

SCIENTIFIC REPORTS



OPEN

Visual information and expert's idea in Hurst index estimation of the fractional Brownian motion using a diffusion type approximation

Received: 23 August 2016
Accepted: 11 January 2017
Published: 14 February 2017

Ali R. Taheriyoun & Meisam Moghimbeygi

An approximation of the fractional Brownian motion based on the Ornstein-Uhlenbeck process is used to obtain an asymptotic likelihood function. Two estimators of the Hurst index are then presented in the likelihood approach. The first estimator is produced according to the observed values of the sample path; while the second one employs the likelihood function of the incremental process. We also employ visual roughness of realization to restrict the parameter space and to obtain prior information in Bayesian approach. The methods are then compared with three contemporary estimators and an experimental data set is studied.

General motivation

Fractional Brownian motion (fBm) appears in modeling wide classes of non-stationary stochastic processes. The statistical self-similarity and the ability of to fine-tune the order of Hölder continuity are the famous advantages of this process that make it typically one of the greatest interest in modeling natural phenomena. Typically, the creation of the fBm is attributed to ref. 1 since it investigated the basic properties of fBm and stressed its role in modeling of natural phenomena. However, it had been introduced during the generation of Gaussian spirals in Hilbert space². Let $\{B_H(t), t \in \mathbb{R}_+\}$ be an fBm that is a Gaussian zero-mean continuous process with stationary increments and the homogeneous stationary incremental variance function

$$\sigma_{B_H}^2(t) := E|X(t+s) - X(s)|^2 = \sigma^2 |t|^{2H},$$

for any $t, s > 0$. Consequently, the covariance function is

$$\gamma_H(t, s) := E[B_H(t)B_H(s)] = \frac{\sigma^2}{2}(t^{2H} + s^{2H} - |t - s|^{2H}), \quad (1)$$

where $0 < H < 1$, and $\sigma > 0$ are the parameters of the process. σ is a scale parameter and we then assume that $\sigma = 1$. Therefore, the interesting parameter is H , called fractal or Hurst (by Benoit B. Mandelbrot) index, in honor of Harold E. Hurst (for spending 62 years in Egypt carrying out a project on the hydrology of the Nile river) and also in honor of Ludwig O. Hölder.

The features of H are more far-reaching than a parameter of covariance function. For instance, we know that the differentiability of a process indicates the smoothness of the corresponding realizations. According to the last term of (1), fBm is not differentiable in \mathbb{L}^2 , while it satisfies the Hölder continuity of order H and it has derivatives of order α for any $0 < \alpha \leq H$. Thus, an fBm with greater H induces the smoother realizations and this is our general theoretic motivation for considering H as the crucial parameter of the process. Incremental stationary Gaussian processes with a Hölder continuity of order H construct an important class of H -index processes³. The most important feature of this family is the closed form of Hausdorff or fractal dimension for the graph of

Shahid Beheshti University, G.C., Department of Statistics, Tehran, 1983969411, Iran. Correspondence and requests for materials should be addressed to A.R.T. (email: a_taheriyoun@sbu.ac.ir)

Gaussian processes of this family. More clearly, the Hausdorff dimension of an H -index process is equal to $2 - H$. This means that the dimension and consequently the roughness of the fBm is decreasing in H (which is why H is called the fractal index) and stresses the importance of the estimation of H .

A very simple approach to obtaining the ML estimate is considering the likelihood function as the density function of a multivariate normal distribution. According to the last term of the covariance function (1), the joint density function behaves very roughly and is not hoped to employ even numerical methods to obtain the ML estimate of H . For a given index t , we simply use the series decomposition⁴ $B_H(t) = \sum_n Y_n(t)$, such that for each n the summand process $\{Y_n(t)\}_t$ has an Itô integral representation. The choice of Y_n depends on the true value of H . When $H < 1/2$, the summand process is denoted by Y_{1n} and for the case $H > 1/2$, the summand is in the form of $Y_n = Y_{2n} + Y_{3n}$ where Y_{jn} , $j = 1, 2$ and 3 , is defined in the section entitled Decomposition of fBm. We then follow two scenarios in the ML computation: The first one is the simple use of the more computationally compatible covariance functions of the independent processes $\{Y_n(t)\}$, $n = 1, 2, \dots$ and then using the ML estimate of the covariances of Y_n and the invariance property of ML estimates, we compute the ML estimate of H . The second scenario is based on the covariance function of the incremental process $Y_n(t_{i+1}) - Y_n(t_i)$ that is more smooth in comparison with fBm. Again, using the independence of Y_n and $Y_{n'}$ for each $n, n' \in \mathbb{N}$, the likelihood function becomes more comfortable to apply numerical maximization algorithms. We hope also to accelerate the Bayesian computation by the use of this approximation of the likelihood function. The MCMC algorithm is used to capture the posterior information from the multiplication of the approximation of likelihood function by a prior distribution. Meanwhile, the visual information of the observed path of fBm is transferred to the estimation procedure by use of informative prior distributions. For instance, when the sample path shows an obvious short range dependence property, we consider the interval $(0, 1/2]$ as the support of a prior distribution.

Practical motivation

We discussed above the theoretical motivation for estimating H and now we briefly explain the enthusiasm of physicists and environmental scientists for H . Measuring the smoothness is an inseparable part of almost all surface analysis investigations. Due to the *power* and *scaling* laws, a variety of high resolution observations has been studied as a realization of fractal random fields. The Lévy fractional Brownian sheet (LFBs) of fractal index H is usually appointed to represent fractal behavior. The zero-mean Gaussian random field $L_H(\mathbf{t})$ is an LFBs of index H if

$$E[L_H(\mathbf{t})L_H(\mathbf{s})] = \frac{\sigma^2}{2}(\|\mathbf{t}\|^{2H} + \|\mathbf{s}\|^{2H} - \|\mathbf{t} - \mathbf{s}\|^{2H}), \quad (2)$$

for $\mathbf{t}, \mathbf{s} \in \mathbb{R}_+^d$. There are many methods to estimate the fractal index H based on the d -dimensional data of realization, but using almost all of them requires both computational and technical skills. A discussion on this subject could be found in ref. 5 and the references therein. Using the line transect data, we hope to decrease the computational cost of estimating H by reducing the dimension of data which tends to estimating H based on one-dimensional processes. For clarity, suppose that a picture has been taken of a surface ($d = 2$) and the observation is gathered in an $r \times c$ matrix. We are looking for an estimation of H that uses the data as c samples of length r . The question is: ‘is it possible to consider the data matrix as c sample paths from an fBm of index H ?’ To answer this question, note that the Hausdorff dimension of the graph of $L_H(\mathbf{t})$ is equal to $d + 1 - H$. Extract the simple one dimensional stochastic process $\{Z(t_i); t_i \in \mathbb{R}_+\}$ from $L_H(\mathbf{t})$, $\mathbf{t} = (t_1, \dots, t_d)$, by letting the $d - 1$ number of the elements of $\mathbf{t} \in \mathbb{R}_+^d$ to be fixed. Let σ_Z^2 be the incremental variance of Z , then

$$\lim_{|t_i| \rightarrow 0} \frac{\sigma_Z(t_i)}{|t_i|^H} = c',$$

where $c' > 0$ does not depend on \mathbf{t} , and hence Z is a one-dimensional index- H process³. Thereupon, it is helpful to look at the line-transect data of the realizations of LFBs when the aim of the study involves only the estimation of the fractal index and/or measuring surface roughness.

The estimators of H have often been made for estimating the roughness of surfaces using the line transect data. Several methods have been proposed to estimate H , the oldest among them is R/S-statistic introduced by ref. 6. Further results for this estimator were provided by ref. 7. In ref. 8 a robust estimation of H has been introduced for stable distributions. It has been shown that this estimator is not efficient in comparison with the ML estimators in the Gaussian case. Perhaps the first serious attempt to make the ML estimate in image textures is provided by ref. 9. One may find a Bayesian solution for the same problem in ref. 10. Using the asymptotic behavior of the k -th absolute moment of discrete variations of a sampled path, a class of consistent estimators of the Hurst index has been constructed¹¹. Based on the Bahadur representation for sample quantiles of nonlinear transformations of Gaussian processes, another consistent estimator of the Hurst index for the general class of locally self-similar Gaussian processes has been presented¹². Generally, there is a class of estimators based on a method, called discrete variations, in which the oscillation of a quantity is employed in the estimation procedure. In statistical literature, refs 13 and 14, simultaneously originated this method and it has been highlighted many times in the other works^{11,12,15}. There also exists a consistent estimator of H , based on the Karhunen-Loève expansion of a Gaussian process¹⁶. We compare our estimators with the estimators of refs 11, 12 and 15.

