package. *Comput. Statist.*, 28(4):1529–1547, 2013. ISSN 0943-4062. doi: 10.1007/s00180-012-0365-6. URL `https://doi.org/10.1007/s00180-012-0365-6`.
2. P. Cheridito, H. Kawaguchi, and M. Maejima. Fractional Ornstein-Uhlenbeck processes. *Electron. J. Probab.*, 8:no. 3, 14, 2003. ISSN 1083-6489. doi: 10.1214/EJP.v8-125. URL `https://doi.org/10.1214/EJP.v8-125`.
3. I. Daubechies. *Ten lectures on wavelets*, volume 61 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992. ISBN 0-89871-274-2. doi: 10.1137/1.9781611970104. URL `https://doi.org/10.1137/1.9781611970104`.
4. R. M. Dudley and R. Norvaiša. *Concrete functional calculus*. Springer Monographs in Mathematics. Springer, New York, 2011. ISBN 978-1-4419-6949-1. doi: 10.1007/978-1-4419-6950-7. URL `https://doi.org/10.1007/978-1-4419-6950-7`.

5. R. Giacometti, S. Ortobelli, and M. Bertocchi. A stochastic model for mortality rate on Italian data. *J. Optim. Theory Appl.*, 149(1):216–228, 2011. ISSN 0022-3239. doi: 10.1007/s10957-010-9771-5. URL `https://doi.org/10.1007/s10957-010-9771-5`.
6. Y. Hu and D. Nualart. Parameter estimation for fractional Ornstein-Uhlenbeck processes. *Statist. Probab. Lett.*, 80(11-12):1030–1038, 2010. ISSN 0167-7152. doi: 10.1016/j.spl.2010.02.018. URL `https://doi.org/10.1016/j.spl.2010.02.018`.
7. H. E. Hurst. Long-term storage capacity of reservoirs. *Trans. Amer. Soc. Civil Eng.*, 116:770–799, 1951.
8. J. Istas and G. Lang. Quadratic variations and estimation of the local Hölder index of a Gaussian process. *Ann. Inst. H. Poincaré Probab. Statist.*, 33(4): 407–436, 1997. ISSN 0246-0203. doi: 10.1016/S0246-0203(97)80099-4. URL `https://doi.org/10.1016/S0246-0203(97)80099-4`.
9. P. Jevtić, E. Luciano, and E. Vigna. Mortality surface by means of continuous time cohort models. *Insurance Math. Econom.*, 53(1):122–133, 2013. ISSN 0167-6687. doi: 10.1016/j.insmatheco.2013.04.005. URL `https://doi.org/10.1016/j.insmatheco.2013.04.005`.
10. A. N. Kolmogoroff. Wienersche Spiralen und einige andere interessante Kurven im Hilbertschen Raum. *C. R. (Doklady) Acad. Sci. URSS (N.S.)*, 26: 115–118, 1940.
11. K. Kubilius and Y. Mishura. The rate of convergence of Hurst index estimate for the stochastic differential equation. *Stochastic Process. Appl.*, 122(11): 3718–3739, 2012. ISSN 0304-4149. doi: 10.1016/j.spa.2012.06.011. URL `https://doi.org/10.1016/j.spa.2012.06.011`.
12. A. W. Lo. Long-term memory in stock market prices. *Econometrica*, 59(5): 1279–1313, 1991.
13. B. B. Mandelbrot. Limit theorems on the self-normalized range for weakly and strongly dependent processes. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 31:271–285, 1974/75. doi: 10.1007/BF00532867. URL `https://doi.org/10.1007/BF00532867`.
14. B. B. Mandelbrot and J. W. Van Ness. Fractional Brownian motions, fractional noises and applications. *SIAM Rev.*, 10:422–437, 1968. ISSN 0036-1445. doi: 10.1137/1010093. URL `https://doi.org/10.1137/1010093`.
