



# Modeling Stock Market Dynamics with Stochastic Differential Equation Driven by Fractional Brownian Motion: A Bayesian Method

N. Harnpornchai<sup>†,1</sup> and K. Autchariyapanitkul<sup>‡</sup>

<sup>†</sup>Faculty of Economics, Chiang Mai University

e-mail : [napateconcmu@gmail.com](mailto:napateconcmu@gmail.com) (N. Harnpornchai)

<sup>‡</sup>Faculty of Economics, Maejo University

**Abstract :** A Bayesian method is proposed for the parameter identification of a stock market dynamics which is modeled by a Stochastic Differential Equation (SDE) driven by fractional Brownian motion (fBm). The formulation for the identification is based on the Wick-product solution of the SDE driven by an fBm. The determination of the solution is carried out using an independence Metropolis Hastings algorithm. The historical record of SET index is employed for the purpose of method demonstration. For the SET index example, the estimate of the Hurst exponent is approximately 0.5. Consequently, the market is considered efficient.

**Keywords :** Stochastic differential equation; fractional Brownian motion; Bayesian inference; Hurst exponent; stock market dynamics.

**2010 Mathematics Subject Classification :** 62P20; 91B84.

---

## 1 Introduction

The behaviors of stock markets have continuously been a focus of studies in the area of financial risk management. The market behaviors are generally characterized via the market indices. A typical and widely-used model for characterizing

---

Thanks! This research was supported by the National Research Council of Thailand.

<sup>1</sup>Corresponding author.

the dynamics of a market index is in the form of a Stochastic Differential Equation driven (SDE) driven by a standard Brownian motion (sBm) (see e.g. [1]),

$$\frac{dS_t}{S_t} = \mu dt + \sigma dB_t \quad (1.1)$$

$S_t$  is the market index at time  $t$  for  $t \geq 0$ ,  $\mu$  is the mean change rate of the index,  $\sigma$  is the volatility, and  $B_t$  is the sBm.  $\frac{dS_t}{S_t}$  is the index return. An underlying assumption for the validity of model (1.1) is that the index returns are independent, i.e. not temporally correlated. In other words, the index returns have to follow a random walk. When the assumption is violated, the results from employing the model (1.1) can be to a certain extent different from the true market dynamics. An example of such an assumption violation definitely includes the financial data with long-range dependence. The long-range dependence was observed and manifested in different types of assets [2–4]. To mitigate the modeling risk, the random walk behavior should not be assumed a priori without rationale justification. The generalization of model (1.1) can be accomplished by replacing the sBm by a fractional Brownian motion (fBm):

$$\frac{dS_t}{S_t} = \mu dt + \sigma dB_t^H \quad (1.2)$$

$B_t^H$  is the fBm with a Hurst exponent  $H$ , where  $H \in (0, 1)$ . The value of  $H$  determines the dynamic characteristics of the process [5]. For  $H = 0.5$ , the process becomes an sBm. The differences of the process in disjoint intervals are then independent according to the properties of Brownian motion. If,  $0 < H < 0.5$ , the process belongs to the class of anti-persistent or mean-reverting processes. An increment is likely to be followed by a decrement. The anti-persistent intensity increases when  $H$  approaches zero. When,  $0.5 < H < 1$  the process has a persistent characteristic. An increment (decrement) is likely to be followed by an increment (a decrement). In other words, for  $H \neq 0.5$ , the differences of the process are correlated. Thus, the classification of the process in terms of its covariance structure becomes realized using the value of  $H$ .

