ROBUST ESTIMATION OF THE MEMORY PARAMETER OF A GAUSSIAN TIME SERIES USING WAVELETS

Olaf Kouamo, Céline Lévy-Leduc, Eric Moulines

Institut Telecom & CNRS, LTCI, Telecom ParisTech

ABSTRACT

We propose in this paper robust estimators of the memory parameter d of a (possibly) non stationary Gaussian time series with generalized spectral density f. This generalized spectral density is characterized by the memory parameter d and by a function f^* which specifies the short-range dependence structure of the process. The memory parameter d is estimated by regressing the logarithm of the estimated variance of the wavelet coefficients at different scales. The two robust estimators of d that we consider are based on robust estimators of the variance of the wavelet coefficients, namely the square of the scale estimator proposed by [**?**] and the median of the square of the wavelet coefficients. We establish a Central Limit Theorem for these robust estimators as well as for the estimator of d based on the classical estimator of the variance proposed by [**?**]. The properties of these estimators are also compared on publicly available Internet traffic packet counts data.

Index Terms— Memory parameter estimator, long-range dependence, robustness, wavelet analysis.

1. INTRODUCTION

During the last decades, long-range dependence has been observed in many different fields, including financial econometrics, hydrology or analysis of Internet traffic. In most of these applications, however, the presence of atypical observations is quite common. These outliers might be due to gross errors in the observations but also to unmodeled disturbances; see for example [**?**] and [**?**] for possible explanations of the presence of outliers in Internet traffic analysis. It is well-known that even a few atypical observations can severely affect estimators, leading to incorrect conclusions. Hence, defining robust estimators of the memory parameter which are less sensitive to the presence of additive outliers is a challenging practical problem.

In this paper, we consider the class of fractional processes, denoted $M(d)$ defined as follows. Let $X = \{X_k\}_{k\in\mathbb{Z}}$ be a real-valued Gaussian process, not necessarily stationary and denote by $\mathbf{\Delta} X$ the first order difference of X, defined by $[\mathbf{\Delta} X]_n = X_n - X_{n-1}$, $n \in \mathbb{Z}$. Define, for an integer $K \geq 1$, the Kth order difference recursively as follows : $\mathbf{\Delta}^K = \mathbf{\Delta} \circ \mathbf{\Delta}^{K-1}$. Let f^* be a bounded non-negative symmetric function which is bounded away from zero in a neighborhood of the origin. Following [**?**], we say that X is an $M(d)$ process if for any integer $K > d - 1/2$, $\mathbf{\Delta}^K X$ is stationary with spectral density function

$$f_{\mathbf{\Delta}^K X}(\lambda) = |1 - \mathrm{e}^{-\mathrm{i}\lambda}|^{2(K-d)}\, f^*(\lambda), \quad \lambda \in (-\pi, \pi)\ . \tag{1}$$

Observe that $f_{\mathbf{\Delta}^K X}(\lambda)$ in (**??**) is integrable since $-(K-d) < 1/2$. When $d \geq 1/2$, the process is not stationary. One can nevertheless associate to X the function

$$f(\lambda) = |1 - \mathrm{e}^{-\mathrm{i}\lambda}|^{-2d} f^*(\lambda)\ , \tag{2}$$

which is called a *generalized spectral density function*. In the sequel, we assume that $f^* \in \mathcal{H}(\beta, L)$ with $0 < \beta \leq 2$ and $L > 0$ where $\mathcal{H}(\beta, L)$ denotes the set of non-negative and symmetric functions g satisfying, for all $\lambda \in (-\pi, \pi)$, $|g(\lambda) - g(0)| \leq L\, g(0)\, |\lambda|^\beta$. Our setting is semi-parametric in that both d and f^* in (**??**) are unknown. Here, f^* can be seen as a nuisance parameter whereas d is the parameter of interest. This assumption on f^* is typical in the semi-parametric estimation setting; see for instance [**?**] and [**?**] and the references therein.

