Maximum likelihood estimators from discrete data modeled by mixed fractional Brownian motion with application to the Nordic stock markets

Author(s): Dufitinema, Josephine; Pynnönen, Seppo; Sottinen, Tommi

Title: Maximum likelihood estimators from discrete data modeled by mixed fractional Brownian motion with application to the Nordic stock markets

Year: 2020

Version: Final draft (post print, aam, accepted manuscript)

Maximum likelihood estimators from discrete data modeled by mixed fractional Brownian motion with application to the Nordic stock markets

Josephine Dufitinema[a*] , Seppo Pynnönen[a] and Tommi Sottinen[a]

[a] School of Technology and Innovation, Mathematics and Statistics Unit, University of Vaasa, Finland

ARTICLE HISTORY
Compiled April 29, 2020

ABSTRACT
Mixed fractional Brownian motion is a linear combination of Brownian motion and independent Fractional Brownian motion that is extensively used for option pricing. The consideration of the mixed process is able to capture the long–range dependence property that financial time series exhibit. This paper examines the problem of deriving simultaneously the estimators of all the unknown parameters for a model driven by the mixed fractional Brownian motion using the maximum likelihood estimation method. The consistency and asymptotic normality properties of these estimators are provided. The performance of the methodology is tested on simulated data sets, and the outcomes illustrate that the maximum likelihood technique is efficient and reliable. An empirical application of the proposed method is also made to the real financial data from four Nordic stock market indices.

KEYWORDS
Mixed fractional Brownian motion; long–range dependence; Maximum likelihood estimation; Nordic stock market indices.

1. Introduction

The construction of financial models that capture the fluctuation of financial assets has been recently based on the mixed fractional Brownian motion (henceforth mfBm); a linear combination of Brownian motion and independent fractional Brownian motion. The reason is that financial time series have been found to exhibit long–range dependence [11,16,27]. And, in the classical Black and Scholes model, the standard Brownian motion process is unable to capture the stock price movements; because of its independent increment property [4]. To overcome this issue, the fractional Black–Scholes model was introduced where Brownian motion is substituted by fractional Brownian motion (henceforth fBm); a parameterized ($H \in (0,1)$) extension of Brownian motion with short–range dependence for $H < 1/2$ and long–range dependence increments for $H > 1/2$ [12,19,21]. However, except in the Brownian motion case ($H = 1/2$), fBm is neither a semi–martingale nor a Markov process; the reason why the arbitrage opportunities appeared in the proposed model [7,30]. Therefore, to take into consideration the long memory property; Cheridito [6] replaced the Brownian motion with mfBm

*Corresponding author. Email: josephine.dufitinema@uwasa.fi

for $H \in (3/4, 1)$, and proved that the mfBm model is equivalent to the one driven by Brownian motion, and thus it is arbitrage–free.

The suitable and promising use of mfBm to describe some phenomena in different areas raised interests in identifying the parameters of this process. Filatova [14] suggested that mfBm process reflected better the real network traffic properties; and moreover, the author developed the mfBm parameters estimation method for stochastic modeling for computer network traffic. In the discrete framework, as in practical terms, observations are only available in discrete time [32]; for instance, in finance, where stock prices are collected once a day. The estimation problem of the mfBm parameters was studied in [36] by the use of the maximum likelihood method; and proofs of the estimators' asymptotic properties were provided. Furthermore, paper [37] combined the maximum likelihood approach with Powell's method to compute these estimators efficiently. Recall that the articles mentioned above investigated the parametric estimation problem of the mfBm process itself. To the best of our knowledge, the problem of estimating simultaneously all the unknown parameters in a model driven by the mfBm process (mixed fractional Black–Scholes model), and its application to the real data has not been studied before. Unlike in the case of the traditional geometric Brownian motion [5], the drift fractional Brownian motion [18,34], and the geometric fractional Brownian motion (see [20,35], and more recently [31]). Therefore, this article aims to bridge that gap. Furthermore, in literature, the separate estimation for the Hurst parameter H has been extensively studied using different estimation methods such as the R/S (rescaled analysis) [3,13,24]. In this article, inspired by Weilin et al. [34] and Misiran et al. [20] work in the fractional Brownian motion case; we propose a hybrid simultaneous estimation approach of all the unknown parameters including the Hurst index in the mixed fractional Brownian case; where the optimal value of the Hurst parameter H is obtained by maximizing numerically the profile likelihood function.

The main contributions of this paper are first, to construct in the discrete framework, the estimators of all the unknown parameters of the model driven by the mfBm process using the maximum likelihood estimation method. The choice of the maximum likelihood method in the range of other existing methods (least squares estimation, minimum distance estimation, to cite few) is based on its well–known desirable asymptotic properties such as consistency, normality, and efficiency. Secondly, to study the consistency and the asymptotic normality of these estimators. Thirdly, to illustrate the efficiency performance of our algorithm through Monte Carlo simulations. Numerical computations indicate that the proposed method performs significantly well and provides reasonable estimations of all the unknown parameters of the studied model. Finally, to show the application of our approach in a realistic context through an empirical study.

This article is organized as follows. Section 2 addresses the estimation problem of all the unknown parameters of the model governed by the mfBm process and provides their maximum likelihood estimators. Section 3 deals with the study of the convergence and the asymptotic normality of these estimators. The proposed approach is assessed in section 4 by some numerical experiments. Section 5 presents our empirical results of four Nordic stock market indices. Section 6 concludes the article and offers suggestions for further research.

2. Parameter identification

2.1. Model specification

On a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, consider a Geometric Brownian motion market model with investment in risky–asset whose price satisfies $dS_t = \mu S_t dt + \sigma S_t dB_t,\ t \in [0, T]$, where $\mu \in (-\infty, \infty)$ and $\sigma > 0$ are the drift and the volatility parameter respectively. To allow the long–memory property, replace the Brownian driver process σB_t by the mixed process $M_t = \sigma B_t + \tau B_t^H$; a linear combination of B_t and independent fBm B_t^H of Hurst parameter $H \in (0, 1)$, where $(\sigma, \tau) \neq (0, 0)$ are two real constants. For more elaborate details on the properties of mfBm, we refer to [38].
Hence, the model for the risky–asset governed by mfBm is given by the following equation:

$$S_t = S_0 \exp\left((\mu - \frac{1}{2}\sigma^2)t + \sigma(B_t + \lambda B_t^H) \right) = \exp(Y_t), \tag{1}$$

with $S_0 > 0$ and $\lambda = \frac{\tau}{\sigma}$.

Since $H > 1/2$, the quadratic variation of the mixed process is $\sigma^2 t$. Consequently, (1) is the solution of the Itô–Föllmer forward–type pathwise (ω–by–ω)(see [15]) stochastic differential equation:

$$dS_t = \mu S_t dt + \sigma S_t dM_t,\ t \geq 0. \tag{2}$$

It is worth emphasizing that (2) requires $H \geq \frac{1}{2}$ to admit a solution. However, in this article, we are interested in the study of long–range dependence. Therefore, the Hurst parameter is strictly greater than $\frac{1}{2}$, that is, $\frac{1}{2} < H < 1$.
Hence estimating the parameters from (2) is equivalent to estimating the unknown parameters from the following model:

$$Y_t = (\mu - \frac{1}{2}\sigma^2)t + \sigma(B_t + \lambda B_t^H), \quad t \geq 0. \tag{3}$$

Y_t is equal to the RHS of (3) in $\mathbb{L}^2$ and with probability one.

Assume that the studied process is observed at discrete–time points $(t_1, t_2, ..., t_N)$ and $t_k = kh, k = 1, 2, ..., N$ for a fixed interval $h > 0$ for the notation simplification purpose. Thus the observation vector is $\mathbf{Y} = (Y_{t_1}, Y_{t_2}, ..., Y_{t_N})^T$. $(^T)$ denotes the vector transposition. Hence, the discrete–time observation can be formulated in the form of the vector as follows:

$$\mathbf{Y} = (\mu - \frac{1}{2}\sigma^2)t + \sigma(B + \lambda B^H), \tag{4}$$

where $\mathbf{Y} = (Y_h, Y_{2h}, ..., Y_{Nh})^T$, $t = (t_1, t_2, ..., t_N)^T = (h, 2h, ..., Nh)^T$, $B = (B_h, B_{2h}, ..., B_{Nh})^T$, and $B^H = (B_h^H, B_{2h}^H, ..., B_{Nh}^H)^T$. We aim to construct estimators of all the unknown parameters μ, σ, λ and H and study their asymptotic properties as $N \to \infty$.

2.2. Estimation procedures

We use the maximum likelihood estimation (MLE) of the unknown parameters of the mfBm in (1) based on discrete observations discussed above. The reason behind the choice of MLE is its efficient application in a large set [26].

In (4), let $m = \mu - \frac{1}{2}\sigma^2$. Then, consider a general form of the model as follows:

$$\mathbf{Y} = mt + \sigma(B + \lambda B^H). \tag{5}$$

The estimates for the drift parameter μ will be deduced from the estimates of m. The evaluation of the likelihood function of $\mathbf{Y}$ is explicit, as the law of $\mathbf{Y}$ is Gaussian, its joint probability density function is

$$f(\mathbf{Y}; m, \sigma^2, \lambda^2, H) = (2\pi\sigma^2)^{-\frac{N}{2}} |\Gamma|^{-\frac{1}{2}} \exp\left[-\frac{1}{2\sigma^2}(\mathbf{Y} - mt)^T \Gamma^{-1} (\mathbf{Y} - mt)\right],$$

where $|\Gamma|$ is the determinant of the covariance matrix

$$\begin{aligned} \Gamma = \Gamma(H, \lambda^2) &= [\mathrm{Cov}[B_{ih}, B_{jh}] + \lambda^2 \mathrm{Cov}[B^H_{ih}, B^H_{jh}]]_{i,j=1,2,\ldots,N} \\ &= [h(i \wedge j) + \frac{\lambda^2}{2} h^{2H} (i^{2H} + j^{2H} - |i-j|^{2H})]_{i,j=1,2,\ldots,N}; \end{aligned}$$

with $\lambda = \dfrac{\tau}{\sigma}$ and $i \wedge j$ denotes the minimum between i and j.