Several methods have been proposed for the determination of the Hurst exponent. The methods include, for examples, the aggregate variance method [6], the absolute moments method [6], the discrete variations [7], the Higuchi method [8], the periodogram method [9], the variance of the regression residuals [10], the R/S analysis and its variant [11], [12], the variance of the regression residuals [13], the detrended fluctuation analysis [14], the detrended moving average [15], the Whittle method [16], and the wavelet method as well as its variant [17–19]. It should be noted that those methods yield either a point estimator of the Hurst exponent or a confidence interval indicating the Hurst estimates. Recently, a Bayesian method has been proposed for the estimation of the Hurst exponent [20]. The uncertainty of the inferred parameter is naturally taken into account. Accordingly, the Bayesian method can yield both the point estimator and confidence intervals at the same time. It has been also shown in [20] that the Bayesian method results in

more accurate estimation of Hurst exponent synthesized data, compared with the periodogram method, the detrended fluctuation analysis, and the wavelet method.

The present work proposes a Bayesian inference of the parameters in an SDE driven by an fBm. The SDE has the form of (1.2), the application of which can be generally found in the context of financial risk modeling and analysis. The Bayesian method is considered herein due to its afore-mentioned virtues. The contribution of the current work is different from [20] in which the Bayesian method was employed for the purpose of data analysis, i.e. to only infer the Hurst exponent associated with given time series. The evolution of the time series and its modeling are not the subjects of interest in [20]. This work, instead, applies the Bayesian method to the estimation of the parameters in a stochastic dynamic system, based on the prescribed model of an SDE driven by an fBm and on the measurement of the dynamic output, i.e. the index returns. Consequently, not only the Hurst exponent is estimated, but also the stock market dynamics is mathematically modeled, which is definitely beneficial from the view point of financial risk modeling and analysis.

After this introduction, the methodology will be described. The application of the proposed methodology is next illustrated via a numerical example. Finally, the conclusions are made at the end.

## 2 Methodology

### 2.1 Fractional Brownian Motion (fBm)

The fBm was first introduced by Kolmogorov in his work spirals of Wiener [21]. It was later Mandelbrot and Van Ness who provided a stochastic integral representation of the fBm in terms of an sBm [22] and coined the term fractional Brownian motion. The fBm is be defined as follows (confer, for example, [23]):

**Definition 2.1.** Let  $H \in (0, 1)$  and is referred to as a Hurst exponent. An fBm  $B^H(t) = B_t^H$ , where  $t \geq 0$ , is a continuous process with the following properties:

1.  $B_t^H$  is a Gaussian process for  $t \geq 0$ .
2.  $B_0^H = 0$  and  $E[B_t^H] = 0$  for  $t \geq 0$ .
3.  $E[B_t^H B_s^H] = \frac{1}{2}(t^{2H} + s^{2H} - |t - s|^{2H})$  for all  $t \geq 0$  and  $s \geq 0$ .

The term Hurst exponent was named after the hydrologist Hurst who had observed the water run-offs of the Nile river [24] by Mandelbrot.

The solution of the SDE driven by an fBm, based on the Wick product [25], is

$$S_t = S_0 \exp(\mu t - \frac{\sigma^2}{2} t^{2H} + \sigma B_t^H), \quad (2.1)$$

equivalently

$$Y_t = (\mu t - \frac{\sigma^2}{2} t^{2H} + \sigma B_t^H) \quad (2.2)$$

in which

$$Y_t = \ln \left( \frac{S_t}{S_0} \right). \quad (2.3)$$

## 2.2 Formulation for Bayesian Inference

The parameter estimation of the model (1.2) under the framework of the Bayesian inference can be obtained in the form of the posterior distribution from the Bayes theorem. The posterior distribution is given as

$$f(\mu, \sigma, H|D) \propto L(D|\mu, \sigma, H)f(\mu, \sigma, H) \quad (2.4)$$

where  $D$  is the data and is defined as

$$D = (t_1, Y_{t_1}), \dots, (t_N, Y_{t_N}) \quad (2.5)$$

$N$  is the number of data points used in the inference.  $L(D|\mu, \sigma, H)$  is the likelihood and  $f(\mu, \sigma, H)$  is the prior, respectively. The likelihood can be explicitly written as