Different approaches have been proposed for building robust estimators of the memory parameter for M(d) processes in the semi-parametric setting outlined above. [**?**] have proposed a robustified wavelet based-regression estimator developed by [**?**]; the robustification is achieved by replacing the estimation of the wavelet coefficients variance at different scales by the median of the square of the wavelet coefficients. Another technique to robustify the wavelet regression technique has been outlined in [**?**] which consists in regressing the logarithm of the square of the wavelet coefficients at different scales. [**?**] proposed a robustified version of the log-periodogram regression estimator introduced in [**?**]. The method replaces the log-periodogram of the observation by a robust estimator of the spectral density in the neighborhood of the zero frequency, obtained as the discrete Fourier transform of a robust autocovariance estimator defined in [**?**]. This procedure works well in practice but there is no theoretical support in the semi-parametric context (note however that the consistency and the asymptotic normality of the robust estimator of the covariance have been discussed in [**?**] in several dependence frameworks).

The two robust estimators of d that we propose consist in regressing the logarithm of robust variance estimators of the wavelet coefficients of the process X on a range of scales. We use as robust variance estimators the square of the scale estimator proposed by [**?**] and the square of the *mean absolute deviation* (MAD). These estimators are a robust alternative to the estimator of d proposed by [**?**] which uses the same method but with the classical variance estimator. Here, we derive a Central Limit Theorem (CLT) for the two robust estimators of d and also for the estimator of d proposed by [**?**].

The paper is organized as follows. In Section **??**, we introduce the wavelet setting and define the wavelet based regression estimators of d. Section **??** is dedicated to the asymptotic properties of the robust estimators of d. An Internet traffic packet counts dataset collected from the University of North Carolina, Chapel is studied as an application in Section **??**.

2. DEFINITION OF THE ESTIMATORS OF THE MEMORY PARAMETER

2.1. The wavelet setting

The wavelet setting involves two functions ϕ and ψ in $\mathrm{L}^2(\mathbb{R})$ and their Fourier transforms

$$\widehat{\phi}(\xi) \stackrel{\mathrm{def}}{=} \int_{-\infty}^{\infty} \phi(t) \mathrm{e}^{-\mathrm{i}\xi t}\,\mathrm{d}t \quad \text{and} \quad \widehat{\psi}(\xi) \stackrel{\mathrm{def}}{=} \int_{-\infty}^{\infty} \psi(t) \mathrm{e}^{-\mathrm{i}\xi t}\,\mathrm{d}t\,. \tag{3}$$

Assume the following:

(W-1) ϕ and ψ are compactly-supported, integrable, and $\widehat{\phi}(0) = \int_{-\infty}^{\infty} \phi(t)\,\mathrm{d}t = 1$ and $\int_{-\infty}^{\infty} \psi^2(t)\,\mathrm{d}t = 1$.

(W-2) There exists $\alpha > 1$ such that $\sup_{\xi\in\mathbb{R}} |\widehat{\psi}(\xi)|\,(1+|\xi|)^{\alpha} < \infty$.

(W-3) The function ψ has M vanishing moments, *i.e.* $\int_{-\infty}^{\infty} t^m \psi(t)\,dt = 0$ for all $m = 0, \dots, M-1$.

(W-4) The function $\sum_{k\in\mathbb{Z}} k^m \phi(\cdot - k)$ is a polynomial of degree m for all $m = 0, \dots, M-1$.