Methods

Decomposition of fBm. A stochastic integral representation of B_H has been introduced¹⁷ as follows:

$$B_H(t) = \int_0^t K_H(t, s) dB(s), \quad t \geq 0, \tag{3}$$

where $\{B(t), t \geq 0\}$ is the standard Brownian motion and

$$K_H(t, s) = c_H s^{1/2-H} \int_s^t |u - s|^{H-3/2} u^{H-1/2} du,$$

when $1 > H > 1/2$ and

$$K_H(t, s) = b_H \left[\left(\frac{t(t-s)}{s} \right)^{H-1/2} - \left(H - \frac{1}{2} \right) s^{1/2-H} \int_s^t (u-s)^{H-1/2} u^{H-1/2} du \right],$$

for $0 < H \leq 1/2$ where

$$c_H = [H(2H - 1)/\text{Beta}(2 - 2H, H - 1/2)]^{1/2},$$

$$b_H = [(1 - 2H)\text{Beta}(1 - 2H, H + 1/2)/(2H)]^{-1/2}.$$

The mentioned representation in (3) induces the idea of achieving a diffusion process based on the stochastic integration. Since the solution of an Itô integral is a semi-martingale and since B_H , as the solution of the stochastic integral in (3), is not a semi-martingale for $H \neq 1/2$, the idea of constructing an exact diffusion representation is canceled. On the other hand, there is a series of diffusion processes, $\sum_{n=1}^N Y_n(t)$, such that it converges⁴ to $B_H(t)$. The process $\{Y_n(t)\}$ is the answer of an Itô integral and strongly depends on the value of H . We would like to compute the likelihood function using the joint distribution of the diffusion processes $\{Y_n(t)\}_t$ but for finite N . Precisely, the series decomposition for $B_H(t)$ is obtained from the following convergences:

$$\sum_{n=1}^N Y_{1n}(t) := \sum_{n=1}^N \int_{-\infty}^{\log t} \alpha_{1n} t^{H-\beta_n} e^{\beta_n s} dB^{(n)}(s) \xrightarrow{\mathbb{L}^2} B_H(t), \tag{4}$$

for $0 < H \leq 1/2$ and

$$\begin{aligned} \sum_{n=1}^N (Y_{2n}(t) + Y_{3n}(t)) &= \sum_{n=1}^N \alpha_{2n} \int_{-\infty}^{\log t} (e^{Hs} - t^{H(1-\gamma_{2n})} e^{\gamma_{2n} Hs}) dB^{(n)}(s) \\ &\quad + \sum_{n=1}^N \alpha_{3n} \int_{-\infty}^{\log t} (t^{2H-1} e^{(1-H)s} - t^{H-(1-H)\gamma_{3n}} e^{\gamma_{3n}(1-H)s}) dB^{(n)}(s) \\ &\xrightarrow{\mathbb{L}^2} B_H(t), \end{aligned} \tag{5}$$

for $1/2 < H < 1$, as $N \rightarrow \infty$ where $B^{(n)}$ and $B'^{(n)}$ are independent sequences of independent Brownian motions. Also,

$$\begin{aligned} \alpha_{1n}^2 &= (-1)^n \sigma^2 \binom{2H}{n-1} (n-1-H), \\ \alpha_{2n}^2 &= \frac{\sigma^2}{3-2H} (-1)^{n+1} \binom{2H}{n+1} \frac{(n+1)(n+1-H)}{n+1-2H}, \\ \alpha_{3n}^2 &= \frac{3\sigma^2(1-H)}{3-2H} (-1)^{n+1} \binom{2H}{n+1} \frac{(n+2-2H)(n+1-H)}{n}, \end{aligned}$$

where $\binom{b}{a} \stackrel{\text{def}}{=} b(b-1) \dots (b-a+1)/a!$ and $\beta_n = |n-H-1|$. Clearly, for given $t \in (0, 1)$ and $n \in \{1, \dots, N\}$, $Y_{2n}(t)$ and $Y_{3n}(t)$ are defined as

$$Y_{2n}(t) = \alpha_{2n} \int_{-\infty}^{\log t} (e^{Hs} - t^{H(1-\gamma_{2n})} e^{\gamma_{2n} Hs}) dB^{(n)}(s)$$

$$Y_{3n}(t) = \alpha_{3n} \int_{-\infty}^{\log t} (t^{2H-1} e^{(1-H)s} - t^{H-(1-H)\gamma_{3n}} e^{\gamma_{3n}(1-H)s}) dB'^{(n)}(s).$$

Based on the mentioned convergences in (4) and (5), we would like to construct the likelihood function by applying the Euler discretization on the diffusion processes. In the sequel denote the global parameter space $(0, 1)$ by Θ and denote the restricted parameter spaces $(0, 1/2]$ and $(1/2, 1)$ by Θ_1 and Θ_2 , respectively. One may suggest simpler approximations¹⁸ where the fBm is approximated using a semi-martingale process but the approximation does not have a diffusion-type representation. The presented approach strongly depends on this representation and therefore substitution of the approximation mentioned by ref. 4 with other approximations may lack the procedure of likelihood computation.

Estimation of H : simple use of convergences. Assume that a sample path of $B_H(t)$ is observed at discrete times $t_i \in [0, 1)$ for $i = 1, \dots, m$, where m is not a prime number. The restriction of time index into the interval $t \in [0, 1)$ is somehow a loss of generality. But, using the self-similarity of the fBm, it is possible to rescale the sample path into this interval. In fact, this assumption is important in computing the variance of Y_{jn} , for $j = 1, 2$ and 3 . Moreover, only for this subsection let $m = kp$, where $k, p \in \mathbb{N}$. A very elementary approach is to consider the $B_H(t_i)$ as a summation of independent Gaussian random variables.

Proposition 1. Let $\lambda_m(t, s)$ denotes $cov(Y_m(t), Y_m(s))$ for $i = 1, 2, 3$, and $n = 1, \dots, N$. For the fBm, B_H , the covariance between two arbitrary points $0 < s < t$ satisfies

$$\sum_{n=1}^N \lambda_{1n}(t, s) \rightarrow cov(B_H(t), B_H(s)), \tag{6}$$

uniformly on t and s as $N \rightarrow \infty$ when $H \in \Theta_1$. The convergence for $H \in \Theta_2$ is

$$\sum_{n=1}^N \sum_{l=2}^3 \lambda_{ln}(t, s) \rightarrow cov(B_H(t), B_H(s)), \tag{7}$$

where

$$\begin{aligned} \lambda_{2n}(t, s) &= R_2(t, s; \alpha_{2n}, \gamma_{2n}, H) \\ &= \frac{\alpha_{2n}}{2H} \left(s^{2H} + s^{H(1+\gamma_{2n})} \left(\frac{t^{H(1-\gamma_{2n})}}{\gamma_{2n}} - \frac{2}{H(1+\gamma_{2n})(t^{H(1-\gamma_{2n})} + s^{H(1-\gamma_{2n})})} \right) \right), \end{aligned}$$

and

$$\lambda_{3n}(t, s) = (ts)^{2H-1} R_2(t, s; \alpha_{3n}, \gamma_{3n}, 1 - H).$$

The proof of this proposition is grouped with all the other proofs in the supplementary file.

For a large enough N , the left-hand sides of (6) and (7) return an approximation of the covariance function of the fBm in terms of the mentioned series decomposition. We now consider the sample path as a dependent sample of size k , from a p -variate zero-mean normal random vector with covariance matrix $\Sigma = \sum_n [\lambda_m(t_i, t_j)]_{ij}$ for appropriate $l = 1, 2$, where the elements are computed with respect to Proposition 1. There are various methods to compute the ML estimation of H , called \widehat{H} , based on the approximation of likelihood. Only for this subsection assume that the sample path is observed regularly on $t_i = i/m$. The first method begins with the ML estimator of the autocovariance of the incremental stationary process with observations $\Delta B_H(i/m) = B_H((i+1)/m) - B_H(i/m)$ for $i = 1, \dots, m-1$. Then, we compute the ML estimate of the covariance matrix of the p -dimensional random vector $(\Delta B_H(i/m), \dots, \Delta B_H((i+p-1)/m))$ where the estimated matrix does not depend on i . This covariance matrix consists of p estimated autocovariances at lags $0, 1/m, \dots, (p-1)/m$. Using the invariance property of ML estimators one may recover \widehat{H} from each component of this matrix. The resulted estimates depend on the selected lag to recover the H ; thus, based on this observation we may have p different values for the ML estimate of H . Simply speaking, if $\widehat{\gamma}(0), \dots, \widehat{\gamma}((p-1)/m)$ is the estimated autocovariances at the mentioned lags, then the \widehat{H} 's are the solutions of nonlinear equations