The log–likelihood function for $\theta = (m, \sigma^2, \lambda^2, H)$ is

$$l(\mathbf{Y}; \theta) = -\frac{N}{2}\ln(2\pi) - \frac{N}{2}\ln(\sigma^2) - \frac{1}{2}\ln|\Gamma| - \frac{1}{2\sigma^2}(\mathbf{Y} - mt)^T \Gamma^{-1} (\mathbf{Y} - mt).$$

The ML estimate $\hat{\theta}$ of θ is obtained as the solution to the following system of equations

$$\frac{\partial l(\mathbf{Y}; \hat{\theta})}{\partial \hat{m}} = \frac{1}{2\hat{\sigma}^2} t^T \Gamma^{-1} (\mathbf{Y} - \hat{m}t) = 0, \tag{6a}$$

$$\frac{\partial l(\mathbf{Y}; \hat{\theta})}{\partial \hat{\sigma}^2} = -\frac{N}{2\hat{\sigma}^2} + \frac{1}{2\hat{\sigma}^4}(\mathbf{Y} - \hat{m}t)^T \Gamma^{-1} (\mathbf{Y} - \hat{m}t) = 0, \tag{6b}$$

$$\frac{\partial l(\mathbf{Y}; \hat{\theta})}{\partial \hat{\lambda}^2} = -\frac{1}{2}\mathrm{tr}(\hat{\Gamma}^{-1}\Gamma_H) + \frac{1}{2\hat{\sigma}^2}(\mathbf{Y} - \hat{m}t)^T \hat{\Gamma}^{-1} \Gamma_H \hat{\Gamma}^{-1} (\mathbf{Y} - \hat{m}t) = 0, \tag{6c}$$

$$\frac{\partial l(\mathbf{Y}; \hat{\theta})}{\partial \hat{H}} = -\frac{1}{2}\mathrm{tr}(\hat{\Gamma}^{-1}\hat{\Gamma}') + \frac{1}{2\hat{\sigma}^2}(\mathbf{Y} - \hat{m}t)^T \hat{\Gamma}^{-1} \hat{\Gamma}' \hat{\Gamma}^{-1} (\mathbf{Y} - \hat{m}t) = 0, \tag{6d}$$

where tr denotes 'trace',

$$\hat{\Gamma} := \Gamma(\hat{H}, \hat{\lambda}^2), \quad \hat{\Gamma}' = \frac{\partial \Gamma(\hat{H}, \hat{\lambda}^2)}{\partial \hat{H}},$$

and

$$\Gamma_H = \text{Cov}[B^H_{ih}, B^H_{jh}]_{i,j=1,2,\ldots,N} = \frac{1}{2}h^{2H}(i^{2H} + j^{2H} - |i-j|^{2H})_{i,j=1,2,\ldots,N}.$$

To obtain (6c) and (6d), we use the differentiation formulas of a matrix with respect to given parameter. For more details, see [25,28].
(6a) and (6b) result in the estimators

$$\hat{m}_{\lambda^2,H} = \frac{t^T\Gamma^{-1}\mathbf{Y}}{t^T\Gamma^{-1}t}, \tag{7}$$

$$\hat{\sigma}^2_{\lambda^2,H} = \frac{1}{N}\left(\frac{(\mathbf{Y}^T\Gamma^{-1}\mathbf{Y})(t^T\Gamma^{-1}t) - (t^T\Gamma^{-1}\mathbf{Y})^2}{t^T\Gamma^{-1}t}\right). \tag{8}$$

Observe that the accuracy of the estimators $\hat{m}_{\lambda^2,H}$ and $\hat{\sigma}_{\lambda^2,H}$ depend crucially on λ^2 and H, which also need to be estimated. Unlike (6a) and (6b); (6c) and (6d) do not lead to an explicit form. Therefore, following Xiao et al. [36], we apply a hybrid approach in which we replace σ and m by their maximum likelihood solutions ((7) and (8) respectively) in the log–likelihood equation; and maximize the resulting function in terms of the remaining parameters. Carrying out the substitution and dropping constant terms, we get the profile log–likelihood function

$$-\frac{1}{2}\ln|\Gamma| - \frac{N}{2}\ln\left(\frac{(\mathbf{Y}^T\Gamma^{-1}\mathbf{Y})(t^T\Gamma^{-1}t) - (t^T\Gamma^{-1}\mathbf{Y})^2}{t^T\Gamma^{-1}t}\right). \tag{9}$$

To obtain the estimators of the parameters λ^2 and H, we numerically solve the partial optimization problem given by (9) above. Note that maximization of (9) is equivalent to the minimization of its negative log–likelihood. Hence, the optimal values of $\hat{\lambda}^2$ and $\hat{H}$ are obtained by using the function fminsearch in MATLAB. Next, we calculate the values of the estimators $\hat{m}$ and $\hat{\sigma}$ by substituting H with $\hat{H}$ and λ^2 with $\hat{\lambda}^2$ in (7) and (8) respectively. Finally, we deduce the values of the estimators $\hat{\mu} = \hat{m} + \frac{1}{2}\hat{\sigma}^2$ and $\hat{\tau} = \hat{\lambda}\hat{\sigma}$.

3. Asymptotic properties

This section studies the asymptotic behaviors, namely the L^2–consistency, the strong consistency, and the asymptotic normality of the maximum likelihood estimators with closed forms presented in (7) and (8). The proofs of the asymptotic properties of the estimators $\hat{m}_{\lambda^2,H}$ and $\hat{\sigma}^2_{\lambda^2,H}$ hold under the assumption that λ^2 and H are known constants. Regarding the estimators of λ^2 and H, as they do not have closed forms; their optimal values are obtained, and their asymptotic properties are studied through a numerical approach. That is, using a simulation technique presented in section 4.

First, consider the L^2–consistency of the drift parameter defined by (7). We need the following technical results.

Lemma 3.1. *If $A \in \mathbb{R}^{N\times N}$ is a positive definite matrix, $x \in \mathbb{R}^N$, $x \neq 0$ is a non–zero vector, then*

$$x^T A^{-1} x \geq \frac{\|x\|^4}{x^T A\, x}.$$

Proof. See [22, Lemma 2.4]. □

Remark 1. As mentioned above, the studied process Y_t is observed at discrete–time points $t_k = kh, k = 1, 2, ..., N$ for a fixed interval $h > 0$. Therefore, following Mishura et al. [22] as the driver process M_t has stationary increments; the $N \times N$ covariance matrix Γ has Toeplitz structure, that is,
$\Gamma_{k+l,l} = \Gamma_{l,k+l} = \mathbb{E}\left(M_{(k+l)h} - M_{(k+l-1)h}\right)\left(M_{lh} - M_{(l-1)h}\right) = \mathbb{E}\, M_{(k+1)h}M_h - \mathbb{E}\, M_{kh}M_h$ does not depend on l due to the stationarity of increments.

In particular, the following results hold under the assumption that λ^2 and H are known constants. Factually, it is the L^2–consistency that needs the assumption, while the expectation in the unbiasedness can be considered as the conditional expectation that does not depend on these parameters (λ^2 and H); thereby being equal to the unconditional expectation. Thus, the unbiasedness holds whether or not λ^2 and H are estimated.

Theorem 3.2. *The maximum likelihood estimator $\hat{m}_{\lambda^2,H}$ (defined by (7)) is unbiased and converge in mean square to m as $N \to \infty$.*

Proof. Substituting $\mathbf{Y}$ by $mt + \sigma(B_t + \lambda B_t^H)$ in (7), we have

$$\hat{m}_{\lambda^2,H} = \frac{t^T\Gamma^{-1}[mt + \sigma(B_t + \lambda B_t^H)]}{t^T\Gamma^{-1}t} = m + \sigma\, \frac{t^T\Gamma^{-1}(B_t + \lambda B_t^H)}{t^T\Gamma^{-1}t}. \tag{10}$$

Thus,

$$\mathbb{E}[\hat{m}_{\lambda^2,H}] = m + \sigma\, \frac{t^T\Gamma^{-1}\,\mathbb{E}(B_t + \lambda B_t^H)}{t^T\Gamma^{-1}t} = m,$$

and hence $\hat{m}_{\lambda^2,H}$ is unbiased.

Moreover,

$$\begin{aligned}\mathrm{Var}[\hat{m}_{\lambda^2,H}] &= \sigma^2 \mathrm{Var}\left[\frac{t^T\Gamma^{-1}(B_t + \lambda B_t^H)}{t^T\Gamma^{-1}t}\right]\\ &= \sigma^2\,\mathbb{E}\left[\frac{t^T\Gamma^{-1}(B_t + \lambda B_t^H)(B_t + \lambda B_t^H)^T\Gamma^{-1}t}{(t^T\Gamma^{-1}t)^2}\right]\\ &= \sigma^2\,\frac{t^T\Gamma^{-1}\Gamma\,\Gamma^{-1}t}{(t^T\Gamma^{-1}t)^2} = \frac{\sigma^2}{t^T\Gamma^{-1}t}.\end{aligned}$$

By Remark 1 and from Mishura et al. discrete scheme results [22, Theorem 2.5], we have

$$t^T\Gamma\; t = h^2 \sum_{l=1}^{N}\sum_{m=1}^{N} \Gamma_{l,m} \quad \text{and} \quad \|t\| = h\sqrt{N},$$

where $\Gamma_{l,m} = \mathbb{E}\left(M_{(|l-m|+1)h} - M_{|l-m|h}\right)M_h$, with $M_t = \sigma B_t + \tau B_t^H$ the mixed process. As the process M_t has stationary increments and by Toeplitz theorem,

$$\frac{1}{N^2}\sum_{l=1}^{N}\sum_{m=1}^{N}\Gamma_{l,m} = \frac{1}{N}\,\mathbb{E}(M_h)^2 - \sum_{k=2}^{N}\frac{2(N+1-k)}{N^2}\,\mathbb{E}(M_{kh} - M_{(k-1)h})M_h$$
$$\to \lim_{k\to\infty}\mathbb{E}(M_{kh} - M_{(k-1)h})M_h = 0 \quad \text{as} \quad N \to \infty.$$

Hence, with the use of Lemma 3.1,

$$\mathrm{Var}[\hat{m}_{\lambda^2,H}] = \frac{\sigma^2}{t^T\Gamma^{-1}t} \leq \sigma^2\,\frac{t^T\Gamma\;t}{\|t\|^4} = \frac{\sigma^2}{h^2N^2}\sum_{l=1}^{N}\sum_{m=1}^{N}\Gamma_{l,m} \to 0 \quad \text{as} \quad N \to \infty.$$

□

Next, we study the estimator $\hat{\sigma}^2_{\lambda^2,H}$ in (8).
Again, in what follows, the expectation in the unbiasedness can be considered as the conditional expectation that does not depend on parameters λ^2 and H; and thereby being equal to the unconditional expectation. Thus, in the following results, both unbiasedness and mean square convergence hold whether or not λ^2 and H are estimated.

Lemma 3.3. *Let $Z = (Z_1, \ldots, Z_N)^T$ be an N–vector of independent $\mathcal{N}(0,1)$ random variables, then*

$$\mathbb{E}\left[(Z^TZ)(a^TZ)^2\right] = (N+2)a^Ta, \tag{11}$$

where a is a real valued N–vector.

