Physica A 419 (2015) 513–525

Contents lists available at ScienceDirect

Physica A

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

Finite sample properties of power-law cross-correlations estimators

Ladislav Kristoufek *

Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Pod Vodarenskou Vezi 4, 182 08, Prague 8, Czech Republic

Institute of Economic Studies, Faculty of Social Sciences, Charles University, Opletalova 26, 110 00, Prague 1, Czech Republic

HIGHLIGHTS

- Finite sample properties of power-law cross-correlations estimators are studied.
- DCCA, DMCA and HXA methods are compared.
- Each of the methods is better suited for specific characteristics.
- There is no clear winner.

ARTICLE INFO

Article history:
Received 20 June 2014
Received in revised form 24 September 2014
Available online 22 October 2014

Keywords:
Power-law cross-correlations
Long-term memory
Econophysics

ABSTRACT

We study finite sample properties of estimators of power-law cross-correlations – detrended cross-correlation analysis (DCCA), height cross-correlation analysis (HXA) and detrending moving-average cross-correlation analysis (DMCA) – with a special focus on short-term memory bias as well as power-law coherency. We present a broad Monte Carlo simulation study that focuses on different time series lengths, specific methods' parameter setting, and memory strength. We find that each method is best suited for different time series dynamics so that there is no clear winner between the three. The method selection should be then made based on observed dynamic properties of the analyzed series.

1. Introduction

Power-law cross-correlations have become a popular and frequently analyzed topic in various disciplines covering seismology [1], hydrology [2], (hydro)meteorology [3,4], biology [5], biometrics [6], DNA sequences [7], neuroscience [8], electricity [9], finance [10–12], commodities [13,14], traffic [15–17], geophysics [18] and others. The analysis is standardly based on an estimation of the bivariate Hurst exponent H_{xy} which is connected to an asymptotic power-law decay of the cross-correlation function or a divergent (again following a power-law) at origin cross-power spectrum. Specifically, a power-law cross-correlated process has the cross-correlation function of a form $\rho_{xy}(k) \propto k^{2H_{xy}-2}$ for lag $k \to +\infty$ and the cross-power spectrum of a form $|f_{xy}(\lambda)| \propto \lambda^{1-2H_{xy}}$ for frequency $\lambda \to 0+$. In a similar way as for the univariate case, the bivariate Hurst exponent of 0.5 is characteristic for no power-law cross-correlations. Processes with $H_{xy} > 0.5$ are then cross-persistent and they tend to move together whereas for $H_{xy} < 0.5$ they are more likely to move in opposite directions.

Most of the literature focusing on power-law cross-correlations is empirical and there are no studies of statistical properties of the utilized estimators. Here, we try to fill this gap and we present a broad Monte Carlo simulation study

* Correspondence to: Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Pod Vodarenskou vezi 4, 182 08, Prague 8, Czech Republic.

E-mail address: kristouf@utia.cas.cz.

http://dx.doi.org/10.1016/j.physa.2014.10.068
0378-4371/

of performance of three popular bivariate Hurst exponent estimators—detrended cross-correlation analysis [19–21], height cross-correlation analysis [22] and detrending moving-average cross-correlation analysis [23,14]. Specifically, we focus on an ability of the estimators to precisely estimate the bivariate Hurst exponent not only under a simple setting of standard power-law cross-correlations when the bivariate Hurst exponent equals to an average of the Hurst exponent of the separate processes but also under potential short-term memory bias and under power-law coherency. The paper is organized as follows. In Section 2, we introduce all three analyzed estimators. In Section 3, the Monte Carlo simulation setting is described. In Section 4, the results are presented in detail. Section 5 concludes.