$$L(D|\mu, \sigma, H) = \frac{1}{(2\pi)^{N/2}|\Sigma|^{1/2}} \exp[-\mathbf{Z}^T \Sigma^{-1} \mathbf{Z}] \quad (2.6)$$

where

$$Z_k = \frac{1}{\sigma} [Y_{t_k} - \mu t_k + \frac{\sigma^2}{2} t_k^{2H}]; k = 1, \dots, N, \quad (2.7)$$

and

$$\Sigma_{ij} = E[B_i^H B_j^H] = \frac{1}{2} (t_i^{2H} + t_j^{2H} - |t_i - t_j|^{2H}); i, j = 1, \dots, N \quad (2.8)$$

## 2.3 Computational Procedure

The determination of the posterior distribution according to Eq.(2.4) can be accomplished using the Markov Chain Monte Carlo (MCMC) algorithms. In this paper, an MCMC algorithms, namely the Metropolis Hastings (MH) will be employed. Let  $\theta = [\theta_1 \dots \theta_{Npar}]$  be a vector of the parameters  $\theta_j (j = 1, \dots, Npar)$ , where  $Npar$  is the total number of the parameters to be estimated. The MH algorithm makes the generation of the samples from a target density function  $\pi(\theta)$  possible. Therefore, if  $\pi(\theta)$  is the posterior distribution, the samples obtained by the MH algorithm will distribute according to the posterior distribution. The MH algorithm requires the specification of a proposal or candidate density  $q(\theta_{l+2}|\theta_l)$ . The general framework of the MH algorithm is as follows [26]. Repeat the following steps for  $l = 1, \dots, M$  where  $M$  is the total number of sampling times.

1. Draw a candidate  $\theta_{l+1}$  from  $q(\theta_{l+1}|\theta_l)$ .

2. Accept  $\theta_{l+1}$  with the probability of  $\alpha(\theta_l, \theta_{l+1})$ , otherwise set  $\theta_{l+1}$  to be  $\theta_l$ .

Then return to step 1.

The probability of acceptance is

$$\alpha(\theta_l, \theta_{l+1}) = \min \left( \frac{\pi(\theta_{l+1})/q(\theta_{l+1}|\theta_l)}{\pi(\theta_l)/q(\theta_l|\theta_{l+1})}, 1 \right) \quad (2.9)$$

The accepted  $\theta_{l+1}$  will form samples whose limiting distribution is  $\pi(\theta)$ . Instead of using the proposal density which depends on the previous candidate, the proposal density can be independent from the previous state, i.e.  $q(\theta_{l+1}|\theta_l) = q(\theta_{l+1})$ . Accordingly, this becomes the so-called independence MH algorithm with the computational steps:

1. Draw  $\theta_{l+1}$  from  $q(\theta_{l+1})$ .
2. Accept  $\theta_{l+1}$  with the probability of  $\alpha(\theta_l, \theta_{l+1})$ , where

$$\alpha(\theta_l, \theta_{l+1}) = \min\left(\frac{\pi(\theta_{l+1})/q(\theta_{l+1})}{\pi(\theta_l)/q(\theta_l)}, 1\right) \quad (2.10)$$

Again, the accepted  $\theta_{l+1}$  will form samples whose limiting distribution is  $\pi(\theta)$ .

### 3 Numerical Example

#### 3.1 Data

The daily closing values of the SET index from 2<sup>nd</sup> of January 2014 to 28<sup>th</sup> of June 2016, altogether 606 data points (see Figure 1).



Figure 1: Daily closing values of the SET index from 2/1/2014 to 28/6/2016.

### 3.2 Computation and Results

The inference applies the following priors:

$$f(\mu, \sigma, H) = f(\mu)f(\sigma)f(H), \tag{3.1}$$

where

$$f(\mu) \sim U(1.00 \times 10^{-4}, 5.00 \times 10^{-4}) \tag{3.2}$$

$$f(\sigma) \sim LOGN(-5.81, 9.98 \times 10^{-4}) \tag{3.3}$$

$$f(H) \sim U(0, 1) \tag{3.4}$$

The uniform distributions are used as the priors for  $\mu$  and  $H$  because there is no information about those parameters, whereas the log-normal prior distribution is taken as the prior for  $\sigma$  to ensure its non-negative property.