Proof. We can write

$$(Z^TZ)(a^TZ)^2 = a^T(Z^TZ)(ZZ^T)a.$$

In the matrix $(Z^TZ)(ZZ^T)$ the diagonal elements are of the form

$$(Z^TZ)Z_i^2 = Z_i^4 + Z_i^2\sum_{j\neq i}Z_j^2$$

and the off diagonal elements, with $i \neq j$, are of the form

$$(Z^TZ)Z_iZ_j = Z_i^3Z_j + Z_iZ_j^3 + Z_iZ_j\sum_{k\neq i,j}Z_k^2.$$

By independence and properties of $\mathcal{N}(0,1)$ random variables, the expected values of the diagonal elements are,

$$\mathbb{E}\left[Z_i^4 + Z_i^2\sum_{j\neq i}Z_j^2\right] = E\left[Z_i^4\right] + \mathbb{E}\left[Z_i^2\right]\sum_{j\neq i}\mathbb{E}\left[Z_j^2\right] = 3 + (N-1) = N+2,$$

and the off diagonal elements are all zero. Thus, $\mathbb{E}\left[(Z^T Z)(ZZ^T)\right]$ is a diagonal matrix with diagonal elements $N+2$. Therefore,

$$\mathbb{E}\left[(Z^T Z)(a^T Z)^2\right] = a^T \mathbb{E}\left[(Z^T Z) Z Z^T\right] a = (N+2) a^T a,$$

which completes the proof. □

Theorem 3.4. *The maximum likelihood estimator $\hat{\sigma}^2_{\lambda^2,H}$ (defined by (8)) is asymptotically unbiased and converges in mean square to σ^2 as $N \to \infty$.*
Furthermore,

$$\mathbb{E}\left[\hat{\sigma}^2_{\lambda^2,H}\right] = \frac{N-1}{N}\sigma^2 \tag{12}$$

and

$$\mathrm{Var}\left[\hat{\sigma}^2_{\lambda^2,H}\right] = \frac{2(N-1)}{N^2}\sigma^4. \tag{13}$$

Proof. Substituting $\mathbf{Y}$ by $mt + \sigma(B_t + \lambda B_t^H)$ in (8), we have

$$\hat{\sigma}^2_{\lambda^2,H} = \frac{\sigma^2}{N}\left[(B_t + \lambda B_t^H)\Gamma^{-1}(B_t + \lambda B_t^H) - \frac{[t^T \Gamma^{-1}(B_t + \lambda B_t^H)]^2}{t^T \Gamma^{-1} t}\right]. \tag{14}$$

Because $B_t + \lambda B_t^H \sim \mathcal{N}(0, \Gamma)$, then $\Gamma^{-1/2}(B_t + B_t^H) \sim \mathcal{N}(0, I_N)$, where $\Gamma^{-1/2}$ is the inverse of $\Gamma^{1/2}$ which is a symmetric matrix such that $\Gamma^{1/2}\Gamma^{1/2} = \Gamma$, called the square root of Γ, and I_N is the $N \times N$ identity matrix. Furthermore, $(B_t + \lambda B_t^H)^T \Gamma^{-1}(B_t + \lambda B_t^H) \sim \chi^2_N$, $t^T \Gamma^{-1}(B_t + \lambda B_t^H) \sim \mathcal{N}(0, t^T \Gamma^{-1} t)$, and $\frac{t^T \Gamma^{-1}(B_t + \lambda B_t^H)}{\sqrt{t^T \Gamma t}} \sim \mathcal{N}(0,1)$.
As a result, by these arguments, the first component in (14) is a chi–squared random variable with N degrees of freedom, which also coincides its expected value. The second argument is a squared standard normal variable, so that the expected value is its variance, which equals one.
Therefore, we get immediately

$$\mathbb{E}[\hat{\sigma}^2_{\lambda^2,H}] = \frac{\sigma^2}{N}(N-1) = \frac{N-1}{N}\sigma^2,$$

which converges to σ^2 as $N \to \infty$, implying the asymptotic unbiasedness.

Next we prove the convergence in mean square. Because of the asymptotic unbiasedness, we need to prove only that $\mathrm{Var}[\hat{\sigma}^2_{\lambda^2,H}] \to 0$ as $N \to \infty$.

Now,

$$\begin{aligned}
\mathrm{Var}[\hat{\sigma}^2_{\lambda^2,H}] &= \mathbb{E}[(\hat{\sigma}^2_{\lambda^2,H})^2] - (\mathbb{E}[\hat{\sigma}^2_{\lambda^2,H}])^2 \\
&= \frac{\sigma^4}{N^2}\Bigg[\mathbb{E}\left[\left((B_t+\lambda B_t^H)^T\Gamma^{-1}(B_t+\lambda B_t^H)\right)^2\right] \\
&\quad -2\mathbb{E}\left[(B_t+\lambda B_t^H)^T\Gamma^{-1}(B_t+\lambda B_t^H)\frac{[t^T\Gamma^{-1}(B_t+\lambda B_t^H)]^2}{t^T\Gamma^{-1}t}\right] \\
&\quad +\mathbb{E}\left[\left(\frac{[t^T\Gamma^{-1}(B_t+\lambda B_t^H)]^2}{t^T\Gamma^{-1}t}\right)^2\right] - (N-1)\Bigg].
\end{aligned}$$

The first expectation in the latter part is the expected value of a squared chi–squared variable with N degrees of freedom. Thus, by straightforward calculations we have,

$$\mathbb{E}\left[\left((B_t+\lambda B_t^H)^T\Gamma^{-1}(B_t+\lambda B_t^H)\right)^2\right] = 2N + N^2 = N(N+2).$$

For the second expectation, we utilize Lemma 3.3 with $Z = \Gamma^{-1/2}(B_t + \lambda B_t^H)$ and $a = \Gamma^{-1/2}t$. Noting that $t^T\Gamma^{-1}(B_t - \lambda B_t^H) = t^T\Gamma^{-1/2}[\Gamma^{-1/2}(B_t - B_t^H)]$ and using Lemma 3.3, we get

$$\mathbb{E}\left[(B_t+\lambda B_t^H)^T\Gamma^{-1}(B_t+\lambda B_t^H)[t^T\Gamma^{-1}(B_t+\lambda B_t^H)]^2\right] = (N+2)(t^T\Gamma t),$$

so that the second expectation becomes $N + 2$.
The last expectation is the expected value of the fourth moment of a standard normal random variable, i.e., the kurtosis, which equals three. Thus,

$$\mathbb{E}\left[\left(\frac{[t^T\Gamma^{-1}(B_t+\lambda B_t^H)]^2}{t^T\Gamma^{-1}t}\right)^2\right] = 3.$$

Collecting the results, we get finally,

$$\mathrm{Var}[\hat{\sigma}^2_{\lambda^2,H}] = \frac{\sigma^4}{N^2}\left(N(N+2) - 2(N+2) + 3 - (N-1)^2\right) = \frac{2(N-1)}{N^2}\sigma^2,$$

which converges to zero as $N \to \infty$.
This completes the proof of the mean square convergence of $\hat{\sigma}^2_{\lambda^2,H}$. □

From a practical point of view, since the values of the parameters λ^2 and H are unknown; it is crucial to point out the continuity of the estimators $\hat{m}_{\lambda^2,H}$ and $\hat{\sigma}_{\lambda^2,H}$ in these parameters which support their usability. With the use of Lemma 3.1 in (10) and (14), it can be noted that the two estimators are continuous in λ^2 and H. Moreover, referring to the simulation results in section 4, the average of the $\hat{\mu}$ estimates is close to the true parameters. The parameter $\hat{\mu}$ is deduced from $\hat{m}$ and $\hat{\sigma}$ as $\hat{\mu} = \hat{m} + \frac{1}{2}\hat{\sigma}^2$. This result also supports the usability of estimators $\hat{m}_{\lambda^2,H}$ and $\hat{\sigma}_{\lambda^2,H}$.

We now show the strong consistency of the MLE $\hat{m}_{\lambda^2,H}$ and $\hat{\sigma}^2_{\lambda^2,H}$ as $N \to \infty$. In particular, the following results hold under the assumption that λ^2 and H are known constants.
we need the following auxiliary statement.

Lemma 3.5. *let $h > 0$ and $\{\hat{m}^{(N)}_{\lambda^2,H}, N = 1, 2, ...\}$ be the ML estimator of the parameter m of the model (3) by the observations $Y_{kh}, k = 1, 2, ...N$. Then the random process $\hat{m}^{(N)}_{\lambda^2,H}$ has independent increments.*

Proof. See [22, Lemma 2.6] □

Theorem 3.6. *The estimators $\hat{m}_{\lambda^2,H}$ and $\hat{\sigma}^2_{\lambda^2,H}$ defined by (7) and (8), respectively, are strong consistent, that is,*

$$\hat{m}_{\lambda^2,H} \to m \quad a.s \quad N \to \infty, \tag{15}$$

$$\hat{\sigma}^2_{\lambda^2,H} \to \sigma^2 \quad a.s \quad as \quad N \to \infty. \tag{16}$$

Proof. First, we discuss the convergence of $\hat{m}_{\lambda^2,H}$.
By Theorem 3.2, $\mathrm{Var}[\hat{m}^{(N)}_{\lambda^2,H}] \to 0$ as $N \to \infty$, therefore, following Mishura et al. [22, Theorem 2.7], we have

$$\begin{aligned}\mathrm{Var}\left[\hat{m}^{(N)}_{\lambda^2,H} - \hat{m}^{(N_0)}_{\lambda^2,H}\right] &= \mathrm{Var}[\hat{m}^{(N)}_{\lambda^2,H}] - \mathrm{Var}[\hat{m}^{(N_0)}_{\lambda^2,H}] \\ &- 2\sqrt{\mathrm{Var}[\hat{m}^{(N)}_{\lambda^2,H}]\mathrm{Var}[\hat{m}^{(N_0)}_{\lambda^2,H}]}\mathrm{corr}\left(\hat{m}^{(N)}_{\lambda^2,H}, \hat{m}^{(N_0)}_{\lambda^2,H}\right) \to \mathrm{Var}[\hat{m}^{(N_0)}_{\lambda^2,H}] \text{ as } N \to \infty.\end{aligned}$$

The process $\hat{m}^{(N)}_{\lambda^2,H}$ has independent increments. Therefore by Kolmogorov's inequality, for $\epsilon > 0$ and $N \in \mathbb{N}$

$$P\left(\sup_{N \geq N_0} \left|\hat{m}^{(N)}_{\lambda^2,H} - \hat{m}^{(N_0)}_{\lambda^2,H}\right| > \frac{\epsilon}{2}\right) \leq \frac{4}{\epsilon^2} \lim_{N\to\infty} \mathrm{Var}\left[\hat{m}^{(N)}_{\lambda^2,H} - \hat{m}^{(N_0)}_{\lambda^2,H}\right] = \frac{4}{\epsilon^2}\mathrm{Var}[\hat{m}^{(N_0)}_{\lambda^2,H}].$$

Then, using the unbiasedness of the estimator, we get

$$\begin{aligned}P\left(\sup_{N \geq N_0} \left|\hat{m}^{(N)}_{\lambda^2,H} - m\right| \geq \epsilon\right) &\leq P\left(\sup_{N \geq N_0} \left|\hat{m}^{(N_0)}_{\lambda^2,H} - m\right| \geq \frac{\epsilon}{2}\right) \\ &+ P\left(\sup_{N \geq N_0} \left|\hat{m}^{(N)}_{\lambda^2,H} - \hat{m}^{(N_0)}_{\lambda^2,H}\right| \geq \frac{\epsilon}{2}\right) \leq \frac{4}{\epsilon^2}\mathrm{Var}[\hat{m}^{(N_0)}_{\lambda^2,H}] + \frac{4}{\epsilon^2}\mathrm{Var}[\hat{m}^{(N_0)}_{\lambda^2,H}] \\ &= \frac{8}{\epsilon^2}\mathrm{Var}[\hat{m}^{(N_0)}_{\lambda^2,H}] \to 0 \quad \text{as} \quad N_0 \to \infty,\end{aligned}$$

hence $\left|\hat{m}^{(N)}_{\lambda^2,H} - m\right| \to 0$ as $N \to \infty$ almost surely.