$$\begin{aligned} \widehat{\gamma}(i/n) &= cov\left(B_H\left(\frac{j+i+1}{m}\right), B_H\left(\frac{j+1}{m}\right)\right) - cov\left(B_H\left(\frac{j+i+1}{m}\right), B_H\left(\frac{j}{m}\right)\right) \\ &\quad - cov\left(B_H\left(\frac{j+i}{m}\right), B_H\left(\frac{j+1}{m}\right)\right) + cov\left(B_H\left(\frac{j+i}{m}\right), B_H\left(\frac{j}{m}\right)\right), \end{aligned} \tag{8}$$

for $i = 0, \dots, p-1$ and the equations do not depend on j . The covariances are replaced from the left hand side of (6) or (7) when $H \in \Theta_1$ or $H \in \Theta_2$, respectively. This equation produces p different values for the ML estimate of H , denoted by $\widehat{H} = (\widehat{H}_1, \dots, \widehat{H}_p)$, and one may use the average of all the resulting solutions of (8) for various values of i to obtain a new consistent estimator. To this end, we propose the estimator

$$\widehat{H}_c = \sum_{i=1}^p w(i) \widehat{H}_i, \tag{9}$$

where w is an appropriate weight function. The idea of using the weighted average comes from the numerical properties of \widehat{H}_i . Numerically, those \widehat{H}_i 's that correspond to the larger values of i return better estimates for H . It is worth mentioning that, although the larger values of H eventuate smoother realizations for the fBm, the error of approximation becomes larger for $H > 1/2$. Hence, the conjecture of better results for smoother realizations may not be true which is due to the weakness of convergence when $H > 1/2$.

Estimation of H : Incremental approach. The approximation of the covariance function of fBm induces the idea of maximizing the likelihood function with respect to H , directly. The new problem is the complicated form of the inverse and determinant of the approximated covariance matrix. Since the process of increments of fBm is stationary, we expect more explicit form of the relationship between H and the covariances for the

incremental process. The increments of process $\{\sum_{n=1}^N Y_{1n}(t)\}$ do not constitute a stationary process; however, for large enough N , they converge in \mathbb{L}^2 to the stationary incremental process of fBm. Using the second approach, we see that the appearance of H in the likelihood becomes more suitable via the smoothing effect of difference operator. This may accelerate the numerical methods of the maximization of likelihood.

Let \mathbf{B}_H denotes the vector of sample path $(B_H(t_1), \dots, B_H(t_m))^T$. Also, let \mathcal{F}_t be the standard filter generated by the Brownian motion that is the smallest σ -field induced by $\{B^{(n)}(s), s \leq t\}$. We attempt to figure out the finite-dimensional distribution of increments $\Delta B_H(t_i) = B_H(t_{i+1}) - B_H(t_i)$, for $i = 1, \dots, m - 1$ and hence the likelihood function. The random variables $Y_{1n}(t_i)$ and $Y_{1n}(t_{i+1})$ are $\mathbb{L}^2(\Omega, \mathcal{F}_{t_{i+1}}, P)$, for $n \in \mathbb{N}$ and $\{B_H(t)\}$ is a quadratic variation process. Thus, we get the convergence

$$\sum_{n=1}^N (Y_{1n}(t_{i+1}) - Y_{1n}(t_i)) \xrightarrow{\mathbb{L}^2} B_H(t_{i+1}) - B_H(t_i),$$

as $N \rightarrow \infty$ for $0 < H \leq 1/2$. Convergence in \mathbb{L}^2 implies the convergence in distribution and we then look at the increments as the limit of the bi-indexed stochastic process $\{\sum_{n=1}^N (Y_{1n}(t_{i+1}) - Y_{1n}(t_i))\}_{N,i}$. The following Theorem helps us to approximate the distribution of the increments of a sample path of fBm with Hurst index $H \leq 1/2$.

Theorem 1. Suppose that \mathbf{B}_H is a sample path of an fBm with fractal index $H \in \Theta_1$. The incremental process $\{\Delta B_H(t_i), i = 1, \dots, m - 1\}$ is the limit of a bi-indexed zero-mean Gaussian process $\{\xi_{t,N}; t \in \mathbb{R}_+, N \in \mathbb{N}\}$ with the non-stationary covariance function

$$\begin{aligned} \sigma_{t_i t_j} &:= \text{cov}(\xi_{t_i, N}, \xi_{t_j, N}) \\ &:= \text{cov}\left(\sum_{n=1}^N (Y_{1n}(t_{i+1}) - Y_{1n}(t_i)), \sum_{n=1}^N (Y_{1n}(t_{j+1}) - Y_{1n}(t_j))\right) \\ &= \sum_{n=1}^N \frac{\alpha_{1n}^2}{2\beta_n} \Delta(t_i^{H-\beta_n}) \Delta(t_j^{H+\beta_n}), \end{aligned} \tag{10}$$

for any $t_j < t_i$ as $N \rightarrow \infty$.

Let $\sum_{\text{inc}} = [\sigma_{t_i t_j}]_{i,j=1}^{m-1}$ be the covariance matrix of the increments vector $\mathbf{D} = (\Delta B_H(t_1), \dots, \Delta B_H(t_{m-1}))^T$ and thus the likelihood function is

$$L_1(H|\mathbf{B}_H) := |\sum_{\text{inc}}|^{-1/2} \exp\left(-\frac{1}{2} \mathbf{D}^T \sum_{\text{inc}}^{-1} \mathbf{D}\right), \tag{11}$$

for $H \in \Theta_1$. So, the ML estimate of Hurst index, called \widehat{H}_{inc} , can be obtained by

$$\widehat{H}_{\text{inc}} = \arg \max_{H \in \Theta_1} L_1(H|\mathbf{B}_H). \tag{12}$$

The same approach is employed to calculate the ML estimate of fractal index when the parameter space is Θ_2 and we denote the induced likelihood function by $L_2(H|\mathbf{B}_H)$.

Theorem 2. Suppose that \mathbf{B}_H is a sample path of an fBm with fractal index $H \in \Theta_2$. The incremental process $\{\Delta B_H(t_i), i = 1, \dots, m - 1\}$ is the \mathbb{L}^2 limit of a bi-indexed zero-mean Gaussian process $\{\xi'_{t,N}; t \in \mathbb{R}_+, N \in \mathbb{N}\}$ with the non-stationary covariance function as same as the process

$$\left\{ \sum_{n=1}^N (\alpha_{2n} [Z_{1n}(t_i) + Z_{2n}(t_i)] + \alpha_{3n} [Z_{3n}(t_i) + Z_{4n}(t_i)]), i = 1, \dots, m - 1 \right\},$$

as $N \rightarrow \infty$ where $(Z_{1n}(\cdot), \dots, Z_{4n}(\cdot))$, $n = 1, \dots, N$ are independent zero-mean Gaussian random vectors and the covariance elements are given through the equations (22)–(27) in the supplementary file.

Therefore, it suffices to update the elements of $\sum_{\text{inc}} = [\sigma_{t_i t_j}]$ in (11) according to Theorem 2 and consequently

$$\widehat{H}_{\text{inc}} = \arg \max_{H \in \Theta_2} L_1(H|\mathbf{B}_H).$$

A criticism is the complicating form of the resulting likelihood as a function of H . One may claim to compute the exact likelihood using (1). Note that the direct use of (1) yields an in-differentiable function of H ; while for $H > 1/2$, β_n does not appear in the approximation of the likelihood function. Thus, the target function for maximization belongs to the class of $C_{(0,1)}^1$ functions and is smooth enough for using numerical methods in computational procedures.

Bayesian discussion. The approximation of the likelihood function depends on whether $H \leq 1/2$ or not. According to (1), the fBm is short-range dependent when $H \in \Theta_1$ and it becomes a long-range dependent process

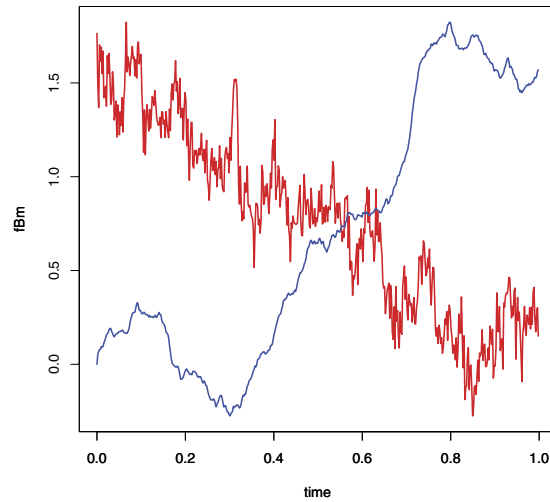


Figure 1. Two realizations of fBm with Hurst indexes $H = 1/4$ (red line) and $H = 3/4$ (blue line). One of the figures are translated through the vertical axis to demonstrate both the realizations in the same Cartesian system.

as well as H crosses the threshold $1/2$. Thus, the realization of fBm for $H \leq 1/2$ has a rough graph; while it has a low oscillation with a polynomial trend when $H \in \Theta_2$. This visually difference in the short-range and long-range dependent fBm's may help us in constraining the parameter space to accelerate the calculation of ML estimate or to employ priori information for Bayesian inference. Based on the type of dependence, we constrain the parameter space, Θ , into one of the sub-intervals Θ_1 or Θ_2 . Consequently, the class of prior distributions is categorized into the set of distributions with support on Θ_1 or Θ_2 .