Total number of realizations is 20000 with the 10000-realization burn-in. Thus, only 10000 realizations are taken into account for the inference purpose. The histogram corresponding to the marginal posterior distribution of each parameter is shown in Figure2, Figure 3 and Figure 4, respectively.

When applying MCMC, it is required that the convergence to the target density function be validated. Accordingly, the trace plot, running mean plot, and autocorrelation plot of respective inferred parameters are performed for the purpose of the convergence diagnostic.

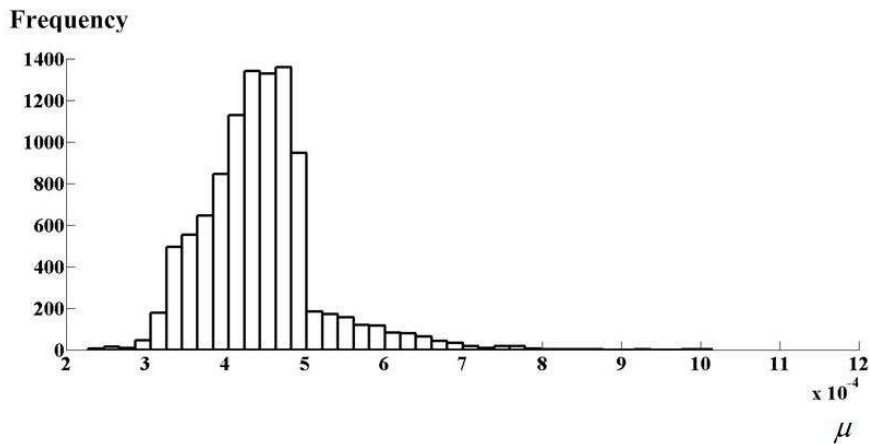


Figure 2: Histogram of  $\mu$  from independence MH with sample size of 10000.

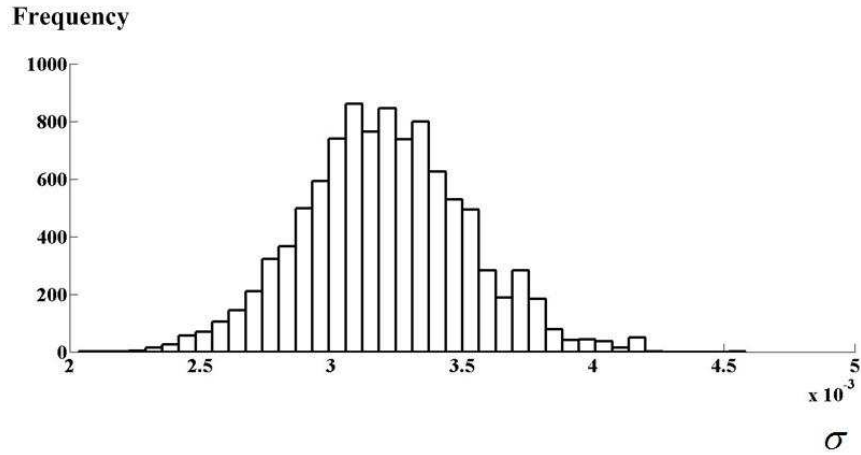


Figure 3: Histogram of  $\sigma$  from independence MH with sample size of 10000.