Moreover, we will show that

$$\sum_{N \geq 1} P\left(\left|\hat{\sigma}^2_{\lambda^2,H} - \sigma^2\right| > \frac{1}{N^\delta}\right) < \infty, \tag{17}$$

for some $\delta > 0$.

The Chebyshev's inequality combined with the mean square convergence calculations in Theorem 3.4 implies that for any small positive δ

$$P\left[\left|\frac{\sigma^2}{N}\left[\frac{(Y^T\Gamma^{-1}Y)(t^T\Gamma^{-1}t)-(t^T\Gamma^{-1}Y)^2}{t^T\Gamma^{-1}t}\right]-\sigma^2\right|>\frac{1}{N^\delta}\right]$$
$$\leq \frac{\sigma^4}{N^{2\delta}}\,\mathbb{E}\left[1-\frac{1}{N}\left[\frac{(Y^T\Gamma^{-1}Y)(t^T\Gamma^{-1}t)-(t^T\Gamma^{-1}Y)^2}{t^T\Gamma^{-1}t}\right]\right]^2=\frac{\sigma^4}{N^{2\delta+2}}.$$

Thus (17) is proven, which implies (16) by the Borel–Cantelli lemma. □

We now move on to the study of the asymptotic normality of the estimators $\hat{m}_{\lambda^2,H}$ and $\hat{\sigma}^2_{\lambda^2,H}$.
Regarding the asymptotic normality of $\hat{m}_{\lambda^2,H}$, we note that the estimator is normal with expectation m and variance

$$\mathrm{Var}[\hat{m}_{\lambda^2,H}]=\frac{\sigma^2}{t^T\Gamma^{-1}t}.$$

Consequently,

$$\sqrt{t^T\Gamma^{-1}t}(\hat{m}_{\lambda^2,H}-m)\sim\mathcal{N}(0,\sigma^2).$$

Hence, in what follows, we study the asymptotic distribution of $\hat{\sigma}^2_{\lambda^2,H}$.
The proof of the $\hat{\sigma}^2_{\lambda^2,H}$ asymptotic normality requires a criterion from the Malliavin calculus; plus precisely, the results of the Malliavin derivative D with respect to the Gaussian process $M_t = B_t + \lambda B^H_t$. These results are very well presented in Xiao et al. [36]. Therefore, borrowing the idea of Xiao et al., we will make use of the following technical lemma.

Lemma 3.7. *For a time interval $[0,T]$, we denote by $\mathcal{E}$ the set of real–valued step functions on $[0,T]$ and let $\mathcal{H}$ be the Hilbert space defined as the closure of $\mathcal{E}$ with respect to the scalar product*

$$\langle \mathbf{1}_{[0,t]},\mathbf{1}_{[0,s]}\rangle_{\mathcal{H}}=R_H(t,s),$$

where $R_H(t,s)$ is the covariance function of the mfBm. We will denote by $\|.\|_{\mathcal{H}}$ the norm in $\mathcal{H}$ induced by $\langle\cdot,\cdot\rangle_{\mathcal{H}}$; define

$$F_N=\frac{1}{\sigma^2}\sqrt{\frac{N}{2}}(\hat{\sigma}^2_{\lambda^2,H}-\sigma^2)=\frac{1}{\sqrt{2N}}\left[(B_t+\lambda B^H_t)^T\Gamma^{-1}(B_t+\lambda B^H_t)\right]-\sqrt{\frac{N}{2}}.$$

Then we have

$$\|DF_N\|^2_{\mathcal{H}}=\frac{2\hat{\sigma}^2_{\lambda^2,H}}{\sigma^2}.$$

Proof. See [36, Lemma 4.3]. □

The asymptotic distribution of $\hat{\sigma}^2_{\lambda^2,H}$ is embedded in the following theorem.

Theorem 3.8. *We have*

$$\frac{1}{\sigma^2}\sqrt{\frac{N}{2}}(\hat{\sigma}^2_{\lambda^2,H}-\sigma^2)\xrightarrow{\mathcal{L}}\mathcal{N}(0,1)\quad as\quad N\to\infty.$$

Proof. Using the results of Theorem 3.4, we get

$$\lim_{N\to\infty}\mathbb{E}[F_N^2]=\lim_{N\to\infty}\mathbb{E}\left[\frac{1}{\sigma^2}\sqrt{\frac{N}{2}}(\hat{\sigma}^2_{\lambda^2,H}-\sigma^2)\right]^2=\lim_{N\to\infty}\frac{N}{2\sigma^4}\mathbb{E}[\hat{\sigma}^4_{\lambda^2,H}-2\hat{\sigma}^2_{\lambda^2,H}\sigma^2+\sigma^4]=1.$$

By Theorem 3.4 and Lemma 3.7 $\|DF_N\|^2_{\mathcal{H}}$ converges in L^2 to the constant 2. Applying Theorem 4 in Nualart and Ortiz–Latorre [23], the proof is complete. □

4. Simulation Results

This section investigates the efficiency of the constructed maximum likelihood estimators through Monte Carlo simulation studies. The simulation studies aim to illustrate the proven properties of the estimators. That is, to show that in fact, the estimated parameters do converge to the actual values as the sample size increases. The crucial phase of Monte Carlo simulation is the construction of the mixed fractional Brownian motion path. The mfBm simulation procedure used in this article is as follows: First, based on the method of [17], generate the standard Brownian motion. Next, using Wood's method: circulant matrix [33], generate fractional Gaussian noise using the algorithm proposed in [29]. The path of fractional Brownian motion is obtained by taking the cumulative sums of the fractional Gaussian noise. Finally, mfBm path is constructed. For a sample size of 500 and a sampling interval of 0.002, Figure 1 to 3 illustrates the path of the mixed process M_t for different σ, τ, and H parameters values. The figures highlight the fact that the Brownian motion process with Hölder index α and the fractional Brownian motion with Hölder index β, the path of a linear combination between them will be a curve with Hölder index which is the minimum between α and β (see [2]).

In the following, we describe the complete μ, σ, τ, and H parameters estimation procedure from discrete observations $Y_{ih}, 1\le i\le N$, as presented in section 2. The observations $\mathbf{Y}=Y_h,Y_{2h},...,Y_{Nh}$ are simulated for different values of μ, σ, τ,and H, with fixed sampling interval $h=1/252,1/52,1/12$ (data collected by daily, weekly and monthly observations, respectively), and sample size $N=100,200,300$, and 500. For each case, sample sizes are replicated 100 times from the actual model.
In the numeric maximization of (9), as we are dealing with an optimization problem involving an inverse of a large–scale matrix with exponential elements; we follow one of the Coeurjolly's approach [9]; the use of the Cholesky decomposition of the covariance matrix Γ. That is, as Γ is a symmetric positive definite matrix; it can be written as LL^T, where L is a lower triangular matrix. In terms of storage, it is much easier to compute the inverse of a lower triangular matrix. Hence, $\Gamma^{-1}=(L^{-1})^TL^{-1}$ and the determinant of Γ; $|\Gamma|=|L^2|$. To deal with the fact that eigenproblem might be ill–conditioned and hard to compute even when the matrix itself is well–conditioned with respect to the inversion; we use the Multiprecision Computing Toolbox [1]; a MATLAB extension for computing with arbitrary precision. The toolbox allows solving of numerically unstable problems such as ill–conditioned matrices and minimizes rounding and cancellation errors in computations.

The algorithm of our estimation method used in this article is shown in Figure 4 and summarized as follows:

Step 1. Set N: the sample size and h: the sampling interval;

Step 2. Set the values of μ, σ, τ, and H parameters;

Step 3. Generate mixed fractional Brownian motion based on the above–cited algorithms;

Step 4. Construct the path of model 4;

Step 5. Numerically maximize (9) to get the estimators $\hat{H}$ of H and $\hat{\lambda}^2$ of λ^2;

Step 6. Compute $\hat{m}$ by substituting H with $\hat{H}$ and λ^2 with $\hat{\lambda}^2$ in (7);

Step 7. Calculate $\hat{\sigma}$ by substituting H with $\hat{H}$ and λ^2 with $\hat{\lambda}^2$ in (8);

Step 8. Deduce the drift estimator $\hat{\mu} = \hat{m} + \frac{1}{2}\hat{\sigma}^2$;

Step 9. Deduce the estimator $\hat{\tau} = \hat{\lambda}\hat{\sigma}$.

Note that in the case of empirical analysis, step 1 to 4 are skipped and proceed from 5 to 9.

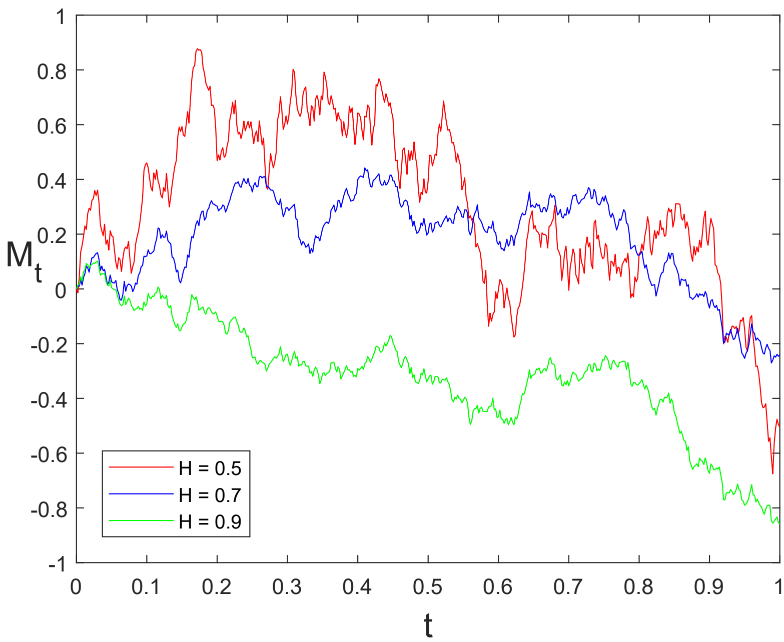

Figure 1. Simulated mfBm paths for different values of H ($\sigma = 0.4$, $\tau = 1.4$).