The procedure of using a priori information is quite simple. We need a likelihood function as the joint distribution of observations given the unknown parameter H and a prior distribution. The likelihood function could be explained in terms of the sample path \mathbf{B}_H or its increments, \mathbf{D} . Simply speaking, one can obtain the likelihood function in terms of \mathbf{B}_H by employing Proposition 1 or calculate the likelihood in terms of increments by Theorems 1 and 2. Both methods express the likelihood function for $H \leq 1/2$ and $H > 1/2$, separately.

Let $L_1(\cdot|\mathbf{B}_H)$ and $L_2(\cdot|\mathbf{B}_H)$ denote the likelihood functions in terms of \mathbf{B}_H for $H \in \Theta_1$ and $H \in \Theta_2$, respectively (just like the notation which used in (11)) and define $L_1(\cdot|\mathbf{D})$ and $L_2(\cdot|\mathbf{D})$ in the same way. Let $f_D(\cdot|H)$ denotes the sample distribution of increments that is rewritten as

$$f_D(\cdot|H) = \begin{cases} L_1(H|\cdot)I_{\Theta_1}(H) + L_2(H|\cdot)I_{\Theta_2}(H) & 0 < H < 1, \\ 0 & \text{otherwise,} \end{cases} \quad (13)$$

where $I_A(\cdot)$ is the indicator function. Although there is no conjugate prior distribution, reducing the parameter space may change the mixture likelihood into the simple multivariate normal. Experts opinions and conjectures on the sample path produce the prior information and may reduce the parameter space. For instance, consider a realization of an fBm with fractal index $H = 1/4$ which the short-range dependence property is very obvious (Fig. 1).

In this case, the expert opinion hopefully suggests to consider H in the interval $(0, 1/2]$ and so the second term in (13) is omitted. According to the bounded parameter space, we deduce the use of $1 - 1/2$ Beta(a, b) and $1/2$ Beta(a, b) as a priori distributions for long and short-range dependent sample paths, respectively. When the determination of the range is not possible by visual evidences, using a prior distribution on $(0, 1)$ inevitably yields to working with the mixture form of (13). In this situation, we use the Beta(a, b) distribution as the prior. Moreover, the suggested priors could be replaced by the uniform distribution as a non-informative prior. The rest is using good candidates for the proposal distributions and performing the MCMC algorithm to report the posteriori issues. The implementation and numerical results of the Bayesian approach are discussed in the following section.

Results

We separate this section into three main parts. First, the numerical behavior of the likelihood based estimator is studied and then we explain the procedure of Bayesian approach. The third part contains the computation of estimators for a real data.

Likelihood computation. The consistency of the estimators strongly depends on the mean squared error (MSE) of the approximation of fBm. The structure of the two dimensional Brownian motions in the case $H > 1/2$ could be taken into account to effectively express the terms of series representation. Although the sample path is smoother in the case $H > 1/2$, the MSE of the approximation of the fBm is greater in comparison with the case $H \leq 1/2$. There is another approximation for fBm that is more effective⁴ than (5) when $H \in \Theta_2$. However, it is

	\widehat{H}_c	\widehat{H}_{inc}	DV	SQ	TM	Bayesian	
						Uniform prior	Beta type prior
$H=0.05$	13.693	9.633	15.2544	37.7292	18.162	4.566	4.817
0.1	13.631	10.838	17.0344	42.5174	24.017	4.855	4.706
0.2	16.124	11.996	24.8491	47.2209	33.702	4.519	5.440
0.3	18.649	10.488	23.8394	57.5849	28.419	6.259	4.844
0.4	16.653	10.608	34.6774	61.1165	36.690	6.585	4.711
0.45	15.638	9.560	30.3553	55.3652	40.734	4.556	5.387
0.55	14.217	12.635	40.8556	62.8026	42.061	5.356	5.067
0.6	14.612	11.469	38.5361	65.9654	47.568	8.152	5.079
0.7	15.130	11.312	36.7738	55.5099	53.171	8.625	4.785
0.8	14.824	10.607	34.7035	59.9203	42.661	11.408	5.242
0.9	13.980	9.981	30.2872	71.0405	44.464	13.690	4.733
0.95	13.963	10.120	42.3179	59.2273	58.332	12.859	5.004

Table 1. The MSE of the Bayesian estimators, \widehat{H}_c and \widehat{H}_{inc} under the restricted parameter space along the MSE of the estimators based on discrete variation (DV¹¹), sample quantiles (SQ¹²) and trimmed mean (TM¹⁵) under the general parameter space are given. The results are multiplied by 10⁴.

difficult to obtain a diffusion type process using that approximation. Therefore, we continue the simulation study with the performed likelihoods in previous sections.

Using the circulant embedding method¹⁹, we first generate the sample path $\mathbf{b} = (b_1, \dots, b_m)^T$ of size $m = 256$ from an fBm with known index H . We only implement the simulation study for regular observations. The likelihood function given the observed sample path is

$$L(H|\mathbf{b}) \propto |\Sigma_l|^{-1/2} \exp\left(-\frac{1}{2}\mathbf{b}^T \Sigma_l^{-1} \mathbf{b}\right), \tag{14}$$

where $\Sigma_l = \sum_{n=1}^{256} [\lambda_{ln}(i, j)]_{ij=1}^{256}$, when the generation is implemented according to $H \in \Theta_p, l = 1, 2$. The use of well-known optimization algorithms becomes very difficult due to the complicating form of the likelihood as a function of H . The solutions of equation system (8) for a sample path with $H = 0.1$ and $p = 4$ are located at the vector of estimators $\widehat{\mathbf{H}} = (0.0789, 0.1092, 0.0920, 0.1080)$ with respect to $(\widehat{\gamma}(0), \widehat{\gamma}(1/256), \widehat{\gamma}(2/256), \widehat{\gamma}(3/256))$. It can be seen that the estimators based on $\widehat{\gamma}(i/m)$ become more precise for larger values of i . In one hand, the replication of this simulation study shows less bias but greater variance for greater values of i . On the other hand, we found less variance but greater bias for estimators based on $\widehat{\gamma}(i/m)$ when i is small. This is the basis of using the weighted estimator (9). Although one may use the cross validation method to construct a better weight function, for this simulation study we partition off the vector $\widehat{\mathbf{H}}$ into two vectors with $p/2$ elements and the following weight function is considered for this simulation study

$$w(i, j) = \begin{cases} 1/(3p) & \text{if } i \leq p/2, \\ 2/(3p) & \text{if } i > p/2. \end{cases}$$

The second column of Table 1 shows the MSE of \widehat{H}_c when the unknown parameter H is attributed to one of the intervals $(0, 1/2]$ or $(1/2, 1)$. Each reported MSE in this paper is computed by replication of estimation procedure for 100 times. Figure 2 shows the position of the ML estimator and \widehat{H}_c based on the autocovariance estimation at lags $0, \dots, p - 1$ for $p = 4, 8$, and 16 . Obviously, \widehat{H}_c becomes closer to the ML estimate by decreasing p . Although the presented estimators \widehat{H}_c defined by (9) are not ML estimates, they are near to the ML estimates computed from the approximated likelihood function of the increments. Moreover, \widehat{H}_c numerically dominates all the simple ML estimates \widehat{H}_i in (9) for $p = 4$.

