Physica A 444 (2016) 43–48

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

# Hurst exponent estimation of self-affine time series using quantile graphs

Andriana S.L.O. Campanharo [a], Fernando M. Ramos [b,*]

[a] Departamento de Bioestatística, Instituto de Biociências, Universidade Estadual Paulista, Botucatu, São Paulo, Brazil

[b] Laboratório de Computação e Matemática Aplicada, Instituto Nacional de Pesquisas Espaciais, São José dos Campos, São Paulo, Brazil

## HIGHLIGHTS

- The quantile graph (QG) method for the estimation of the Hurst exponent of self-affine time series is presented.
- The QG method is applied to the characterization of different fractional Brownian motions.
- Comparison between H estimates using the QG method and the exact values used to generate the motions shows an excellent agreement.
- For a given time series length, H estimation error depends basically on the statistical framework used for determining the exponent of a power-law model.
- The QG method is numerically simple and has only one free parameter, the number of quantile/nodes; with a simple modification, it can be extended to the analysis of fractional Gaussian noises.

## ARTICLE INFO

*Article history:*
Received 26 May 2015
Received in revised form 23 September 2015

*Keywords:*
Self-affine time series
Hurst exponent
Complex networks
Quantile graphs

## ABSTRACT

In the context of dynamical systems, time series analysis is frequently used to identify the underlying nature of a phenomenon of interest from a sequence of observations. For signals with a self-affine structure, like fractional Brownian motions (fBm), the Hurst exponent $H$ is one of the key parameters. Here, the use of quantile graphs (QGs) for the estimation of $H$ is proposed. A QG is generated by mapping the quantiles of a time series into nodes of a graph. $H$ is then computed directly as the power-law scaling exponent of the mean jump length performed by a random walker on the QG, for different time differences between the time series data points. The QG method for estimating the Hurst exponent was applied to fBm with different $H$ values. Comparison with the exact $H$ values used to generate the motions showed an excellent agreement. For a given time series length, estimation error depends basically on the statistical framework used for determining the exponent of the power-law model. The QG method is numerically simple and has only one free parameter, $Q$, the number of quantiles/nodes. With a simple modification, it can be extended to the analysis of fractional Gaussian noises.



## 1. Introduction

In the last two decades, research on complex networks became the focus of widespread attention, with developments and applications spanning different scientific areas, from sociology and biology to physics. One of the reasons behind the growing

* Corresponding author.
*E-mail addresses:* andriana@ibb.unesp.br (A.S.L.O. Campanharo), fernando.ramos@inpe.br (F.M. Ramos).

http://dx.doi.org/10.1016/j.physa.2015.09.094

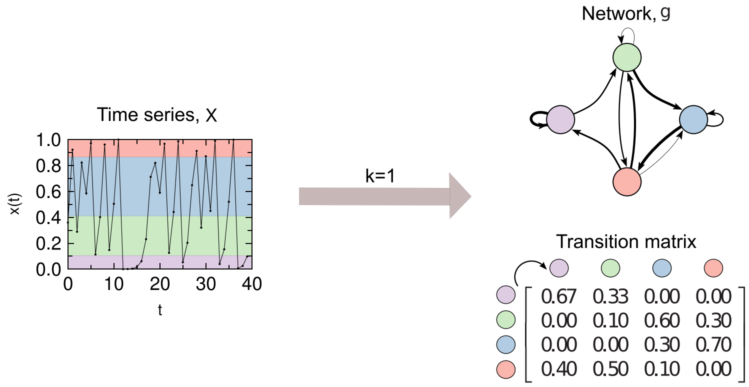


**Fig. 1.** Illustration of the QG method for $k = 1$. A time series $X$ is split into $Q = 4$ quantiles (colored shading) and each quantile $q_i$ is assigned to a node $n_i \in \mathcal{N}$ in the corresponding network $g$. Two nodes $n_i$ and $n_j$ are then connected in the network with a weighted arc $(n_i, n_j, w_{ij}^k) \in \mathcal{A}$ where the weight $w_{ij}$ of the arc is given by the probability that a point in quantile $q_i$ is followed by a point in quantile $q_j$. Repeated transitions between quantiles results in arcs in the network with larger weights (represented by thicker lines) and therefore higher values in the corresponding transition matrix. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

popularity of complex networks is that almost any discrete structure can be suitably represented as a graph, whose features may be then characterized, analyzed and, eventually, related to its respective dynamics [1]. Recently several approaches have been proposed for mapping a time series into a complex network representation, based on concepts such as correlations [2,3], visibility [4,5], recurrence analysis [6], transition probabilities [7–10] and phase-space reconstructions [11–13]. These studies have shown that distinct features of a time series can be mapped onto networks with distinct topological properties, opening the door to the analysis of discrete, time-ordered data sets with mathematical tools usually used in the study of geometric shapes and topological spaces.