Condition **??** ensures that the Fourier transform $\widehat{\psi}$ decreases quickly to zero. Condition **??** ensures that ψ oscillates and that its scalar product with continuous-time polynomials up to degree $M-1$ vanishes. It is equivalent to asserting that the first $M-1$ derivatives of $\widehat{\psi}$ vanish at the origin and hence

$$|\widehat{\psi}(\lambda)| = O(|\lambda|^M)\,, \text{ as } \lambda \to 0\,. \tag{4}$$

Daubechies wavelets (with $M \geq 2$) and the Coiflets satisfy these conditions, see [**?**]. Viewing the wavelet $\psi(t)$ as a basic template, define the family $\{\psi_{j,k}, j \in \mathbb{Z}, k \in \mathbb{Z}\}$ of translated and dilated functions

$$\psi_{j,k}(t) = 2^{-j/2}\,\psi(2^{-j}t - k),\; j \in \mathbb{Z},\, k \in \mathbb{Z}\,. \tag{5}$$

Positive values of k translate ψ to the right, negative values to the left. The *scale index* j dilates ψ so that large values of j correspond to coarse scales and hence to low frequencies. We suppose throughout the paper that

$$(1+\beta)/2 - \alpha < d \leq M\,. \tag{6}$$

We now describe how the wavelet coefficients are defined in discrete time, that is for a real-valued sequence $\{x_k, k \in \mathbb{Z}\}$ and for a finite sample $\{x_k, k = 1, \dots, n\}$. Using the scaling function ϕ, we first interpolate these discrete values to construct the following continuous-time functions

$$\mathbf{x}_n(t) \stackrel{\mathrm{def}}{=} \sum_{k=1}^{n} x_k\,\phi(t-k) \quad \text{and} \quad \mathbf{x}(t) \stackrel{\mathrm{def}}{=} \sum_{k\in\mathbb{Z}} x_k\,\phi(t-k), \quad t \in \mathbb{R}\,. \tag{7}$$

Without loss of generality we may suppose that the support of the scaling function ϕ is included in $[-\mathrm{T}, 0]$ for some integer $\mathrm{T} \geq 1$. Then

$$\mathbf{x}_n(t) = \mathbf{x}(t) \quad \text{for all} \quad t \in [0, n - \mathrm{T} + 1]\,.$$

We may also suppose that the support of the wavelet function ψ is included in $[0, \mathrm{T}]$. With these conventions, the support of $\psi_{j,k}$ is included in the interval $[2^j k, 2^j(k+\mathrm{T})]$. The wavelet coefficient $W_{j,k}$ at scale $j \geq 0$ and location $k \in \mathbb{Z}$ is formally defined as the scalar product in $\mathrm{L}^2(\mathbb{R})$ of the function $t \mapsto \mathbf{x}(t)$ and the wavelet $t \mapsto \psi_{j,k}(t)$:

$$W_{j,k} \stackrel{\mathrm{def}}{=} \int_{-\infty}^{\infty} \mathbf{x}(t)\psi_{j,k}(t)\,\mathrm{d}t = \int_{-\infty}^{\infty} \mathbf{x}_n(t)\psi_{j,k}(t)\,\mathrm{d}t, \quad j \geq 0, k \in \mathbb{Z}\,, \tag{8}$$

when $[2^j k, 2^j k + \mathrm{T}] \subseteq [0, n - \mathrm{T} + 1]$, that is, for all $(j,k) \in \mathcal{I}_n$, where

$$\mathcal{I}_n \stackrel{\mathrm{def}}{=} \{(j,k) :\, j \geq 0, 0 \leq k \leq n_j - 1\} \quad \text{with } n_j = [2^{-j}(n - \mathrm{T} + 1) - \mathrm{T} + 1]\,. \tag{9}$$

If $\mathbf{\Delta}^M X$ is stationary, then from [**?**, Eq (17)] the process $\{W_{j,k}\}_{k\in\mathbb{Z}}$ of wavelet coefficients at scale $j \geq 0$ is stationary but the two–dimensional process $\{[W_{j,k},\, W_{j',k}]^T\}_{k\in\mathbb{Z}}$ of wavelet coefficients at scales j and j', with $j \geq j'$, is not stationary. Here ^{T} denotes the transposition. This is why we consider instead the stationary *between-scale* process

$$\{[W_{j,k},\, \mathbf{W}_{j,k}(j-j')^T]^T\}_{k\in\mathbb{Z}}\,, \tag{10}$$

where $\mathbf{W}_{j,k}(j-j')$ is defined as follows:

$$\mathbf{W}_{j,k}(j-j') \stackrel{\mathrm{def}}{=} \Big[W_{j',2^{j-j'}k},\, W_{j',2^{j-j'}k+1},\, \dots,\, W_{j',2^{j-j'}k+2^{j-j'}-1}\Big]^T\,.$$