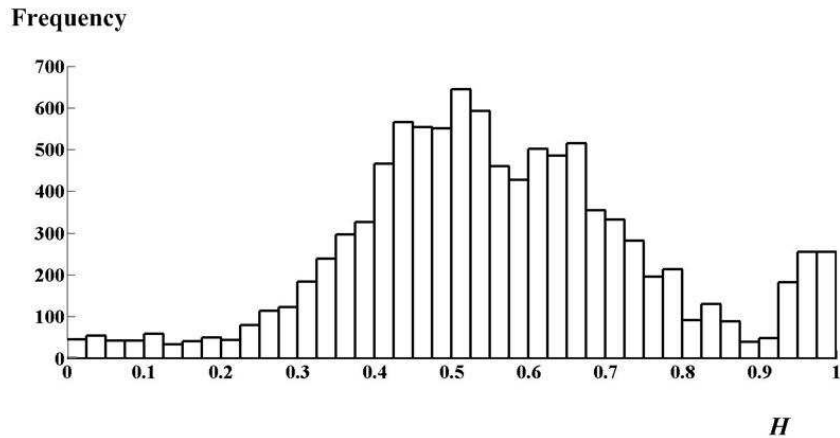


Figure 4: Histogram of  $H$  from independence MH with sample size of 10000.

The results are graphically given in Figure 5, Figure 6, and Figure 7, respectively. The trace plot and running mean plot in all cases indicate the good-mixing behavior of the simulated chain. The autocorrelation associated with each parameter is considerably reducing for higher values of lag. This indicates low degree of correlation among samples. The independency among realizations can be thus accepted.

When the mode is taken as the estimate, the following values are obtained:

$$\hat{H} = 0.51 \tag{3.5}$$

$$\hat{\mu} = 4.736 \times 10^{-4} \tag{3.6}$$

$$\hat{\sigma} = 3.087 \times 10^{-4} \tag{3.7}$$

Since the estimate of  $H$  is approximately 0.5, the driving Brownian motion can be approximately taken as an sBm. Regarding the estimated  $H$ , the market is considered efficient.

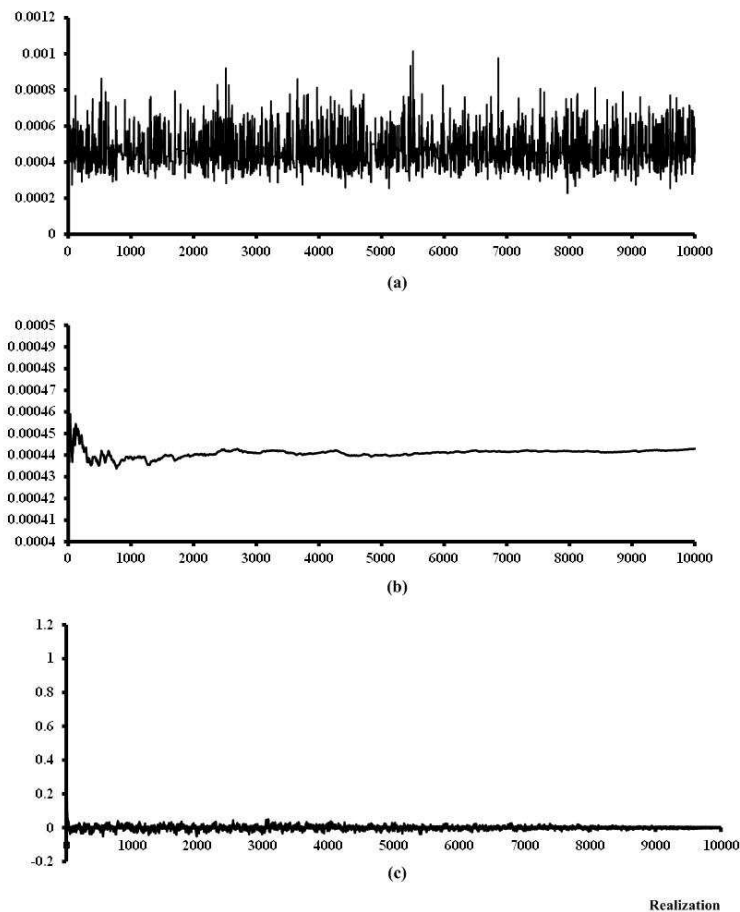


Figure 5: Convergence diagnostic of  $\mu$ : (a) trace plot, (b) running mean plot, (c) autocorrelation plot..



