Commun Nonlinear Sci Numer Simulat 40 (2016) 163–172

Contents lists available at ScienceDirect

Commun Nonlinear Sci Numer Simulat

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

Research paper

Power-law cross-correlations estimation under heavy tails

Ladislav Kristoufek [a,b,*]

[a] Institute of Information Theory and Automation, The Czech Academy of Sciences, Pod Vodarenskou Vezi 4, 182 08 Prague 8, Czech Republic
[b] Institute of Economic Studies, Faculty of Social Sciences, Charles University, Opletalova 26, 110 00 Prague 1, Czech Republic

A R T I C L E   I N F O

Article history:
Received 22 June 2015
Revised 6 February 2016
Accepted 8 April 2016
Available online 24 April 2016

Keywords:
Power-law cross-correlations
Heavy tails
Monte Carlo study

A B S T R A C T

We examine the performance of six estimators of the power-law cross-correlations—the detrended cross-correlation analysis, the detrending moving-average cross-correlation analysis, the height cross-correlation analysis, the averaged periodogram estimator, the cross-periodogram estimator and the local cross-Whittle estimator—under heavy-tailed distributions. The selection of estimators allows to separate these into the time and frequency domain estimators. By varying the characteristic exponent of the α-stable distributions which controls the tails behavior, we report several interesting findings. First, the frequency domain estimators are practically unaffected by heavy tails bias-wise. Second, the time domain estimators are upward biased for heavy tails but they have lower estimator variance than the other group for short series. Third, specific estimators are more appropriate depending on distributional properties and length of the analyzed series. In addition, we provide a discussion of implications of these results for empirical applications as well as theoretical explanations.

1. Introduction

Power-law cross-correlations have brought a new perspective into analysis of bivariate time series with applications across a wide range of disciplines—hydrology [1], (hydro)meteorology [2,3], seismology and geophysics [4,5], economics and finance [6–12], biometrics [13], biology [14], DNA sequences [15], neuroscience [16], electricity [17], traffic [18–21], and others. Analyzing cross-correlations between a pair of time series brings new insights into their dynamics and specifically, the power-law behavior of these may suggest very special features compared to exponential (vanishing) cross-correlations [22–27]. Presence of power-law cross-correlations is a very current topic in various scientific fields but, unfortunately, empirical papers strongly outnumber theoretical ones discussing their emergence and origin. Podobnik et al. [28] suggest that power-law cross-correlated processes can occur as a mixture of correlated long-range correlated processes. Sela & Hurvich [25] and Kristoufek [26,27] add a possibility of long-range correlated processes with cross-correlated error terms as sources of long-range cross-correlations together with a discussion of various properties of such processes. Investigation of power-law cross-correlated processes has also led to an important novelty in analyzing cross-correlations at specific scales [29–32] as well as a new regression framework for specific scales [33].

* Correspondence to: Institute of Information Theory and Automation, The Czech Academy of Sciences, Pod Vodarenskou Vezi 4, 182 08 Prague 8, Czech Republic. Tel.: +420 607684417.
E-mail address: kristoufek@ies-prague.org, kristouf@utia.cas.cz

http://dx.doi.org/10.1016/j.cnsns.2016.04.010
1007-5704/

Formally, processes labeled as power-law (long-range, long-term) cross-correlated can be defined in both time and frequency domain. In the time domain, the long-range cross-correlated processes $\{x_t\}$ and $\{y_t\}$ are characterized by a power-law (hyperbolically) decaying cross-correlation function $\rho_{xy}(k)$ with a time lag k so that $\rho_{xy}(k) \propto k^{2H_{xy}-2}$ for $k \to +\infty$ [34]. In the frequency domain, the processes have a divergent at origin cross-power spectrum. Specifically, the cross-power spectrum $f_{xy}(\lambda)$ with frequency λ scales as $|f_{xy}(\lambda)| \propto \lambda^{1-2H_{xy}}$ for $\lambda \to 0+$ [25]. The bivariate Hurst exponent H_{xy} is a measure of long-range cross-correlations, specifically their decay. For $H_{xy} = 0.5$, the processes are not considered long-range cross-correlated whereas for $H_{xy} > 0.5$, these are referred to as the cross-persistent processes which tend to move together. Time series with $H_{xy} < 0.5$ form a very specific type of processes which have been theoretically only sparsely examined.

In the same way as for the univariate series, the correlations scaling and the underlying distribution tail behavior are tightly interconnected [35,36]. Investigation of the relationship between the two goes back to Mandelbrot & Wallis [37]—and thus hydrology—when examining the emerge of the so-called Hurst effect [38]. Even though the effect emerges both for long-range dependent processes and for processes with heavy tails, it is the correlation function, and specifically its shape, which is crucial. The confusion about the real and the spurious sources of the Hurst effect spreads also into self-similarity and phase transitions [39]. Such spurious effects can easily transfer into bivariate, or in general multivariate, setting which forms a motivational cornerstone of this text.