Fractional Gaussian noises (fGn) and fractional Brownian motions (fBm) have been used as a theoretical framework to investigate time series emerging in different scientific areas. A fBm is a non-stationary self-affine process with stationary increments given by a fGn [14]. Self-affine random processes are characterized by a power-spectrum with a power-law dependency on frequency as $P(\omega) \sim \omega^{-\beta}$, with $-1 \leq \beta \leq 1$ for noises and $1 < \beta \leq 3$ for motions. For $\beta = 0$, we have a white uncorrelated noise [15]. Noises and motions are also characterized by their Hurst exponents ($H$) [16]. The classical Brownian motion is a fBm with $H = 1/2$. Correlation between fBm's increments is negative for $0 < H < 1/2$ and positive for $1/2 < H < 1$ [17]. Because of the intrinsic non-stationarity and long range dependence of fBm, the estimation of $H$ often requires more robust methods than those provided by standard Fourier analysis [18].

In spite of the large number of applications of complex networks methods in the study time series, usually involving the classification of dynamical systems or the identification of dynamical transitions (see Ref. [13] and references therein), establishing a link between a network measure and $H$ remains an open question. Recently, a linear relationship between the exponent of the power law degree distribution of visibility graphs and $H$ has been established for noises and motions [19, 20]. Here, we propose an alternative approach for the computation of the Hurst exponent. This new approach is based on a generalization of the method introduced in Ref. [10], in which time series quantiles are mapped into nodes of a graph (here called quantile graph or QG) and vice-versa. To this end $H$ is computed directly as power-law scaling exponent of the mean jump length performed by a random walker on a QG, for different time differences between the time series data points. $H$ estimates computed with this method are robust, with a standard error that depends basically on the statistical framework used for fitting a power-law model to the random walk data (for a maximum likelihood estimator, see, for example, Ref. [21]).

This paper is organized as follows. After this Introduction, we describe in Section 2 the QG method for computing $H$. Results are presented and discussed in Section 3 while an overall conclusion is given in Section 4.

## 2. Methods

Let the range of values in a time series be coarse-grained into $Q$ quantiles $q_1, \ldots, q_Q$, and let $\mathcal{M}$ be a map from a time series $X \in \mathcal{T}$ to a network $g \in \mathcal{G}$, with $X = \{x(t)|t \in \mathbb{N}, x(t) \in \mathbb{R}\}$ and $g = \{\mathcal{N}, \mathcal{A}\}$ being a set of nodes $\mathcal{N}$ and arcs $\mathcal{A}$. Specifically, $\mathcal{M}$ assigns each quantile $q_i$ to a node $n_i \in \mathcal{N}$ in the corresponding network. Two nodes $n_i$ and $n_j$ are connected with a weighted arc $n_i, n_j, w_{ij}^k \in \mathcal{A}$ whenever two values $x(t)$ and $x(t+k)$ belong respectively to quantiles $q_i$ and $q_j$, with $t = 1, 2, \ldots, T$ and the time differences $k = 1, \ldots, k_{\max} < T$. For an illustration of the QG method for $k = 1$, see Fig. 1.

Weights $w_{ij}^k$ are simply given by the number of times a value in quantile $q_i$ at time $t$ is followed by a point in quantile $q_j$ at time $t+k$, normalized by the total number of transitions. Repeated transitions through the same arc increase the value of the corresponding weight. With proper normalization, the weighted adjacency matrix becomes a Markov transition matrix $\mathbf{W_k}$, with $\sum_j^Q w_{ij}^k = 1$. The resulting network is weighted, directed and connected, with $Q$ being typically much smaller than $T$. Note that by randomly moving from one node to the other with probability given by $\mathbf{W_k}$, and by assigning the corresponding quantile values to $x(t)$, it is possible to reconstruct $X$ from $g$.