For the sampling interval $h = 1/252$, the means and standard deviations (S.Dev.) of the estimators for different samples size are given in Table 1; where the true value is the parameter value used in the Monte Carlo simulation. Tables 2 and 3 report $h = 1/52$ and $h = 1/12$ sampling intervals simulation results; to investigate the effect of the underlying sampling interval.

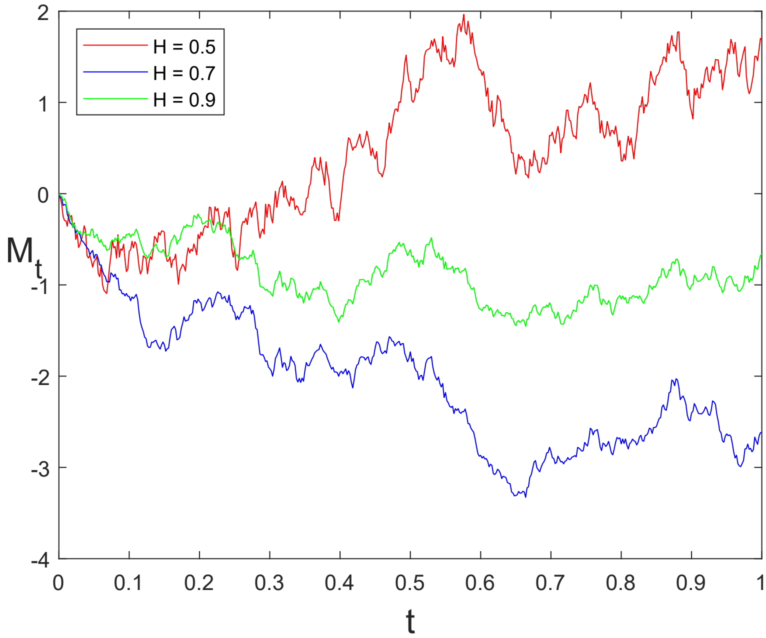

Figure 2. Simulated mfBm paths for different values of H ($\sigma = 1$, $\tau = 3$).

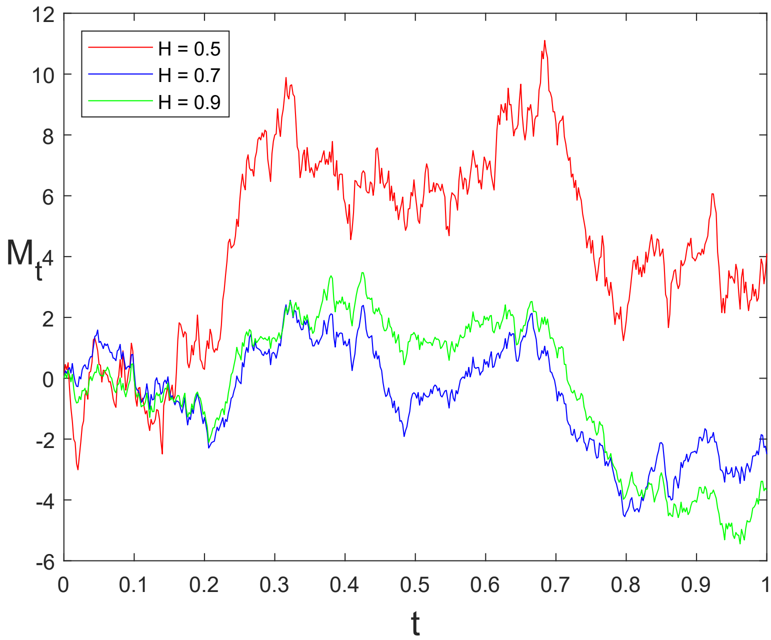

Figure 3. Simulated mfBm paths for different values of H ($\sigma = 5$, $\tau = 12$).

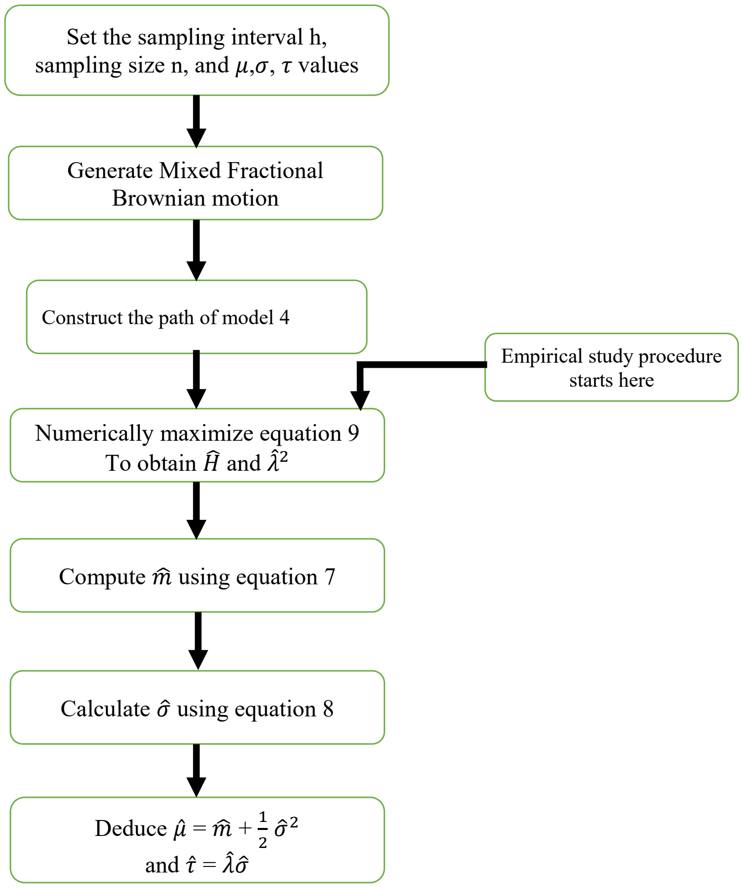

Figure 4. Monte Carlo simulation flow chart

From numerical computations, we observe that on the one hand in all the three sampling intervals cases, as the sample size increases from 100 to 300, the simulated mean of the estimators slowly converges to the true value; on the other hand, the sample variation from 100 to 500 makes the simulated mean converges quickly to the actual value. Therefore, for rapid convergence of the estimated parameters, a large sample size of 500 and above is needed; the larger the sample, the better the estimation. However, overall in all the three cases, the simulated standard deviation and the bias decreases as the number of observations increases. Hence, we can conclude that the Maximum likelihood estimation method proposed in this article performs well since the estimated parameters results move towards their chosen values for $H > 1/2$. Moreover, regarding the effect of the sampling interval, we observe that the obtained estimators are not affected by the sampling interval since neither restriction was placed on the sampling interval nor dependence of it with the estimators.

Regarding the range of the maximization procedure for the parameter $H = 0.5, 0.95$ in the 100 times' simulation replications; for each sampling interval in the four studied sample sizes, it is as follows:

- For $h = 1/252$, when $H = 0.55$, the range's minimum value is 0.5000, and the maximum value is 0.8733. When $H = 0.95$, the range's minimum value is 0.9000, and the maximum is 0.9737.
- For $h = 1/52$, when $H = 0.55$, the minimum value of the range is 0.5000, and

Table 1. Estimation results of the sampling interval $h = 1/252$.

		μ	H	σ	τ	μ	H	σ	τ	μ	H	σ	τ
True value		0.5000	0.5500	0.4000	0.6000	1.5000	0.7500	1.0000	2.5000	4.0000	0.9500	5.0000	12.0000
Block A.	**Sample size N = 100**												
Mean		0.5620	0.5384	0.3981	0.5514	1.7010	0.7921	1.0249	2.7599	4.0443	0.9563	4.8869	16.7499
S.Dev		1.2372	0.0487	0.0299	0.0414	3.8432	0.1192	0.0791	0.2126	15.8592	0.0421	0.3711	1.2721
Block B.	**Sample size N = 200**												
Mean		0.4788	0.5404	0.4138	0.5593	1.3736	0.7631	1.0054	2.6513	3.7057	0.9544	5.0320	14.9444
S.Dev		0.7653	0.0437	0.0220	0.0304	2.6395	0.0915	0.0526	0.1388	12.2513	0.0414	0.2649	0.4709
Block C.	**Sample size N = 300**												
Mean		0.4731	0.5475	0.4128	0.5719	1.4648	0.7520	0.9950	2.5431	3.7410	0.9468	4.9942	10.4727
S.Dev		0.6962	0.0428	0.0174	0.0241	2.4517	0.0776	0.0400	0.1022	11.3128	0.0418	0.2145	0.4498
Block D.	**Sample size N = 500**												
Mean		0.4967	0.5543	0.4066	0.6096	1.4721	0.7514	1.0154	2.5095	3.9050	0.9488	4.9837	12.3924
S.Dev		0.5064	0.0419	0.0131	0.0197	2.2572	0.0749	0.0328	0.0778	12.3879	0.0401	0.1636	0.4069

Table 2. Estimation results of the sampling interval $h = 1/52$.

		μ	H	σ	τ	μ	H	σ	τ	μ	H	σ	τ
True value		0.5000	0.5500	0.4000	0.6000	1.5000	0.7500	1.0000	2.5000	4.0000	0.9500	5.0000	12.0000
Block A.	**Sample size N = 100**												
Mean		0.5330	0.5437	0.4078	0.5649	1.7086	0.7640	1.1097	2.0835	4.0798	0.9666	4.9483	12.8893
S.Dev		0.5907	0.0573	0.0305	0.0423	2.3791	0.0420	0.0849	0.1595	12.5406	0.0188	0.3816	0.9941
Block B.	**Sample size N = 200**												
Mean		0.4924	0.5455	0.4125	0.5714	1.6382	0.7635	1.0460	2.4560	3.7472	0.9640	5.0495	12.2267
S.Dev		0.3684	0.0595	0.0226	0.0313	1.6866	0.0405	0.0558	0.1310	10.4164	0.0171	0.2634	0.6379
Block C.	**Sample size N = 300**												
Mean		0.4917	0.5476	0.4118	0.5746	1.5394	0.7624	1.0340	2.4961	3.8937	0.9629	5.0543	11.6899
S.Dev		0.3331	0.0490	0.0177	0.0245	1.5724	0.0378	0.0426	0.1028	9.7486	0.0160	0.2124	0.4493
Block D.	**Sample size N = 500**												
Mean		0.4933	0.5529	0.4029	0.5857	1.4990	0.7523	1.0056	2.5173	4.0313	0.9523	5.0394	11.7193
S.Dev		0.2455	0.0529	0.0137	0.0190	1.5029	0.0340	0.0323	0.0810	11.2610	0.0155	0.1649	0.3835

Table 3. Estimation results of the sampling interval $h = 1/12$.