To implement the incremental approach, construct the sample increments $\mathbf{d} = (d_1, \dots, d_{256})^T$ where $d_j = b_j - b_{j-1}$, for $j = 1, \dots, 257$. The likelihood function of the observed increments is then

$$L(H|\mathbf{d}) \propto |\Sigma_{inc}|^{-1/2} \exp\left(-\frac{1}{2}\mathbf{d}^T \Sigma_{inc}^{-1} \mathbf{d}\right),$$

where $\Sigma_{inc} = [\sigma_{ij}]_{i,j=1}^{256}$ is computed according to (10) when the generated \mathbf{b} is a sample path of fBm of index $0 < H \leq 1/2$. Fortunately, the smoothing role of differentiation operator makes it possible to maximize the likelihood by employing the simple Particle swarm optimization (PSO) via R package `psoptim`. The implementation of this procedure for $1/2 < H < 1$ is not as simple as before. For convenience, let us denote the right hand sides of (A.5)–(A.8) in the supplementary file by $A_5(t, s) - A_8(t, s)$, and the right hands of (A.9) and (A.10) by $A_9(t)$ and $A_{10}(t)$, respectively. According to the Theorem 2, the sub-diagonal elements $\sigma_{ij}, j < i$ for $1/2 < H < 1$ is obtained by

$$\sigma_{ij} = \sum_{n=1}^N \alpha_{2n}^2 (A_5(i, j) + A_6(i, j)) + \alpha_{3n}^2 (A_7(i, j) + A_8(i, j)),$$

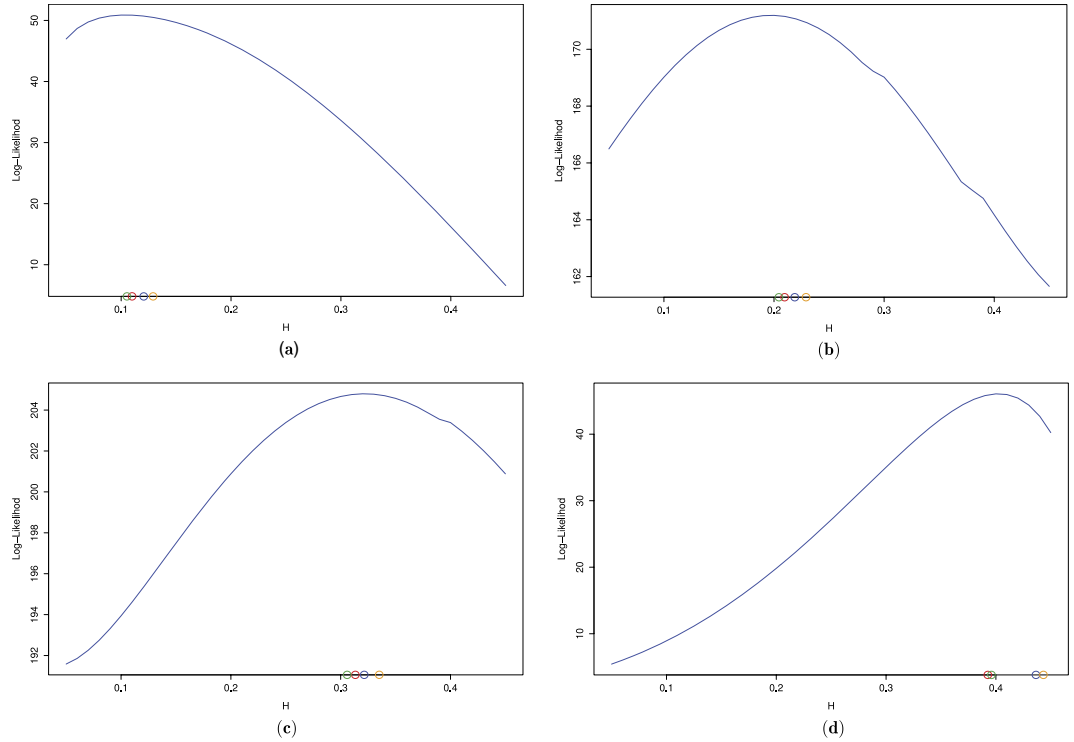


Figure 2. The logarithm of the worst approximated likelihood function based on the incremental approach for sample paths generated with $H = 0.1$ (top left, **a**), $H = 0.2$ (top right, **b**), $H = 0.3$ (bottom left, **c**) and $H = 0.4$, (bottom right, **d**). The plots also demonstrate the resulting \widehat{H}_{inc} (green circle,) for the generated sample path along the resulting \widehat{H}_c 's for $p = 4$ (red circle,), 8 (blue circle,), and 16 (orange circle,).

for $i = 2, \dots, 256, j = 1, \dots, i - 1$ and the top-diagonal components are computed using the symmetric behavior of Σ_{inc} . Furthermore, the diagonal elements, σ_{ii} , are computed as follows:

$$\sigma_{ii} = \sum_{n=1}^N \alpha_{2n}^2 (A_5(i, i) + A_9(i)) + \alpha_{3n}^2 (A_7(i, i) + A_{10}(i)),$$

for $i = 1, \dots, 256$. The covariance matrix Σ_{inc} is now acquired and the `nlm` could find the maximum of likelihood function, numerically. The results of this procedure is shown in the 3rd column of Table 1.

In the procedure mentioned above we have known that H belongs to Θ_1 or Θ_2 . When there is no such information about the parameter, the likelihood function compensates this uncertainty. Let us explain this part for \widehat{H}_c . The likelihood function for this problem is

$$\begin{aligned} L(H|\mathbf{b}) &= |\Sigma_1|^{-1/2} \exp\left(-\frac{1}{2}\mathbf{b}^T(\Sigma_1)^{-1}\mathbf{b}\right) I_{(0,1/2)}(H) \\ &\quad + |\Sigma_2|^{-1/2} \exp\left(-\frac{1}{2}\mathbf{b}^T(\Sigma_2)^{-1}\mathbf{b}\right) I_{(1/2,1)}(H) \\ &= L_1 + L_2, \end{aligned} \tag{15}$$

where $\Sigma_l = \sum_{n=1}^{256} [\lambda_{ln}(i, j)]_{ij=1}^{256}$, for $l = 1, 2$. Thus, $\max_H L = \max\{\max_H L_1, \max_H L_2\}$ that is equivalent to the maximization of L_1 and L_2 separately and then we pick out the one with the greater value. In this way, we allocate the unknown parameter to one of the intervals along the estimation procedure. We can simply modify this method for the incremental approach by replacing Σ_l by $\Sigma_{inc}^{(l)}$, that is the covariance matrix of increments when $H \in \Theta_l$ for $l = 1, 2$. The results are gathered in the second and third columns of Table 2 and in comparison with Table 1, the MSE of estimators are greater and this shows the effect of more precise parameter spaces used in the previous simulation study.

Bayesian computation. The Bayesian computation begins with the multiplication of the likelihood function by the prior distribution. Using the mentioned method in 1, the likelihood function computation is straightforward. Remaining is the selection of prior distribution and then the posterior generation using MCMC algorithm. Again let \mathbf{b} as a sample path of fBm and the parameter space is Θ_1 or Θ_2 . The suggested prior distributions for $H \in \Theta_1$ and $H \in \Theta_2$ are respectively $1/2$ Beta (a, a) and $1 - 1/2$ Beta (a, a) which both return non-informative uniform prior when $a = 1$. Therefore, the posterior distribution for the parameter space $\Theta_l, l = 1, 2$, satisfies

	\widehat{H}_c	\widehat{H}_{inc}	Bayesian	
			skewed beta prior	Uniform prior
$H=0.1$	42.458	42.983	12.882	36.810
0.3	39.595	36.941	12.512	25.951
0.45	35.937	34.629	5.589	23.242
0.55	37.790	33.498	10.423	18.313
0.7	33.129	32.943	9.435	16.225
0.9	33.876	31.025	5.634	12.320

Table 2. The MSE of ML and Bayesian estimators for $H \in (0, 1)$, multiplied by 10^4 .

$$\pi_l(H|\mathbf{b}) = C_b |\Sigma_l|^{-1/2} \exp\left(-\frac{1}{2}\mathbf{b}^T \Sigma_l^{-1} \mathbf{b}\right) (H(1 - 2H)(-1)^{l+1})^{a-1},$$

where Σ_l is the one used in (14) and $C_b > 0$ is a constant depends on \mathbf{b} . The posterior distribution has no closed form and inevitably we need to capture the posterior properties using the sample generated by the Metropolis-Hastings algorithm. We use the truncated normal distribution with small enough variance as the proposal density function. More precisely, when the algorithm is in the j th iteration, the sample H_j is generated from a normal distribution with mean H_{j-1} and variance $1/64$ which is truncated at points zero and $1/2$ when the parameter space is Θ_1 or truncated at $1/2$ and one when the true parameter belongs to Θ_2 . The variance $(1/8)^2$ is chosen to cover an interval of length $1/2$ by the 95 percent of the values that lie in a band of length $4 \times 1/8$, according to the famous thumb rule of $68 - 95 - 99.7$. Let us denote the density of this distribution by $q_l(H_j|H_{j-1})$ where the subscript $l = 1, 2$ represents the parameter space. We accept the generated H_j as a sample of the posterior distribution $\pi_l(\cdot|\mathbf{b})$ with probability $\min(1, \eta_l)$ where

$$\eta_l = \frac{\pi_l(H_j|\mathbf{b})}{\pi_l(H_{j-1}|\mathbf{b})} r_l(H_j, H_{j-1}),$$

and

$$r_1(H_j, H_{j-1}) = \frac{\Phi(4 - 8H_{j-1}) - \Phi(-8H_{j-1})}{\Phi(4 - 8H_j) - \Phi(-8H_j)},$$

$$r_2(H_j, H_{j-1}) = \frac{\Phi(8 - 8H_{j-1}) - \Phi(4 - 8H_{j-1})}{\Phi(8 - 8H_j) - \Phi(4 - 8H_j)},$$

where Φ is the cumulative distribution function of a standard normal random variable. To obtain the true random sample we eliminate the first 2500 generations and for each parameter setting we generate a sample of size 10000 from the corresponding posterior distribution. The results for $a = 1$ and 3 are gathered in the last two columns of Table 1.