For all $j, j' \geq 1$, the covariance function of the between scale process is given by

$$\mathrm{Cov}(\mathbf{W}_{j,k'}(j-j'), W_{j,k}) = \int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}\lambda(k-k')}\,\mathbf{D}_{j,j-j'}(\lambda; f)\,d\lambda\,, \tag{11}$$

where $\mathbf{D}_{j,j-j'}(\lambda; f)$ stands for the cross-spectral density function of this process. For further details, we refer the reader to [**?**, Corollary 1]. The case $j = j'$ corresponds to the spectral density function of the *within-scale* process $\{W_{j,k}\}_{k\in\mathbb{Z}}$.

In the sequel, we shall use that the within- and between-scale spectral densities $\mathbf{D}_{j,j-j'}(\lambda; d)$ of the process X with memory parameter $d \in \mathbb{R}$ can be approximated by the corresponding spectral density of the generalized fractional Brownian motion $B_{(d)}$ defined, for $d \in \mathbb{R}$ and $u \in \mathbb{N}$, by

$$\begin{aligned}\mathbf{D}_{\infty,u}(\lambda; d) &= \Big[\mathbf{D}^{(0)}_{\infty,u}(\lambda; d), \dots, \mathbf{D}^{(2^u-1)}_{\infty,u}(\lambda; d)\Big] \\ &= \sum_{l\in\mathbb{Z}} |\lambda + 2l\pi|^{-2d}\,\mathbf{e}_u(\lambda + 2l\pi)\,\overline{\widehat{\psi}(\lambda + 2l\pi)}\widehat{\psi}(2^{-u}(\lambda + 2l\pi))\,,\end{aligned} \tag{12}$$

where,

$$\mathbf{e}_u(\xi) \stackrel{\mathrm{def}}{=} 2^{-u/2}\,[1, \mathrm{e}^{-\mathrm{i}2^{-u}\xi}, \dots, \mathrm{e}^{-\mathrm{i}(2^u-1)2^{-u}\xi}]^T, \quad \xi \in \mathbb{R}\,.$$

For further details, see [**?**, p. 307].

2.2. Definition of the robust estimators of the memory parameter

Let us now define robust estimators of the memory parameter d of the M(d) process X from the observations $X_1, \dots, X_n$. These estimators are derived from the [**?**] construction, and consists in regressing estimators of the scale spectrum

$$\sigma_j^2 \stackrel{\mathrm{def}}{=} \mathrm{Var}(W_{j,0}) \tag{13}$$

with respect to the scale index j. More precisely, if $\widehat{\sigma}^2_j$ is an estimator of σ^2_j based on $W_{j,0:n_j-1} = (W_{j,0}, \ldots, W_{j,n_j-1})$ then an estimator of the memory parameter d is obtained by regressing $\log(\widehat{\sigma}^2_j)$ for a finite number of scale indices $j \in \{J_0, \ldots, J_0 + \ell\}$ where $J_0 = J_0(n) \geq 0$ is the lower scale and $1 + \ell \geq 2$ is the number of scales in the regression. The regression estimator can be expressed formally as

$$\widehat{d}_n(J_0, \mathbf{w}) \stackrel{\text{def}}{=} \sum_{j=J_0}^{J_0+\ell} w_{j-J_0} \log\left(\widehat{\sigma}^2_j\right) , \tag{14}$$