		μ	H	σ	τ	μ	H	σ	τ	μ	H	σ	τ
True value		0.5000	0.5500	0.4000	0.6000	1.5000	0.7500	1.0000	2.5000	4.0000	0.9500	5.0000	12.0000
Block A.	**Sample size N = 100**												
Mean		0.5203	0.5463	0.4127	0.5717	1.4196	0.7179	0.8498	2.4156	4.7339	0.9626	5.0951	10.7437
S.Dev		0.2975	0.0609	0.0309	0.0427	1.5671	0.0325	0.0640	0.1818	10.9651	0.0161	0.3964	0.8359
Block B.	**Sample size N = 200**												
Mean		0.4977	0.5413	0.4122	0.5710	1.3083	0.7255	0.8869	2.4542	3.7782	0.9591	5.0913	11.8660
S.Dev		0.1876	0.0425	0.0227	0.0314	1.1363	0.0433	0.0484	0.1341	9.4413	0.0130	0.2653	0.6183
Block C.	**Sample size N = 300**												
Mean		0.4999	0.5473	0.4164	0.5768	1.5018	0.7327	0.9915	2.4559	4.3260	0.9579	5.0251	12.5676
S.Dev		0.1681	0.0480	0.0179	0.0248	1.0653	0.0467	0.0427	0.1014	8.8875	0.0120	0.2046	0.5117
Block D.	**Sample size N = 500**												
Mean		0.5008	0.5490	0.4027	0.5827	1.5017	0.7547	1.0015	2.5393	4.1624	0.9568	5.0066	12.1600
S.Dev		0.1257	0.0415	0.0136	0.0189	1.0393	0.0407	0.0323	0.0852	7.4258	0.0105	0.1631	0.4190

the maximum is 0.7213. When $H = 0.95$, the range's minimum value is 0.9000, and the maximum is 0.9666.

- For $h = 1/12$, when $H = 0.55$, the minimum value of the range is 0.5000, and the maximum is 0.7156. When $H = 0.95$, the range's minimum value is 0.9000, and the maximum is 0.9852.

Hence, when $H = 0.5, 0.95$, all the H estimates fall in the interval 0.5 to 1; $H \in (0.5, 1)$.

5. Empirical study

This section presents the application of our method described in the previous sections to the real data. The purpose is to estimate the parameters of the risky–asset market model using the proposed estimation approach based on the daily return series. The utilized data in this empirical analysis have been retrieved from Thomson Reuters DataStream database. They are daily historical data for the Helsinki stock market index (OMXH25), the Norwegian index (AXLT), the Swedish index (OMXS20), all three indices spanning from January 01, 2010 to December 28, 2018, and containing 2347 observations. The Danish index (OMXC20) covers the period from November 25, 2011 to December 31, 2018, which is 1852 observations. As mentioned above, data are collected daily, in other words, the index prices are observed at a time interval of $h = 1/252$. All indices' closing prices were transformed into logarithmic returns using the formula:

$$r_t = \ln X_t - \ln X_{t-1} = \ln\left(\frac{X_t}{X_{t-1}}\right)$$

where X_{t-1} and X_t are two consecutive observations for a time series, and r_t is its return.

Basic descriptive plots of data are presented in Figure 5 to 8 as follows: In all figures, chart(a) represent the timeline plot of the series; they reflect the non–stationarity as no long–run average appears in all the series. Chart(b) illustrates the logarithm transformation trend of the series over time. Log–returns reveal high volatility, prominent spikes and jump; however, they appear to be stable around the mean. Chart(c) demonstrate the closer normal distribution but not completely Gaussian as the normalized histogram try to capture all features. Chart(d) plots the quantile–quantile(Q–Q) of the log–returns distribution against the standard normal distribution quantiles. The plots indicate that log–returns distribution exhibit an S–shape curve compared to the normal distribution; with the presence of significantly larger historical quantiles in the tail of the distribution; which implies the non–normality distribution of the log–returns.

Table 4 reports for the indices mentioned above, the sample size, their minimum/maximum values, the returns sample mean, standard deviation, skewness, and kurtosis. For all the series, the mean returns are marginally positive. They all are left-skewed as they all have negative skewness. Moreover, all series are leptokurtic as they all display a kurtosis coefficient significantly higher than three. Such features (skewness and kurtosis) implying the non–normality distribution; are common in return series. Q–Q plots also confirm these results (see charts(d) in Figure 5 to 8).

Next, we investigate the presence of long–range dependence property in the studied indices. The method used is the plot of the sample autocorrelation functions of the daily returns. As illustrated in Figure 9 and 10, slow and very weak decay of the autocorrelation function is observed.

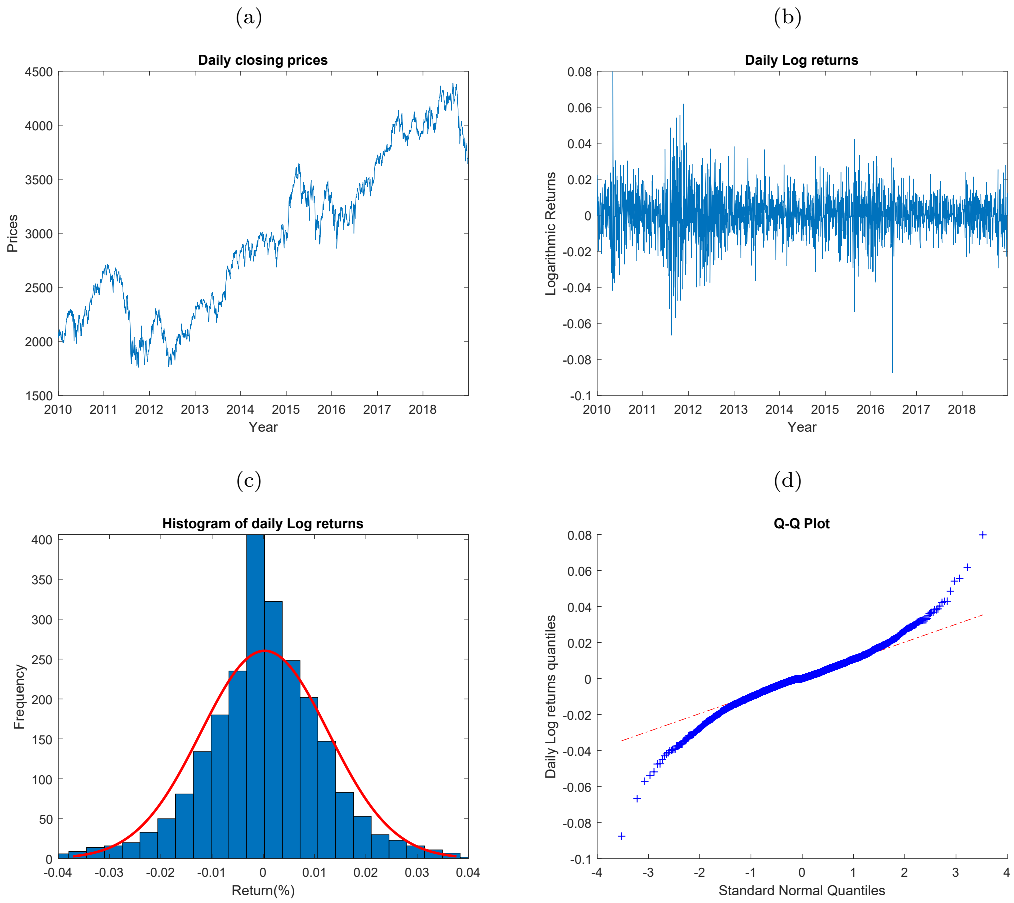

Figure 5. Statistical plots of daily returns for OMXH25 index.

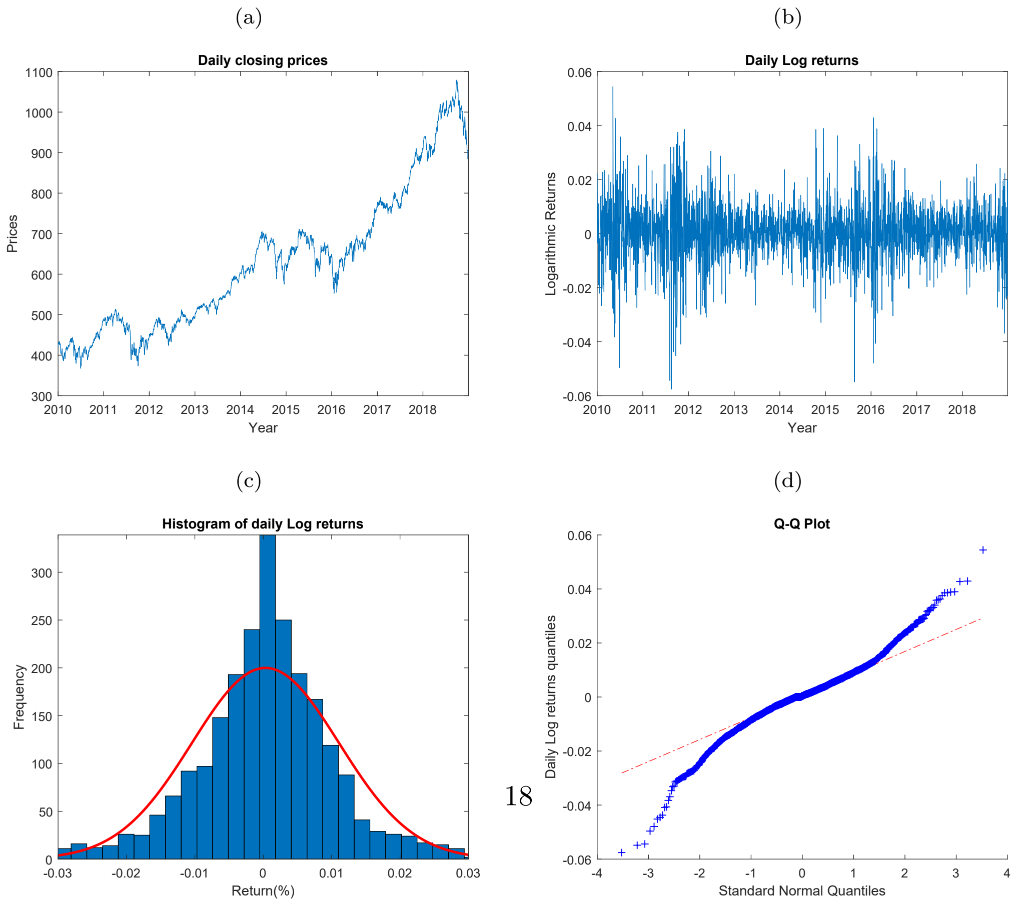

Figure 6. Statistical plots of daily returns for AXLT index.