Bayesian computation for the general parameter space $\Theta = (0, 1)$ is very similar. In this case, we use $\text{Beta}(a_1, a_2)$ as the prior distribution. When the observed sample path, \mathbf{b} , shows an obvious short range dependence behavior, then let $a_1 < a_2$ and for long range dependence employ $a_2 < a_1$. Using the notation of (15), the posterior is

$$\pi(H|\mathbf{b}) = C'_b (L_1 + L_2) H^{a_1-1} (1 - H)^{a_2-1},$$

where $C'_b > 0$ is the normalizing constant. The proposal density $q(\cdot|H_{j-1})$ is a normal distribution with mean H_{j-1} and variance $1/16$ which is truncated at zero and one. Thus, to obtain the j th observation from the posterior, generate an observation from density $q(\cdot|H_{j-1})$ and accept it with probability $\min(1, \eta)$ where

$$\eta = \frac{\pi(H_j|\mathbf{b})(\Phi(4(1 - H_{j-1})) - \Phi(-4H_{j-1}))}{\pi(H_{j-1}|\mathbf{b})(\Phi(4(1 - H_j)) - \Phi(-4H_j))}.$$

Random samples of size 10000 by considering the burning point at 2500 are used to estimate the Hurst index H as the posterior mean and the results are shown in Table 2. The results for $H = 0.45$ and $H = 0.55$ were obtained by use of Beta distribution with parameters $a_1 = a_2 = 3$ as the prior distribution.

Both the Tables 1 and 2 show that the Bayesian estimator with informative prior returns better results particularly for short range dependent sample paths. In Table 1 and in comparison with Table 2, the MSE of estimators are smaller and this shows the effect of more precise parameter spaces. The error of decomposition and hence the risks of \widehat{H}_c and \widehat{H}_{inc} are decreasing in H . The fourth column of Table 2 represents the importance of visual information because it returns smaller MSEs in comparison with the last column. Concerning Table 1, to make the range-preserving property for ML estimates, whenever the values of the estimators were out of the parameter spaces, the results were reset to the nearest endpoint. Surprisingly, except for the cases $H = 0.45$ and 0.55 , there was no need to reset the values. We also note that the \widehat{H}_c and \widehat{H}_{inc} have under and over estimation properties,

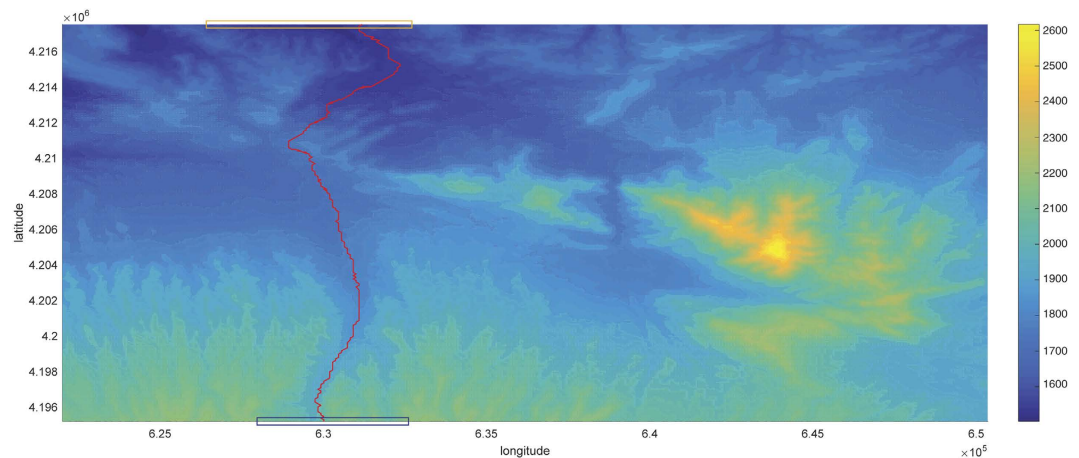


Figure 3. The image of the mountainous area from the Northwest of Iran. The smoothest path between two regions demonstrated by two rectangular in North and South of the region is shown with a red solid line. This path is determined using the Bayesian approach with beta-type prior.

respectively. For each setting, the trimmed mean base estimator¹⁵ r , is computed for further comparison. The simple comparison returns the validity of \hat{H}_c , \hat{H}_{inc} and the Bayesian estimators. Although the computation of Bayes estimators with Beta and Beta type priors are time-consuming, they might compete with other estimators when the lower MSE is the matter.

Practical aspects. According to our real data, we are confronted with the same situation in the practical problem introduced in the section entitled Practical motivation. Firstly, note that a topographic map of region S is a collection of heights $\{Z(s, t)\}$ at geographic coordinates $(s, t) \in S$. Typically, the coordinates are considered as nodes of a regular lattice. Thus, a route begins from a beginning point and is made by connecting some neighboring nodes to achieve the destination node. Precise topographic maps are often considered as fractal surfaces⁵ and estimating the Hurst index is the most interesting problem for such data. We want to construct a road to move across the region S (from the North to the South or vice versa) and the smoothness of the road is the characteristic of interest. Figure 3 shows the aerial image of S that is a mountainous area from the Northwest of Iran restricted to a rectangle with two opposite vertices located at the geographic xy coordinates $(621995.073, 4195182.082)$ and $(650345.073, 4217502.082)$. The image is gridded into a 249×316 lattice. We would like to find the smoothest North-South route between two regions indicated by the orange and blue rectangles. According to the cost of computation: (1) The image is rescaled into a 83×158 lattice; (2) we restrict our study in the new lattice to those routes made by 260 nodes at most. The χ^2 goodness of fit test of fBm is implemented on the resulting route to insure the assumption of fractional behavior. We refer to ref. 20 for a list of goodness of fit tests including the χ^2 method.

There is a narrow valley at the points with the geographic longitude equal to 6.305×10^5 and we expect to detect the smoothest path near these coordinates. Each route is a sequence in the form of $\{Z(t_i, s_i)\}_i$ such that $\{(t_i, s_i)\}_i$ is a set of connected nodes. We first sort all the possible routes with respect to their total oscillation that is simply the sum of the absolute value $|Z(t_i, s_i) - Z(t_j, s_j)|$ for any two successive nodes (t_i, s_i) and (t_j, s_j) . Then, we candidate half of the roads with smaller oscillations to find the smoothest path. If the expert idea allows us to attribute the Hurst index of these routes to Θ_1 or Θ_2 then we can use the beta-type prior to the estimate H ; or else, the simple uniform prior on $(0, 1)$ will be employed. In this case, it is suggested to employ a beta-type prior on Θ_1 . A decomposition of $N = 100$ is used to approximate the fBm and hence the likelihood function of each route. The smoothest route is the path with the larger estimated H . The smoothest detected path is shown with a solid red line in Fig. 3. This route is constructed by 136 nodes and the line transect of the route is provided in Fig. 4. It is worth mentioning that the estimated H parameter of this route is 0.42065 based on the Bayesian computation. The estimators based on the discrete variation, sample quantiles and trimmed mean respectively return the values 0.68479, 0.47103 and 0.42557 for this route.