Presence of heavy tails in distributions, or in other words high occurrence of extreme events, is well documented across disciplines [40–50]. However, the current stream of literature does not take a possible effect on estimators of the bivariate Hurst exponent implied by such heavy-tailed (fat-tailed) distributions into consideration. Here we fill this gap by focusing on three frequency domain estimators—the average periodogram estimator [25], the cross-periodogram estimator and the local cross-Whittle estimator [51]—and three time domain estimators—the detrended cross-correlation analysis [52], the detrending moving-average cross-correlation analysis [7,53] and the height cross-correlation analysis [54]—performance of which we examine under such distributions. Specifically, we study the effect of varying characteristic exponent of the α-stable distributions on bias, variance and mean squared error of the estimators for various time series lengths.

2. Methods

In this section, we shortly describe six estimators of the bivariate Hurst exponent. These are separated into two groups based on their domain of operation—time and frequency. The simulation setting is then described in detail.

2.1. Time domain estimators

The detrended cross-correlation analysis (DCCA, or DXA) [52] is the most popular time domain method of the bivariate Hurst exponent estimation, constructed as a bivariate generalization of the detrended fluctuation analysis (DFA) [55,56]. The method has led to various generalizations and expansions [57–59]. Considering two long-range cross-correlated series $\{x_t\}$ and $\{y_t\}$ with $t = 1, \ldots, T$ and their respective profiles $\{X_t\}$ and $\{Y_t\}$, the DCCA procedure is based on an examination of the detrending covariances scaling with respect to scale s. Specifically, the time series are divided into overlapping boxes of length s and the linear time trend is estimated in each box yielding $\widehat{X_{k,j}}$ and $\widehat{Y_{k,j}}$ for boxes $j \le k \le j+s-1$. The detrended covariance $f^2_{DCCA}(s,j)$ is obtained for each box j of length s and it is further averaged over all boxes of length s to get $F^2_{DCCA}(s)$ as an estimated covariance for scale s. For the power-law cross-correlated processes, we have $F^2_{DCCA}(s) \propto s^{2H_{xy}}$. There are various ways of treating overlapping and non-overlapping boxes for scales s as the method can become computationally highly demanding [35,36,60–62]. Due to this fact, we use non-overlapping boxes with a minimum scale of 10, a maximum scale of $T/5$ and a step between s equal to 10 in the simulations.

The height cross-correlation analysis (HXA) [54] is a bivariate generalization of the height-height correlation analysis (HHCA) [63,64] and the generalized Hurst exponent approach (GHE) [65–67]. The method is based on scaling of the height-height covariance function $K_{xy,2}(\tau) = \frac{\nu}{T^*}\sum_{t=1}^{T^*/\nu} |\Delta_\tau X_t Y_t| \equiv \langle |\Delta_\tau X_t Y_t| \rangle$ of detrended profiles $\{X_t\}$ and $\{Y_t\}$ with time resolution ν and $t = \nu, 2\nu, ..., \nu\lfloor \frac{T}{\nu} \rfloor$, where $\lfloor\rfloor$ is a lower integer sign. We denote $T^* = \nu\lfloor \frac{T}{\nu} \rfloor$, which varies with ν, and we write the τ-lag difference as $\Delta_\tau X_t \equiv X_{t+\tau} - X_t$ and $\Delta_\tau X_t Y_t \equiv \Delta_\tau X_t \Delta_\tau Y_t$ for better legibility. The covariance function scales as $K_{xy,2}(\tau) \propto \tau^{H_{xy}}$ for the power-law cross-correlated processes. Di Matteo et al. [65–67] suggest using the jackknife method to obtain more precise estimates of the Hurst exponent. In our simulations, we set $\tau_{min} = 1$ and vary $\tau_{max} = 5, \ldots, 20$ to get the estimated H_{xy} as an average of these.

The detrended moving-average cross-correlation analysis (DMCA) [7,53] is a generalization of the detrending moving average (DMA) [68,69]. The method is in a way similar to DCCA as it assumes the power-law scaling of detrended covariances. However, there is no box-splitting involved in the procedure, which makes DMCA much less computationally demanding. Specifically, the profiles $\{X_t\}$ and $\{Y_t\}$ are detrended using the moving average of length κ and the covariances of residuals $F^2_{DMCA}(\kappa)$ scale as $F^2_{DMCA}(\kappa) \propto \kappa^{2H_{xy}}$ for the power-law cross-correlated processes. In our setting, we use the non-weighted centered moving averages with $\kappa_{min} = 11$ and $\kappa_{max} = 1 + T/5$ with a step of 2, i.e. parallel to the DCCA setting.

2.2. Frequency domain estimators

Construction of the frequency domain estimators is based on the divergence at origin of the cross power-spectrum of the power-law cross-correlated processes. As the cross-power spectrum is unobservable, its estimation becomes a crucial part