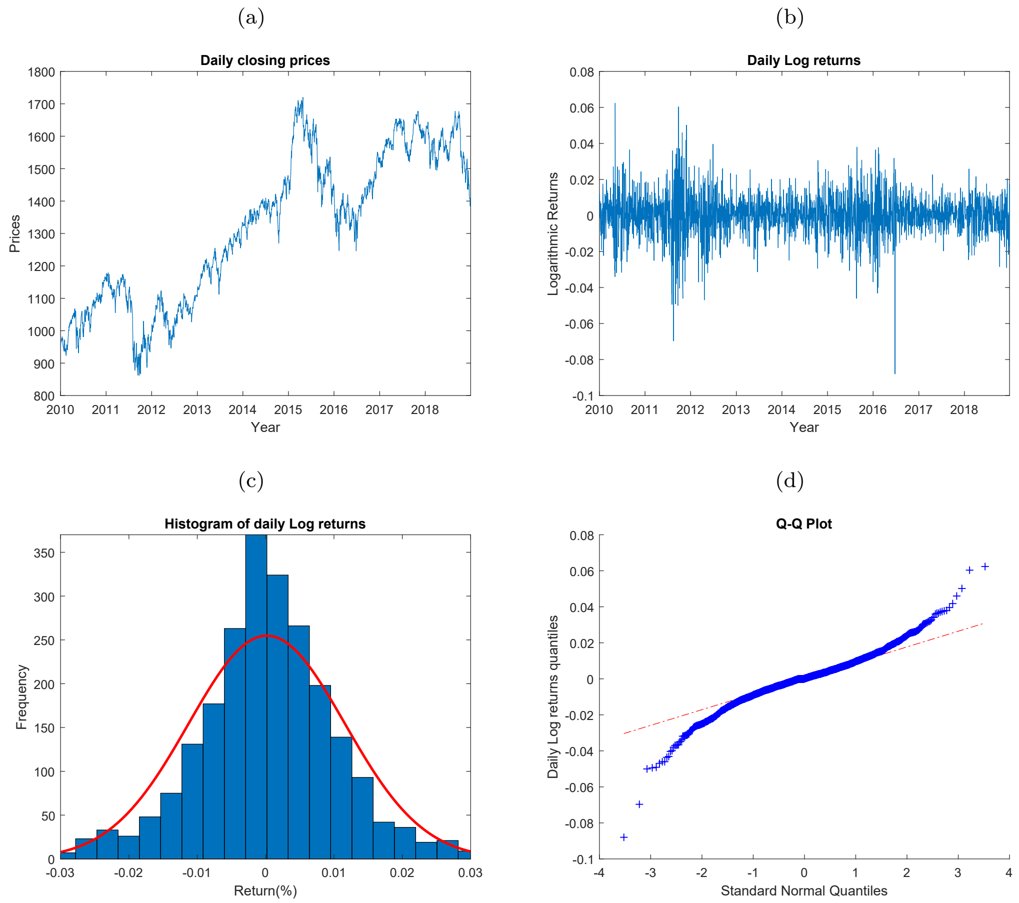

Figure 7. Statistical plots of daily returns for OMXS30 index.

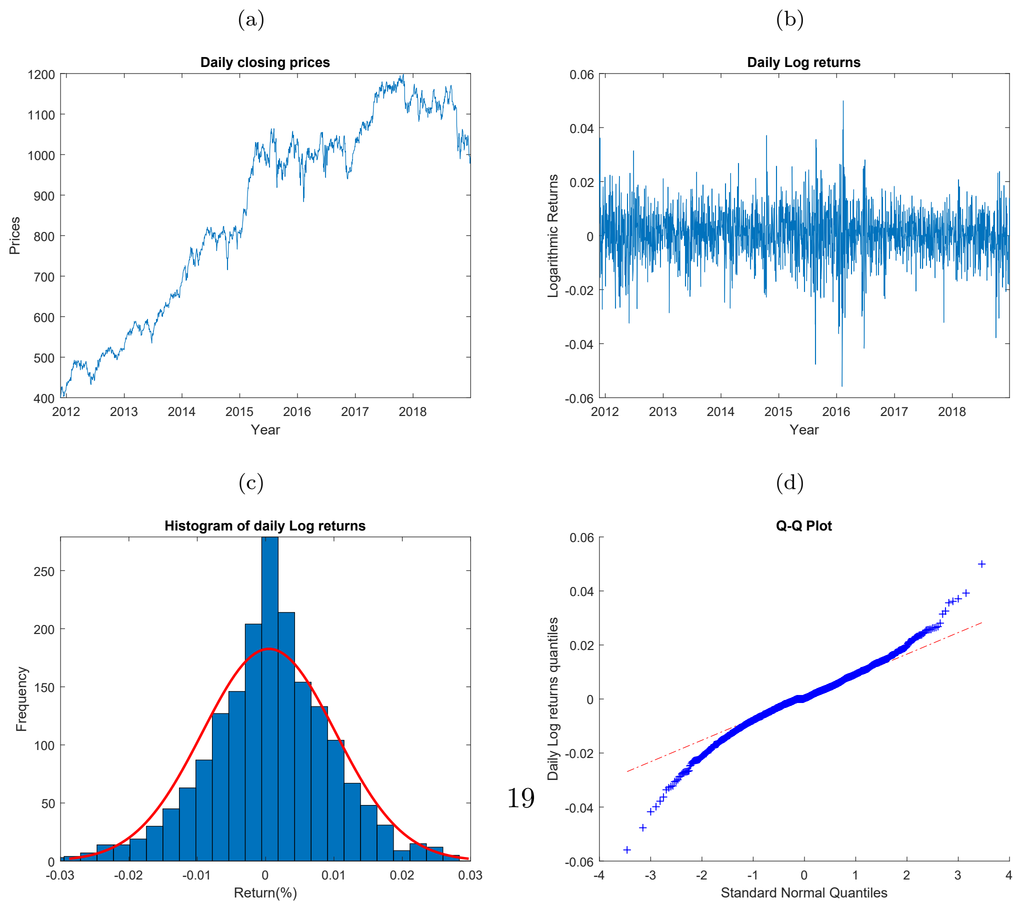

Figure 8. Statistical plots of daily returns for OMXC20 index.

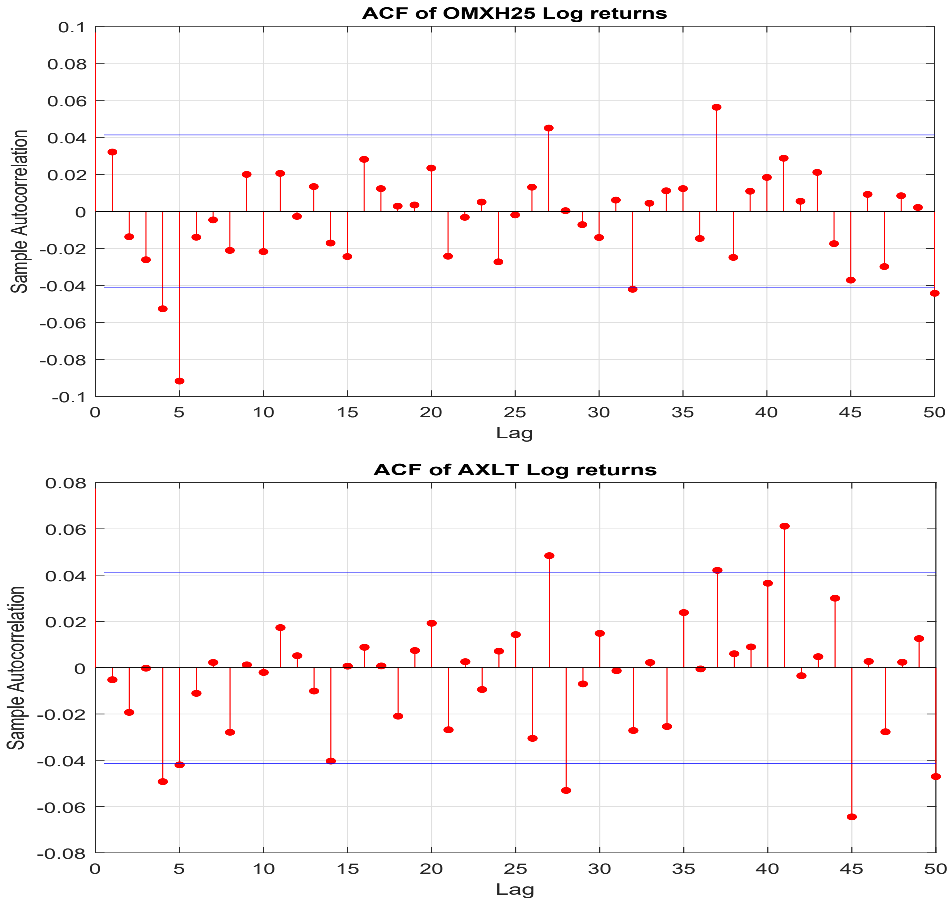

Figure 9. Sample ACF for the OMXH25 and AXLT log-returns.

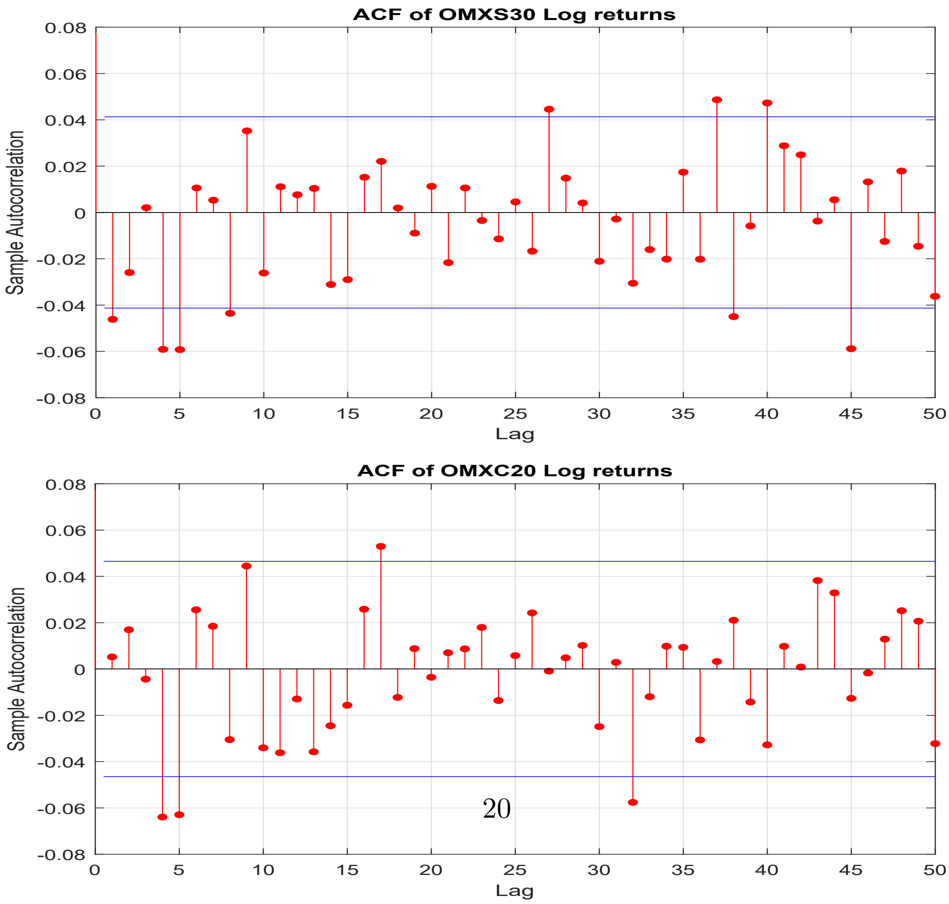

Figure 10. Sample ACF for the OMXS30 and OMXC20 log-returns.