Model diagnosis. The main characteristic of the fBm is the subdiffusion that is

$$\langle X^2(t) \rangle \approx C_H |t|^H, \quad (16)$$

where $0 < H < 1$, is the Hurst index or subdiffusive exponent and $C_H > 0$ is diffusion constant. The class of subdiffusion processes contains three well-known stochastic processes; the fBm, continuous-time random walk (CTRW)²¹; and diffusion on the fractal lattice (DFL)^{22,23}. Although these processes belong to a certain class, their mechanism are rather different²⁴. Applying the presented methods to any other process than the fBm will give some H which is misleading in the sense that the underlying process is not at all an fBm. In this section we want to test whether the underlying process is an fBm or not. To solve this, we first examine some general properties of the fBm and after obtaining positive results, we use the p -variation method²⁵ to discriminate the fBm from the CTRW and DFL. The outline of the algorithm is as follows²⁵:

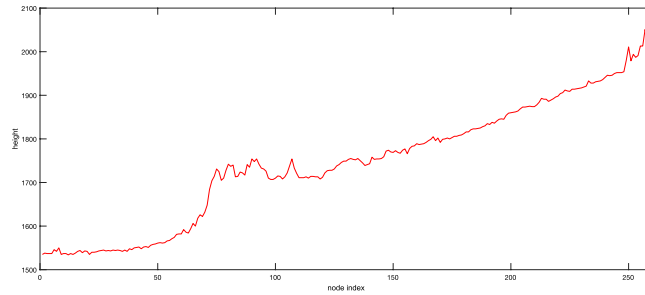


Figure 4. The line transect of the smoothest route detected by the Bayesian approach.

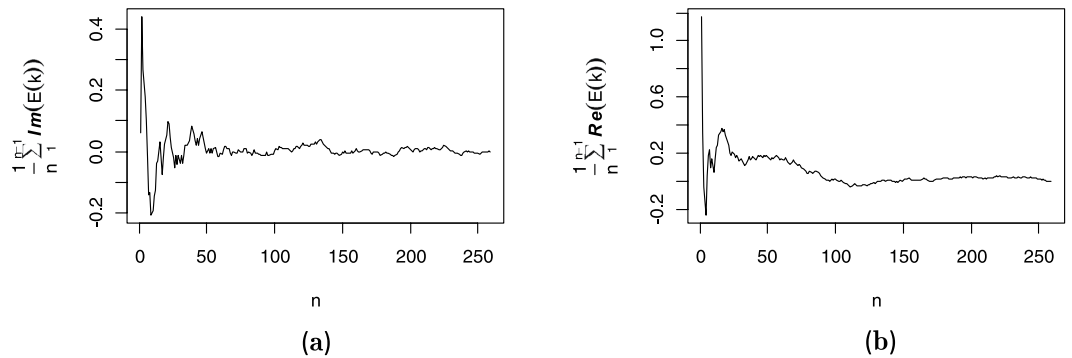


Figure 5. Panel (a) shows the imaginary part of the cumulative mean of $E(k)$ (17) versus n while the right panel (b) demonstrates the real part. In comparison with the simulated processes analyzed by³¹, the range of both the real and imaginary parts are closer to the zero and figures verifies the ergodicity of the underlying process.

- Verifying the stationarity of increments.
- Verifying the normality of the observed path.
- Testing the ergodicity of the increments.
- Testing the p -variation properties for some specific p 's.
- Testing the filling ratio properties.

We have to apply this procedure to all the possible routes, but for the sake of convenience, we do this on some selected routes, and for the sake of conciseness, we demonstrate the results only for the smoothest detected route.

Let's start with the first step of the algorithm. The height of each pixel strongly depends on the height of its neighbors. This induces the unit root effect that causes the non-stationarity. A very familiar testing procedure to test the stationarity versus the unit root is KPSS²⁶. Using the R package `tseries`, the KPSS statistic for the increments of the smoothest route is 0.33641 that produces the p -value equal to 0.1 and thus the incremental stationarity assumption is verified. Testing the stationarity is not restricted to this method and we refer to ref. 25 for another method that requires more sample paths; whereas we only have one realization and the method can not be implemented here.

Note that there are some reasonable effects that can deflect the trajectory of a subdiffusive process from the Gaussian assumption²⁷. That is why we are going to test the normality assumption in the second step. There is plenty of normality testing methods including, Kolmogorov-Smirnov, Anderson-Darling, Shapiro-Wilk, Cramer-von Mises, Shapiro-Francia and Jarque-Bera. There is no certain order for these tests in terms of power, but some of them require the strong assumption of independence of the sample path; which is disaffirmed here under the subdiffusion model. Fortunately, the method of Jarque-Bera is not sensitive to the independence assumption and we employ it to test the normality of increments. We use the R package `fBasics` and we find out that the χ^2 -statistic of Jarque-Bera test is equal to 3.5049 and the corresponding p -value is 0.1733 which confirms the normality of the increments of the smoothest path.

Ergodicity states that the long time averages of physical characteristics provide the same information as the corresponding ensemble average. The ergodicity assumption can be violated for sub²⁸ and super^{29,30} diffusive trajectories even under the power law. It is one of the most important parts of studies, particularly those on single particle tracking. Using the dynamical functional

$$D(n) = \langle \exp\{i(\Delta X(n) - \Delta X(t_0))\} \rangle,$$

where $i = \sqrt{-1}$, an ergodicity test³¹ is employed here. Set $E(n) = D(n) - |\langle \exp(i\Delta X(t_0)) \rangle|^2$. For mixing increments we have³² $\lim_{n \rightarrow \infty} E(n) = 0$ while the increment process is ergodic if and only if the sample mean of $\{E(k)\}_{k=1}^n$ tends to zero, i.e.,

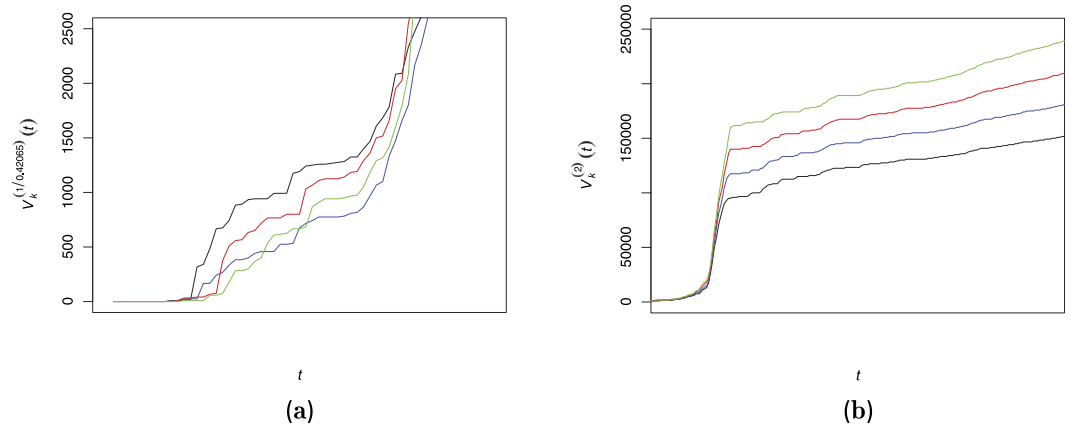


Figure 6. The left panel (a) shows the computed $V_k^{(1/0.42065)}$ of the smoothest path for $k=7$ (black line); $k=8$ (blue line); $k=9$ (red line) and $k=10$ (green line). This panel does not demonstrate any trend between k and $V_k^{(1/0.42065)}$. This is the same behavior of any simulated sample path from an fBm of the same index and size. The panel (b) shows $V_k^{(2)}$ with the same color setting of k . It is clear that the 2-variation is increasing in k and hence the CTRW model is rejected in favor of fBm.

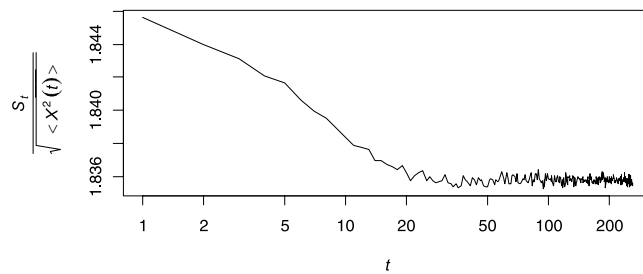


Figure 7. The number of distinct visited sites divided by $\sqrt{\langle X^2(t) \rangle}$ where both the x and y axes are in logarithmic scale. According to the results, $\log(S(t)/\sqrt{\langle X^2(t) \rangle})$ oscillates smoothly a round a constant value by changing $\log t$.

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^{n-1} E(k) = 0. \tag{17}$$

Figure 5 shows the rate of convergence of the cumulative mean to zero. The range of oscillation around zero is very small for the large enough values of n that verifies the ergodicity of the smoothest path.

At this point, the smoothest route passed three steps of the algorithm and in the current step we need to attribute one of the fBm, CTRW or DFL to our data. To this end we use the p -variation method³³ to study the variational behavior of the process. Let $V_k^{(p)}(t)$ denotes the $1/p$ th root of the incremental p -norm, i.e.,

$$V_k^{(p)}(t) = \sum_{j=0}^{n-1} |X(t_{j+1} \wedge t) - X(t_j \wedge t)|^p,$$

where $a \wedge b = \min \{a, b\}$. If $X(t)$ is an fBm then $V_k^{(2)}$ is increasing in k and if $X(t)$ is a CTRW then $\lim_{k \rightarrow \infty} V_k^{(1/\tilde{H})}(t) = 0$ where \tilde{H} is the estimated Hurst index produced by any proper method. Figure 6 exhibits the resulting values of $V_k^{(p)}$ for $p=2$ and $p=1/0.42065$. The results verify that the smoothest route does not follow the CTRW model and thus we remove the CTRW from our list.