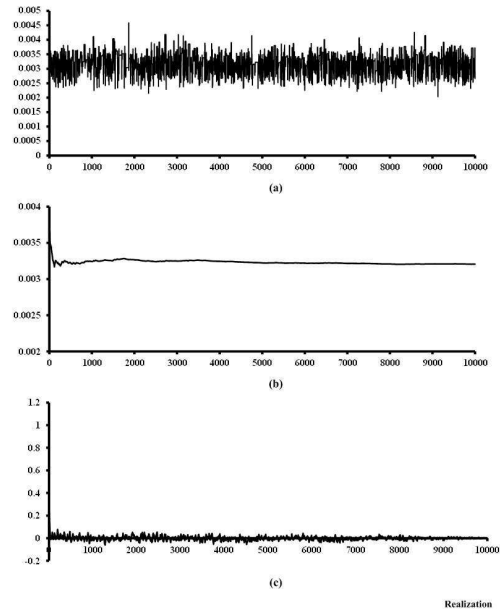


Figure 6: Convergence diagnostic of  $\sigma$ : (a) trace plot, (b) running mean plot, (c) autocorrelation plot.

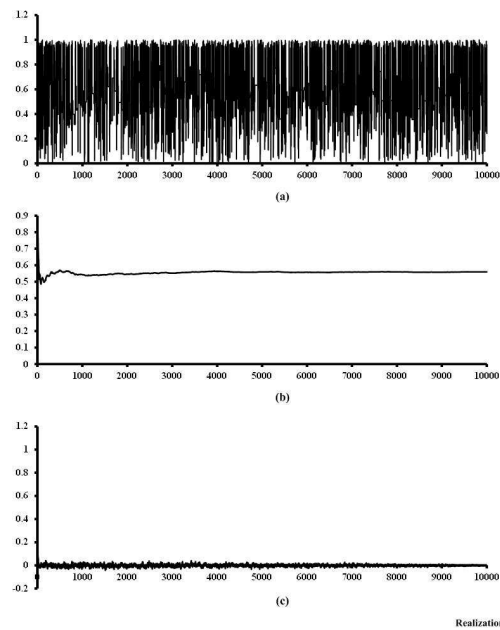


Figure 7: Convergence diagnostic of  $H$ : (a) trace plot, (b) running mean plot, (c) autocorrelation plot.

## 4 Conclusions

A Bayesian method is proposed for the parameter identification of an SDE driven by an fBm. Accordingly, a priori assumption of using the Hurst exponent as 0.5 that can be untrue and erroneous is automatically avoided. MCMC is utilized in order to obtain the numerical values of the identified parameters. The proposed method is demonstrated through the modeling of stock market dynamics in which the SET index is considered. The convergence diagnostic of the identified parameters are satisfactorily verified via the trace plot, running mean plot, and autocorrelation plot.

**Acknowledgements :** The financial support for this research through the Center of Excellence in Econometrics is gratefully appreciated by Napat Harnpornchai.