Table 4. Descriptive statistics of daily returns.

Index	Nobs*	Min	Max	Mean	S.Dev	Skew	Kurt
OMXH25	2347	-0.0875	0.0798	$2.536 * 10^{-4}$	0.0125	-0.2072	6.7377
AXLT	2347	-0.0576	0.0544	$3.258 * 10^{-4}$	0.0108	-0.2637	5.8796
OMXS30	2347	-0.0880	0.0624	$1.671 * 10^{-4}$	0.0114	-0.3544	7.3244
OMXC20	1852	-0.0559	0.0500	$4.906 * 10^{-4}$	0.0097	-0.2984	5.3800

***Nobs**: Number of observations.

Therefore, one can conclude that the studied indices have long–range dependence. To what degree? This will be determined by the long–memory parameter estimated in what follows. The final step is to estimate the unknown parameter of the risky–asset market model; namely H, μ, σ, and τ from the considered financial series, using our estimation procedure proposed in section 2 and the estimation algorithm in section 4. Using the real data and (1), we estimated all the desired parameters.

Table 5 presents the estimation results. The results reveal strong and almost equal evidence of long–range dependence in all the return series during the sample period; as the estimated Hurst parameter is higher than 1/2. In other words, all the indices are persistent, and the closer H is to 1, the higher the degree of persistence. This estimated parameter also suggests a high degree of predictability of the future returns in all the indices based on the historical returns. This finding agrees with the work [8]; where the used estimated long–range parameter (the differencing parameter d) was found significantly positive for the case of Finnish, Danish, and Norwegian stock indices. The expected returns ($\hat{\mu}$), as well as, the volatility ($\hat{\sigma}$) parameters in short memory are relatively close for all indices; implying that all indices have reasonably same daily volatility. The magnitude parameters ($\hat{\tau}$) of the long–range dependence is also slightly equal in all cases; except in the case of OMXC20 where a lower amplitude is observed. However, this is due to the OMXC20's number of observations, which is less compared to the other three indices.

Table 5. Empirical results by our method.

Index	Hurst parameter ($\hat{H}$)	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\tau}$
OMXH25	0.8481	0.0379	0.2754	0.0013
AXLT	0.8279	0.0293	0.2423	0.0026
OMXS30	0.8196	0.0344	0.2622	0.0012
OMXC20	0.8716	0.0238	0.2181	$9.523 * 10^{-4}$

6. Conclusion

Stochastic models driven by long–memory processes have grown into essential tools in the financial world to provide a deep understanding of the behavior of the market. Among the driver processes, the mixed fractional Brownian motion has been found to shed much light on the debate to whether a given market is long–range dependent or not. However, a crucial problem which rises in practice is to provide the optimal parameter estimation procedure in the discrete framework of the model driven by mfBm. This article examines the issue of deriving simultaneously the estimators of all the unknown parameters for a mixed fractional Black–Scholes model in the discrete domain. Using the maximum likelihood methodology; the estimators of the drift, the volatility, the long–range dependent amplitude, and the Hurst parameter are constructed. The asymptotic behavior, namely the consistency and the asymptotic normality of these estimators are also provided. Furthermore, Monte Carlo simulation studies are performed to illustrate the efficiency of our algorithm. Numerical computations indicate the asymptotic convergence of the estimated parameters to the actual value, and that the proposed method achieves the purpose significantly well. Moreover, to show the application of our approach to the real data; an empirical study was done on four Nordic stock markets indices.

Regarding the analysis of the long–range dependence in the studied stock indices; results indicate similar and robust evidence of long–term dependence in the returns of all indices with the estimated Hurst parameter above 0.80. This finding implies that in the Nordic stock markets the no–arbitrage condition (more precisely no free lunch with vanishing risk) of the Fundamental Theorem of Asset Pricing [see for example 10] is satisfied; since by Cheridito [6], the mfBm with $H > 0.75$ is equivalent in law to Brownian motion. Further research includes the use of the proposed approach to investigate the possibility of long memory in other stock markets or exchange rates and a comparison of our method with different stochastic models governed by long–memory processes. Also, the data does not support Gaussianity, adding a jump component to the model will able to capture the heavy tails.

References

[1] Advanpix (2019). *Multiprecision Computing Toolbox for MATLAB.* Latest Version 4.7.0.13560 – November 6, 2019. Available at: https://www.advanpix.com.

[2] Azmoodeh, E., Sottinen, T., Viitasaari, L., and Yazigi, A. (2014). Necessary and sufficient conditions for Hölder continuity of Gaussian processes. *Statistics and Probability Letters*, 94:230–235.

[3] Beran, J. (1994). *Statistics for Long–Memory Processes.* New York: Chapman and Hall.

[4] Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:637–654.

[5] Brigo, D., Dalessandro, A., Neugebauer, M., and Triki, F. (2008). A Stochastic Processes Toolkit for Risk Management. *arXiv. org, Quantitavive Finance Papers.*

[6] Cheridito, P. (2001). Mixed Fractional Brownian motion. *Bernoulli*, 7:913–934.

[7] Cheridito, P. (2003). Arbitrage in fractional Brownian motion models. *Finance Stoch*, 7:533–553.

[8] Christodoulou-Volos, C. and Siokis, F. M. (2016). Long range dependence in stock market returns. *Applied Financial Economics*, 16:18:1331–1338.

[9] Coeurjolly, J.-F. (2000). Simulation and Identification of the Fractional Brownian motion: A Bibliographical andCcomparativeStudy. *Journal of Statistical Software*, 5(7):1–53.

[10] Delbaen, F. and Schachermayer, W. (2004). What is... a Free Lunch? *Notices of the AMS*, 51(5):526–528.

[11] Ding, Z., Granger, C. W. J., and Engle, R. F. (1993). A long memory property of stock market returns and a new model. *J. Empir. Finance.*, 1:83–106.

[12] Duncan, T. E., Hu, Y., and Pasik-Duncan, B. (2000). Stochastic calculus for fractional Brownian motion. I. Theory. *SIAM J Control Optim.*, 38(2):582–612.

[13] Fäy, G., Moulines, E., Roueff, F., and Taqqu, M. (2009). Estimators of long–memory: Fourrier versus wavelets. *J. Econ.*, 151:159–177.

[14] Filatova, D. (2008). Mixed fractional Brownian motion: some related questions for computer network traffic modeling. *ICSES 2008 International Conference on Signals and Electronic Systems, Kraków, Poland*, pages 393–396.

[15] Föllmer, H. (1981). Calcul d'Ito sans probabilités. *Séminaire de probabilités (Strasbourg)*,

Tome 15:143–150.
[16] Henry, Ó. T. (2002). Long memory in stock returns: Some international evidence. *Appl. Financial Econ.*, 12(10):725–729.
[17] Higham, D. J. (2001). An algorithm introduction to numerical simulation of stochastic differential equations. *SIAM Review*, 43:525–546.
[18] Hu, Y., Nualart, D., Xiao, W., and Zhang, W. (2011). Exact maximum likelihood estimator for drift fractional Brownian motion at discrete observation. *Acta Mathematica Scientia*, 31B(5):1851–1859.
[19] Hu, Y. and Øksendal, B. (2003). Fractional white noise calculus and applications to finance. *Infin Dimens Anal Quantum Probab Relat Top.*, 6(1):1–32.
[20] Masnita, M., Zudi, L., and Kok Lay, T. (2010). Fractional Black–Scholes models: Complete MLE with application to fractional option pricing. In *Proceedings of the International Conference on Optimization and Control*, pages 573–588, Guiyang, China.
[21] Mishura, Y. (2008). *Stochastic calculus for fractional Brownian motion.* Berlin: Springer Verlag.
[22] Mishura, Y., Ralchenko, K., and Shklyar, S. (April, 2017). Maximum Likelihood Drift Estimation for Gaussian Process with Stationary Increments. *Austrian Journal of Statistics*, 46:67–78.
[23] Nualart, D. and Ortiz-Latorre, S. (2008). Central limit theorems for multiple stochastic integrals and Malliavin calculus. *Stoch. Proc. Appl.*, 118(4):614–628.
[24] Palma, W. (2007). *Long–Memory Time series: Theory and Methods.* Hoboken, New Jersey: Wiley–Interscience.
[25] Petersen, K. B. and Pedersen, M. S. (Version: November 15,2012). *The Matrix Cookbook.*
[26] Robbertse, W. and Lombard, F. (2012). On maximum likelihood estimation of the long–memory parameter in fractional Gaussian noise. *J Stat Comput Simul*, pages 1–14.
[27] Sadique, S. and Silvapule, P. (2001). Long–term memory in stock market returns: International evidence. *Int. J. Finance Econ.*, 6(1):59–67.
[28] Schott, J. R. (2005). *Matrix Analysis for Statistics.* Wiley, Hoboken, NJ.
[29] Shevchenko, G. (2015). Fractional Brownian motion in a nutshell. *International Journal of Modern Physics: Conference Series*, 16.
[30] Sottinen, T. (2001). Fractional Brownian motion, random works and binary market models. *Finance Stoch*, 5:343–355.
[31] Sun, L., Wang, L., and Fu, P. (2018). Maximum likelihood estimators of a long–memory process from discete observations. *Advanced in Differential Equations*, 154.
[32] Tudor, C. A. and Viens, F. (2007). Statistical aspects of the fractional stochastic calculus. *Ann. statist.*, 35:1183–1212.
[33] Wood, A. and Chan, G. (1994). Simulation of stationary Gaussian processes in $[0, 1]^d$. *Journal of computational and graphical statistics*, 4:409–432.
[34] Xiao, W., Zhang, W., and Zhang, X. (2015a). Parameter Identification for Drift Fractional Brownian Motions with Application to the Chinese Stock Markets. *Communications in Statistics – Simulation and Computation*, 44(8):2117–2136.
[35] Xiao, W., Zhang, W., and Zhang, X. (2015b). Parameter identification for the discretly observed geometric fractional Brownian motion. *Journal of Statistical Computation and Simulation*, 85(2):269–283.
[36] Xiao, W. L., Zhang, W. G., and Zhang, X. L. (2011). Maximum-likelihood estimators in the mixed fractional Brownian motion. *Statistics*, 45(1):73–85.
[37] Zhang, P., Sun, Q., and Xiao, W. L. (2014). Parameter identification in mixed Brownian–fractional Brownian motions using Powell's optimization algorithm. *Economic Modelling*, 40:314–319.
[38] Zili, M. (2006). On the mixed fractional Brownian motion. *Journal of Applied Mathematics and Stochastics Analysis*, 2006:1–9.