We have to make a decision between the fBm model and the DFL in the final step. There is a testing procedure based on the mean maximal excursion (MME) method³⁴ which needs repeated trajectories from a single process. In some practical problems, the mechanism of data allows gathering independent trajectories of a subdiffusive model with a constant Hurst index²⁴. However, in our study, each route is considered as a realization of a subdiffusive process with a corresponding Hurst index. Thus, restricting it to the smoothest route, we only have one trajectory that could be whether an fBm or a DFL. Fortunately, the MME method was modified to be implemented on a single trajectory to decide between the fBm and DFL³⁵. Let S_t denotes the total number of distinct visited sites of the process $X(t)$ up to time t . The ratio $S(t)/\sqrt{\langle X^2(t) \rangle}$ is known as the filling ratio and we have to

compute its slope in a log – log scale. The hypothesis of fBm is rejected versus DFL if the slope is far from zero. Figure 7 shows the values of the filling ratios in a double logarithmic scale and the estimated slope is -4.79×10^{-4} which implies that the logarithm of the filling ratio and time index are uncorrelated. This means that the dimension of the realization was constructed only by a fBm model and cannot be decomposed into an fBm and a fractal noise. So the DFL model is not verified and the final step in the fBm trajectory diagnosis algorithm is passed.

References

- Mandelbrot, B. B. & Van Ness, J. Fractional Brownian motion, fractional noises and applications. *SIAM Rev.* **10**, 422–437 (1968).
- Kolmogorov, A. N. The Wiener spiral and some other interesting curves in a Hilbert space (in German). *C. R. (Doklady) Acad. Sci. URSS* **83**, 115–118 (1940).
- Adler, R. J. *The geometry of random fields* (New York: Wiley, 1981).
- Baxevani, A. & Podgórsky, K. Series decomposition of fractional Brownian motion and its Lamperti transform. *Acta Phys. Polon. B* **40**, 1395–1435 (2009).
- Davies, S. & Hall, P. Fractal analysis of surface roughness by using spatial data. *J Royal Stat Soc Ser B Stat Methodology* **61**, 3–37 (1999).
- Hurst, H. E. Long-term storage capacity of reservoirs. *Trans. Am. Soc. Civ. Eng.* **116**, 770–779 (1951).
- Mandelbrot, B. B. & Wallis, J. R. Computer experiments with fractional Gaussian noises. *Water Resour. Res.* **5**, 228–267 (1969).
- Mandelbrot, B. B. & Taqqu, M. S. Robust R/S analysis of long run serial correlation. *Bull Int Stat Inst* **48**, 59–104 (1979).
- Lundahl, T., Ohley, W. J., Kay, S. M. & Siffert, R. Fractional Brownian motion: A maximum likelihood estimator and its application to image texture. *IEEE Trans. Med. Imaging MI-5*, 152–161 (1986).
- Carlin, J. B. & Dempster, P. Sensitivity analysis of seasonal adjustments: Empirical case studies (with discussion). *J Amer Stat Assoc* **84**, 6–32 (1989).
- Coeurjolly, J.-F. Estimating the parameters of a fractional Brownian motion by discrete variations of its sample paths. *Stat. Inference Stoch. Process.* **4**, 199–227 (2001).
- Coeurjolly, J.-F. Hurst exponent estimation of locally self-similar Gaussian processes using sample quantiles. *Ann. Statist.* **36**, 1404–1434 (2008).
- Kent, J. T. & Wood, A. T. A. Estimating the fractal dimension of a locally self-similar Gaussian process by using increments. *J. Roy. Statist. Soc. Ser. B* **59**, 679–699 (1997).
- Istas, J. & Lang, G. Quadratic variations and estimation of the local Hölder index of a Gaussian process. *Ann. Inst. H. Poincaré Probab. Statist.* **33**, 407–436 (1997).
- Achard, S. & Coeurjolly, J.-F. Discrete variations of the fractional Brownian motion in the presence of outliers and an additive noise. *Stat. Surv.* **4**, 117–147 (2010).
- Salomón, L. A. & Fort, J.-C. Estimation of the Hurst parameter in some fractional processes. *J Stat Comput Simulation* **83**, 542–554 (2013).
- Norros, I., Valkeila, E. & Virtamo, J. An elementary approach to a girsanov formula and other analytical results on fractional brownian motions. *Bernoulli* **5**, 571–587 (1999).
- Dung, N. T. & Thao, T. H. An approximate approach to fractional stochastic integration and its applications. *Braz. J. Probab. Stat.* **24**, 57–67 (2010).
- Wood, A. T. A. & Chan, G. Simulation of stationary Gaussian processes in $[0, 1]^d$. *J. Comput. Graph. Stat.* **3**, 409–432 (1994).
- Beran, J. *Statistics for long-memory processes, vol. 61 of Monographs on Statistics and Applied Probability* (Chapman and Hall, New York, 1994).
- Jeon, J.-H., Barkai, E. & Metzler, R. Noisy continuous time random walks. *J. Chem. Phys.* **139**, 121916 (2013).
- ben Avraham, D. & Havlin, S. *Diffusion and reactions in fractals and disordered systems* (Cambridge University Press, Cambridge, 2000).
- Mardoukhi, Y., Jeon, J.-H. & Metzler, R. Geometry controlled anomalous diffusion in random fractal geometries: looking beyond the infinite cluster. *Phys. Chem. Chem. Phys.* **17**, 30134–30147 (2015).
- Yamamoto, E., Akimoto, T., Yasui, M. & Yasuoka, K. Origin of subdiffusion of water molecules on cell membrane surfaces. *Sci. Rep.* **4**, 4720 (2014).
- Burnecki, K. *et al.* Universal algorithm for identification of fractional Brownian motion. a case of telomere subdiffusion. *Biophys. J.* **103**, 1839–1847 (2012).
- Kwiatkowski, D., Phillips, P. C., Schmidt, P. & Shin, Y. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *J. Econometrics* **54**, 159–178 (1992).
- Wang, B., Kuo, J., Bae, S. C. & Granick, S. When Brownian diffusion is not Gaussian. *Nat. Materials* **11**, 481–485 (2012).
- Bodrova, A., Chechkin, A. V., Cherstvy, A. G. & Metzler, R. Quantifying non-ergodic dynamics of force-free granular gases. *Phys. Chem. Chem. Phys.* **17**, 21791–21798 (2015).
- Cherstvy, A. G. & Metzler, R. Nonergodicity, fluctuations, and criticality in heterogeneous diffusion processes. *Phys. Rev. E* **90**, 012134 (2014).
- Reverey, J. F. *et al.* Superdiffusion dominates intracellular particle motion in the supercrowded cytoplasm of pathogenic *acanthamoeba castellanii*. *Sci. Rep.* **5**, 11690 (2015).
- Magdziarz, M. & Weron, A. Anomalous diffusion: testing ergodicity breaking in experimental data. *Phys. Rev. E* **84**, 051138 (2011).
- Magdziarz, M. & Weron, A. Ergodic properties of anomalous diffusion processes. *Ann. Phys.* **326**, 2431–2443 (2011).
- Magdziarz, M., Weron, A., Burnecki, K. & Klafter, J. Fractional Brownian motion versus the continuous-time random walk: A simple test for subdiffusive dynamics. *Phys. Rev. Lett.* **103**, 180602 (2009).
- Tejedor, V. *et al.* Quantitative analysis of single particle trajectories: mean maximal excursion method. *Biophys. J.* **98**, 1364–1372 (2010).
- Meroz, Y., Sokolov, I. M. & Klafter, J. Test for determining a subdiffusive model in ergodic systems from single trajectories. *Phys. Rev. Lett.* **110**, 090601 (2013).

Acknowledgements

The authors gratefully thank the anonymous reviewer for her/his careful reading and constructive criticism. The comments and suggestions have helped us make revolutionary improvements to the previous version of the paper.

Author Contributions

A.R. Taheriyoun wrote the manuscript. The idea of \widehat{H}_c was introduced by M. Moghimbeygi who also carried out its computation.

Additional Information

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Taheriyoun, A. R. and Moghimbeygi, M. Visual information and expert's idea in Hurst index estimation of the fractional Brownian motion using a diffusion type approximation. *Sci. Rep.* 7, 42482; doi: 10.1038/srep42482 (2017).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017