where the vector $\mathbf{w} \stackrel{\text{def}}{=} [w_0, \ldots, w_\ell]^T$ of weights satisfies $\sum_{i=0}^{\ell} w_i = 0$ and $2\log(2)\sum_{i=0}^{\ell} i w_i = 1$, see [?] and [?]. For $J_0 \geq 1$ and $\ell > 1$, one may choose for example $\mathbf{w}$ corresponding to the least squares regression matrix, defined by $\mathbf{w} = DB(B^T DB)^{-1}\mathbf{b}$ where

$$\mathbf{b} \stackrel{\text{def}}{=} \begin{bmatrix} 0 & (2\log(2))^{-1} \end{bmatrix}, \quad B \stackrel{\text{def}}{=} \begin{bmatrix} 1 & 1 & \ldots & 1 \\ 0 & 1 & \ldots & \ell \end{bmatrix}^T$$

is the design matrix and D is an arbitrary positive definite matrix. The best choice of D depends on the memory parameter d. However a good approximation of this optimal matrix D is the diagonal matrix with diagonal entries $D_{i,i} = 2^{-i}$, $i = 0 \ldots, \ell$; see [?] and the references therein. We will use this choice of the design matrix in the numerical experiments. A heuristic justification for this choice is that by [?, Eq. (28)],

$$\sigma^2_j \sim C\, 2^{2jd} \text{ , as } j \to \infty , \tag{15}$$

where C is a positive constant. In the sequel, we shall consider three different estimators of d based on three different estimators of the scale spectrum σ^2_j with respect to the scale index j which are defined below.

2.2.1. Classical scale estimator

This estimator has been considered in the original contribution of [?] and consists in estimating the scale spectrum σ^2_j with respect to the scale index j by the empirical variance

$$\widehat{\sigma}^2_{\mathrm{CL},j} = \frac{1}{n_j}\sum_{i=1}^{n_j} W^2_{j,i} , \tag{16}$$

where for any j, n_j denotes the number of available wavelet coefficients at scale index j defined in (**??**).

2.2.2. Median absolute deviation

This estimator is well-known to be a robust estimator of the scale as explained in [?]. Since the wavelet coefficients $W_{j,i}$ are centered Gaussian observations, the square of the median absolute deviation of $W_{j,0:n_j-1}$ is defined by

$$\widehat{\sigma}^2_{\mathrm{MAD},j} = \left(m(\Phi) \operatorname*{med}_{0\leq i\leq n_j-1} |W_{j,i}|\right)^2 , \tag{17}$$

where Φ denotes the c.d.f of a standard Gaussian random variable and

$$m(\Phi) = 1/\Phi^{-1}(3/4) = 1.4826 . \tag{18}$$

2.2.3. The Croux and Rousseeuw estimator

This estimator is another robust scale estimator introduced in [?]. Its asymptotic properties in several dependence contexts have been further studied in [?] and the square of this estimator is defined by

$$\widehat{\sigma}^2_{\mathrm{CR},j} = \left(c(\Phi)\{|W_{j,i} - W_{j,k}|;\ 0 \leq i,k \leq n_j - 1\}_{(k_{n_j})}\right)^2 , \tag{19}$$

where $c(\Phi) = 2.21914$ and $k_{n_j} = \lfloor n^2_j/4 \rfloor$. That is, up to the multiplicative constant $c(\Phi)$, $\widehat{\sigma}_{\mathrm{CR},j}$ is the k_{n_j}th order statistics of the n^2_j distances $|W_{j,i} - W_{j,k}|$ between all the pairs of observations.