15. T. Mikosch. *Elementary stochastic calculus—with finance in view*, volume 6 of *Advanced Series on Statistical Science & Applied Probability*. World Scientific Publishing Co., Inc., River Edge, NJ, 1998. ISBN 981-02-3543-7. doi: 10.1142/9789812386335. URL `https://doi.org/10.1142/9789812386335`.
16. M. A. Milevsky and S. D. Promislow. Mortality derivatives and the option to annuitise. volume 29, pages 299–318. 2001. doi: 10.1016/S0167-6687(01)00093-2. URL `https://doi.org/10.1016/S0167-6687(01)00093-2`. 4th IME Conference (Barcelona, 2000).
17. Y. S. Mishura. *Stochastic calculus for fractional Brownian motion and related processes*, volume 1929 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2008. ISBN 978-3-540-75872-3. doi: 10.1007/978-3-540-75873-0. URL `https://doi.org/10.1007/978-3-540-75873-0`.

18. A. Neuenkirch and S. Tindel. A least square-type procedure for parameter estimation in stochastic differential equations with additive fractional noise. *Stat. Inference Stoch. Process.*, 17(1):99–120, 2014. ISSN 1387-0874. doi: 10.1007/s11203-013-9084-z. URL `https://doi.org/10.1007/s11203-013-9084-z`.
19. D. Nualart. Fractional Brownian motion: stochastic calculus and applications. In *International Congress of Mathematicians. Vol. III*, pages 1541–1562. Eur. Math. Soc., Zürich, 2006.
20. C. Park, F. Hernández-Campos, L. Le, J. S. Marron, J. Park, V. Pipiras, F. D. Smith, R. L. Smith, M. Trovero, and Z. Zhu. Long-range dependence analysis of Internet traffic. *J. Appl. Stat.*, 38(7):1407–1433, 2011. ISSN 0266-4763. doi: 10.1080/02664763.2010.505949. URL `https://doi.org/10.1080/02664763.2010.505949`.
21. E. Pitacco, M. Denuit, S. Haberman, and A. Olivieri. *Modelling longevity dynamics for pensions and annuity business.* Oxford University Press, 2009. ISBN 978-0199547272. URL `https://www.amazon.com/Modelling-Longevity-Dynamics-Pensions-Mathematics/dp/0199547270?SubscriptionId=AKIAIOBINVZYXZQZ2U3A&tag=chimbori05-20&linkCode=xm2&camp=2025&creative=165953&creativeASIN=0199547270`.
22. B. L. S. Prakasa Rao. *Statistical inference for fractional diffusion processes.* Wiley Series in Probability and Statistics. John Wiley & Sons, Ltd., Chichester, 2010. ISBN 978-0-470-66568-8.
23. M. B. Priestley. *Spectral analysis and time series. Vol. 1.* Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], London-New York, 1981. ISBN 0-12-564901-0. Univariate series, Probability and Mathematical Statistics.
24. R. Weron. Estimating long-range dependence: finite sample properties and confidence intervals. *Phys. A*, 312(1-2):285–299, 2002. ISSN 0378-4371. doi: 10.1016/S0378-4371(02)00961-5. URL `https://doi.org/10.1016/S0378-4371(02)00961-5`.
25. F. Yerlikaya-Özkurt, C. Vardar-Acar, Y. Yolcu-Okur, and G.-W. Weber. Estimation of the Hurst parameter for fractional Brownian motion using the CMARS method. *J. Comput. Appl. Math.*, 259(part B):843–850, 2014. ISSN 0377-0427. doi: 10.1016/j.cam.2013.08.001. URL `https://doi.org/10.1016/j.cam.2013.08.001`.
26. C. Zeng, Y. Chen, and Q. Yang. The fBm-driven Ornstein-Uhlenbeck process: probability density function and anomalous diffusion. *Fract. Calc. Appl. Anal.*, 15(3):479–492, 2012. ISSN 1311-0454. doi: 10.2478/s13540-012-0034-z. URL `https://doi.org/10.2478/s13540-012-0034-z`.