## References

- [1] P. Alcs, I.M. Jnosi, Modeling the BUX index by a novel stochastic differential equation, *Physica A* 299 (2001) 273-278.
- [2] D.O. Cajueiro, B.M. Tabak, The long-range dependence phenomena in asset returns: the Chinese case, *Appl. Econ. Lett.* 13 (2) (2006) 131-133.
- [3] E. Panas, Estimating fractal dimension using stable distributions and exploring longmemory through ARFIMA models in Athens Stock Exchange, *Appl.Financ. Econ.* 11 (4) (2001) 395-402.
- [4] B. Podobnik, D. Fu, T. Jagric, I. Grosse, H.E. Stanley, Fractionally integrated process for transition economics, *Physica A* 362 (2) (2006) 465-470.
- [5] B.B. Mandelbrot, J.R. Wallis, Robustness of the rescaled range R/S in the measurement of noncyclic long run statistical dependence. *Water Resources Res.* 5 (1969) 967-987.
- [6] M.S. Taqqu, V. Teverovsky, W. Willinger, Estimators for long-range dependence: an empirical study, *Fractals* 3 (1995) 785-798.
- [7] J.F. Coeurjolly, Infrence Statistique Pour Les Mouvements Browniens Fractionnaires Et Multifractionnaires, PhD thesis, Universit Joseph Fourier, Dec. 2000.
- [8] T. Higuchi, Approach to an irregular time series on the basis of the fractal theory, *Physica D* 31 (1988) 277-283.
- [9] J. Geweke, S. Porter-Hudak, The estimation and application of long memory time series models. *Journal of Time Series Analysis* 4 (1983) 221-238.
- [10] C.K. Peng, S.V. Buldyrev, M. Simons, H.E. Stanley, A.L. Goldberger, Mosaic organization of DNA nucleotides, *Physical Review E* 49 (1994) 1685-1689.
- [11] B.B. Mandelbrot, J.R. Wallis, Computer experiments with fractional Gaussian noises. *Water Resour, Res* 5 (1969).

- [12] A.W. Lo, Long-term memory in stock market prices. *Econometrica* 59 (5) (1991) 1279-1313.
- [13] C.K. Peng, S. Havlin, H.E. Stanley, A.L. Goldberger, Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos* 5 (1995) 82-87.
- [14] J.M. Bardet, I. Kammoun, Asymptotic properties of the detrended fluctuation analysis of long range dependent process. *IEEE Transactions on information theory* 54 (5) (2007) 2041-2052.
- [15] S. Arianos, A. Carbone, Detrending moving average algorithm: A closed-form approximation of the scaling law, *Physica A* 382 (2007) 9-15.
- [16] P. Whittle, The analysis of multiple stationary time series. *Journal of the Royal Statistical Society, Series B (Methodological)* 15 (1) (1953) 125-139.
- [17] J.M. Bardet, P. Bertrand, Definition, properties and wavelet analysis of fractional Brownian motion. *Fractals* 15 (2007) 73-87.
- [18] P. Abry, D. Veitch, Wavelet analysis of long-range-dependent traffic. *IEEE Transactions on Information Theory* 44 (1) (1998) 2-15.
- [19] E.J. McCoy, A.T. Walden, Wavelet analysis and synthesis of stationary long-memory processes, *Journal of Computational and Graphical Statistics* 5 (1996) 26-56.
- [20] N. Makarava, S. Benmehdi, M. Holschneider, Bayesian estimation of self-similarity exponent, *Phys. Rev. E* 84 (2) (2011).
- [21] A.N. Kolmogorov, Wiener'sche Spiralen und einige andere interessante Kurven im Hilbertschen Raum, *Doklady Akad. Nauk SSSR* 26 (1940) 115-118.
- [22] B.B. Mandelbrot, J.W. Van Ness, Fractional Brownian motions, fractional noises and applications, *SIAM Rev.* 10 (1968) 422-437.
- [23] F. Biagini, Y. Hu, B. Ksendal, T. Zhang, *Stochastic Calculus for Fractional Brownian Motion and Applications*, Springer-Verlag London Limited, London, (2008).
- [24] H.E. Hurst, Long-term storage capacity reservoirs, *Trans. Amer. Soc. Civil Eng.* 116 (1951) 400-410.
- [25] Y. Hu, B. Ksendal, Fractional White Noise Calculus and Applications to Finance, *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* 6 (2003) 1-32.
- [26] B. Paolo, *Handbook in Monte Carlo simulation: applications in financial engineering, risk management, and economics*, John Wiley & Sons, Inc. Hoboken, New Jersey, (2014).

(Received 3 August 2016)

(Accepted 28 October 2016)