2. Methodology

2.1. Detrended cross-correlation analysis

Detrended cross-correlation analysis (DCCA, or DXA) is the most frequently used method for the estimation of the bivariate Hurst exponent in the time domain. Podobnik and Stanley [19] construct the method as a bivariate generalization of the detrended fluctuation analysis (DFA), which is again probably the most popular heuristic method of estimating the (generalized) Hurst exponent [24–26]. DCCA was further generalized for the multifractal analysis by Zhou [20] and the multifractal detrended cross-correlation analysis (MF-DXA) was developed. Jiang and Zhou [21] altered the filtering procedure in MF-DXA by using the moving averages to create the multifractal detrending moving average cross-correlation analysis (MF-X-DMA). DCCA was also used to construct statistical tests for the presence of long-range cross-correlations between two series [27–32].

In the DCCA procedure, we consider two long-range cross-correlated series $\{x_t\}$ and $\{y_t\}$ with $t = 1, \ldots, T$. Their respective profiles $\{X_t\}$ and $\{Y_t\}$, defined as $X_t = \sum_{i=1}^{t} x_i - \bar{x}$ and $Y_t = \sum_{i=1}^{t} y_i - \bar{y}$, for $t = 1, \ldots, T$, are divided into overlapping boxes of length s so that $T-s+1$ boxes are constructed. In each box between j and $j+s-1$, the linear fit of a time trend is constructed so that we get $\widehat{X_{k,j}}$ and $\widehat{Y_{k,j}}$ for $j \leq k \leq j+s-1$. The covariance between residuals in each box is defined as

$$f^2_{\mathrm{DCCA}}(s,j) = \frac{\sum_{k=j}^{j+s-1} (X_k - \widehat{X_{k,j}})(Y_k - \widehat{Y_{k,j}})}{s-1}. \tag{1}$$

The covariances are finally averaged over the blocks of the same lengths s and the detrended covariance is obtained as

$$F^2_{\mathrm{DCCA}}(s) = \frac{\sum_{j=1}^{T-s+1} f^2_{\mathrm{DCCA}}(s,j)}{T-s}. \tag{2}$$

For the long-range cross-correlated processes, the covariance scales as

$$F^2_{\mathrm{DCCA}}(s) \propto s^{2H_{xy}}. \tag{3}$$

The estimate of the bivariate Hurst exponent is obtained by the log–log regression on Eq. (3). Similarly to DFA and MF-DFA, there are several ways of treating overlapping and non-overlapping boxes of length s, compare e.g. Refs. [24,26,33–37]. In the simulations, we use non-overlapping boxes with a step between s equal to 10 due to computational efficiency.

2.2. Height cross-correlation analysis

Kristoufek [22] introduces the multifractal height cross-correlation analysis (MF-HXA) as a bivariate generalization of the height–height correlation analysis [38–40] and the generalized Hurst exponent approach [41–43], which are often labeled simply as HHCA and GHE, respectively.

MF-HXA is constructed to analyze the multifractal properties of bivariate series similarly to MF-DXA. We generalize the q-th order height–height correlation function for two simultaneously recorded series. Let us consider two profiles $\{X_t\}$ and $\{Y_t\}$ with time resolution ν and $t = \nu, 2\nu, \ldots, \nu\lfloor\frac{T}{\nu}\rfloor$, where $\lfloor\rfloor$ is a lower integer sign. For better legibility, 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 analysis of power-law cross-correlations, i.e. the case when $q = 2$, the height–height covariance function is then defined as

$$K_{xy,2}(\tau) = \frac{\nu}{T^*} \sum_{t=1}^{T^*/\nu} |\Delta_\tau X_t Y_t| \equiv \langle |\Delta_\tau X_t Y_t| \rangle \tag{4}$$

where time interval τ generally ranges between $\nu = \tau_{\min}, \ldots, \tau_{\max}$. Scaling relationship between $K_{xy,q}(\tau)$ and the generalized bivariate Hurst exponent $H_{xy}(q)$ is obtained as

$$K_{xy,2}(\tau) \propto \tau^{H_{xy}}. \tag{5}$$