3. ASYMPTOTIC PROPERTIES OF THE ROBUST ESTIMATORS OF THE MEMORY PARAMETER

We derive, in this section, a Central Limit Theorem for the robust wavelet-based regression estimators of d defined by

$$\widehat{d}_{*,n}(J_0, \mathbf{w}) \stackrel{\text{def}}{=} \sum_{j=J_0}^{J_0+\ell} w_{j-J_0} \log\left(\widehat{\sigma}^2_{*,j}\right) , \tag{20}$$

where $\widehat{\sigma}^2_{*,j}$ are given for $* = \mathrm{CL}$, MAD and CR by (**??**), (**??**) and (**??**), respectively.

Theorem 1. *Assume that X is a Gaussian $M(d)$ process with generalized spectral density function defined in* (**??**) *such that $f^* \in \mathcal{H}(\beta, L)$ for some $L > 0$ and $0 < \beta \leq 2$. Assume that* **??-??** *hold with d, α and M satisfying* (**??**)*. Let $W_{j,k}$ be the wavelet coefficients associated to X defined by* (**??**)*. If $n \mapsto J_0(n)$ is an integer valued sequence satisfying $J_0(n) \to \infty$ and $n2^{-J_0(n)} \to \infty$, as $n \to \infty$, and if*

$$n2^{-(1+2\beta)J_0(n)} \to 0 \text{ , as } n \to \infty, \tag{21}$$

then, $\widehat{d}_{*,n}(J_0, \mathbf{w})$ satisfies the following Central Limit Theorem:

$$\sqrt{n2^{-J_0(n)}}\left(\widehat{d}_{*,n}(J_0, \mathbf{w}) - d\right) \xrightarrow{d} \mathcal{N}\left(0, \mathbf{w}^T \mathbf{V}_*(d)\mathbf{w}\right) , \tag{22}$$

where $\mathbf{V}_*(d)$ is the $(1 + \ell) \times (1 + \ell)$ matrix defined by

$$\mathbf{V}_{*,i,j}(d) = \sum_{p\geq 2} \frac{4c^2_p(\mathrm{IF}_*)}{p!\,\mathrm{K}(d)^p} 2^{pd|i-j|+i\wedge j}$$
$$\sum_{\tau\in\mathbb{Z}} \sum_{r=0}^{2^{|i-j|}-1} \left(\int_{-\pi}^{\pi} \mathbf{D}^{(r)}_{\infty,|i-j|}(\lambda; d)\mathrm{e}^{\mathrm{i}\lambda\tau}\mathrm{d}\lambda\right)^p , \ 0 \leq i,j \leq \ell . \tag{23}$$

In (**??**)*, $\mathrm{K}(d) = \int_{\mathbb{R}} |\xi|^{-2d}|\widehat{\psi}(\xi)|\mathrm{d}\xi$, $\mathbf{D}_{\infty,|i-j|}(\cdot; d)$ is the cross-spectral density defined in* (**??**)*, $c_p(\mathrm{IF}_*) = \mathbb{E}[\mathrm{IF}(X, *, \Phi)H_p(X)]$, where H_p is the pth Hermite polynomial and $\mathrm{IF}(\cdot, *, \Phi)$ is defined as follows:*

$$\mathrm{IF}\,(x, \mathrm{CL}, \Phi) = \frac{1}{2}H_2(x),$$
$$\mathrm{IF}\,(x, \mathrm{CR}, \Phi) = c(\Phi)\left(\frac{1/4 - \Phi(x + 1/c(\Phi)) + \Phi(x - 1/c(\Phi))}{\int_{\mathbb{R}} \varphi(y)\varphi(y + 1/c(\Phi))\mathrm{d}y}\right) ,$$
$$\mathrm{IF}(x, \mathrm{MAD}, \Phi) = -m(\Phi)\left(\frac{{}_{\{|x|\leq 1/m(\Phi)\}} - 1/2}{2\varphi(1/m(\Phi))}\right) ,$$

where φ denotes the p.d.f of the standard Gaussian random variable, $m(\Phi)$ and $c(\Phi)$ being defined in (**??**) *and* (**??**)*, respectively and $H_2(x) = x^2 - 1$ is the second Hermite polynomial.*

The proof of this theorem is omitted here for reasons of space.

4. APPLICATION TO INTERNET TRAFFIC PACKET COUNTS DATA

In this section, we analyze an Internet traffic packet counts dataset collected at the University of North Carolina, Chapel (UNC) which has also been studied by [**?**]. This dataset is publicly available and can be downloaded from the following website http://netlab.cs.unc.edu/public/old_research/net_lrd/.

Figure **??** displays a packet count time series measured at the link of UNC on April 11, a Thursday, from 1 p.m. to 3 p.m., 2002 (Thu1300). These packet counts were measured every 1 millisecond but, for a better display, we aggregated them at 1 second.

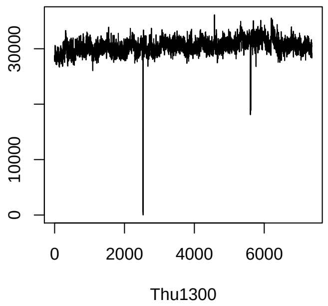

Fig. 1. Packet counts of aggregated traffic every 1 second.

The maximal available scale for this dataset is 20. Since we have less than 4 observations at this scale, we set the coarse scale $J_0 + \ell = 19$ and vary the finest scale J_0 from 1 to 17.

In Figure **??**, we display the estimates $\widehat{d}_{n,\mathrm{CL}}$, $\widehat{d}_{n,\mathrm{CR}}$ and $\widehat{d}_{n,\mathrm{MAD}}$ of the memory parameter d as well as their respective 95% confidence intervals from $J_0 = 1$ to $J_0 = 14$. We propose to choose $J_0 = 9$ for Thu1300 since from these values of J_0 the successive confidence intervals are such that the smallest one is included in the largest one (for the robust estimators). Note that [**?**] chose the same values of J_0 using another methodology. For these values of J_0 we obtain $\widehat{d}_{n,\mathrm{CL}} = 0.43$ (with 95% confidence interval [0.412, 0.443]), $\widehat{d}_{n,\mathrm{CR}} = 0.37$ (with 95% confidence interval [0.358, 0.385]) and $\widehat{d}_{n,\mathrm{MAD}} = 0.38$ with (95% confidence interval [0.362, 0.397]). These values are similar to the one found by [**?**].

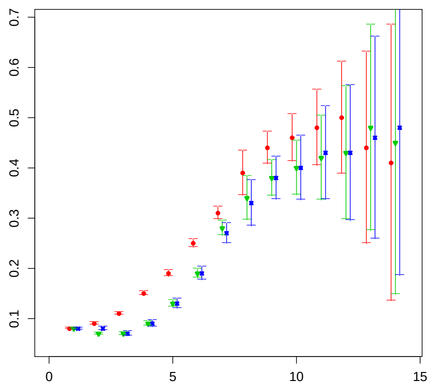

Fig. 2. Confidence intervals of the estimates $\widehat{d}_{n,\mathrm{CL}}$ (red), $\widehat{d}_{n,\mathrm{CR}}$ (green) and $\widehat{d}_{n,\mathrm{MAD}}$ (blue) on the data Thu1300 for $J_0 = 1, \ldots, 14$.

With this choice of J_0 for Thu1300, we observe a significant difference between the classical estimator and the robust estimators. Thus to better understand the influence of outliers on the estimated memory parameter a new dataset with artificial outliers was generated. The Thu1300 time series shows two spikes shooting down. Especially, the first downward spike hits zero. [**?**] have shown that this dropout lasted 8 seconds. Outliers are introduced by dividing by 6 the 8000 observations in this period. The new memory parameter estimators are $\widehat{d}_{n,\mathrm{CL}} = 0.445$, $\widehat{d}_{n,\mathrm{CR}} = 0.375$ and $\widehat{d}_{n,\mathrm{MAD}} = 0.377$. As expected, the classical estimator was affected while the robust estimators remain stable.