Robust estimation in time series

V. A. Reisen

NuMes, D.E; Universidade Federal do Espirito Santo, Brazil.

20 de setembro de 2011

Contents

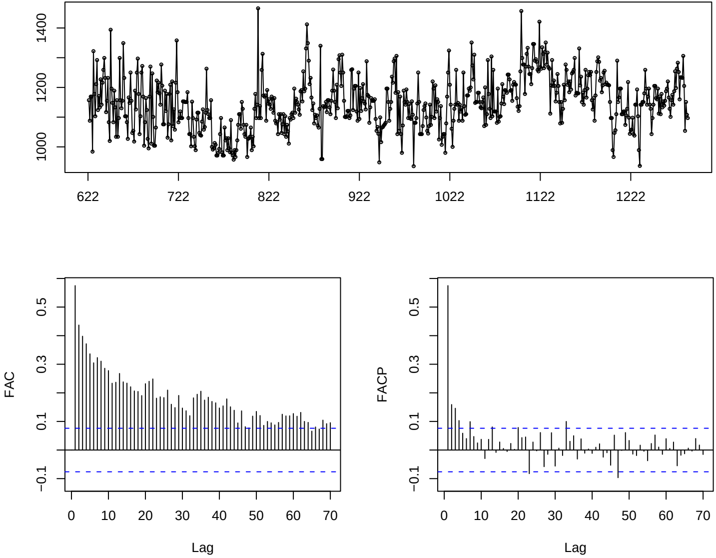
Abstract

A desirable property of an autocovariance estimator is to be robust to the presence of additive outliers. It is well-known that the sample autocovariance, based on the moments, does not own this property. Hence, the definition of an autocovariance estimator which is robust to additive outlier can be very useful for time-series modeling. In this paper, some asymptotic properties of the robust scale and autocovariance estimators proposed by Ma & Genton (2000) is study and applied to time series with different correlation structures.

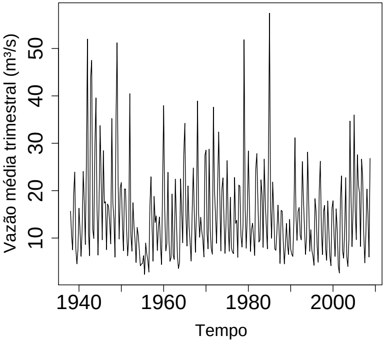
Introduction-References bases of this talk are:

- FAJARDO M., F. A., REISEN, V. A., CRIBARI NETO, Francisco. Robust estimation in Long-memory processes under additive outliers. Journal of Statistical Planning and Inference, 139 , 2511 - 2525, 2009.
- SARGNAGLIA, A. REISEN, V. A, C. LÉVY-LEDUC. Robust estimation in PAR models in the presence of additive outliers, Journal of Multivariate Analysis, 2, 2168-2183, 2010.
- C. LÉVY-LEDUC. H. BOISTARD, MOULINES, E. M. S. TAQQU. and REISEN, V. A. Asymptotic properties of U-processes in long-range dependence. The Annals of Statistics, 39(3),1399-1246, 2011
- C. LÉVY-LEDUC, H. BOISTARD, , MOULINES, E. MURAD S TAQQU and REISEN, V. A. Robust estimation of the scale and the autocovariance functions in short and long-range dependence. Journal of Time Series Analysis,32 (2),135-156. 2011.
- LÉVY-LEDUC, H. BOISTARD, MOULINES, E. MURAD S TAQQU and REISEN, V. A. Large sample behavior of some well-known robust estimators under long-range dependence,Statistics, 45(1),59-71, 2011.

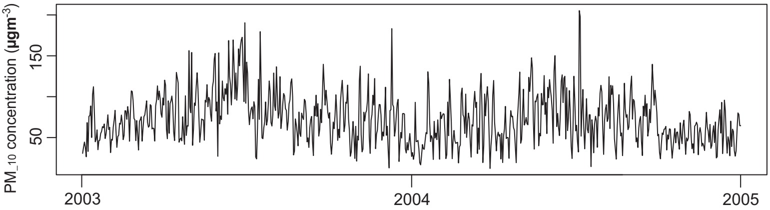
Applications: Nile river 622 - 1281 D.C.



Aplications: Quarterly mean flow of Castelo River, Castelo-ES



Aplications: The daily average PM_{10} concentration- Vitoria-ES



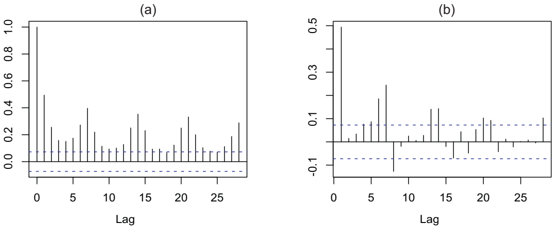


Figura: ACF (a) and PACF (b).

Introduction-some references

- Haldrup & Nielsen (2007) evaluated the impact of measurement errors, outliers and structural breaks on the long-memory parameter estimation.
- Sun & Phillips (2003) suggested the use of a approach adding a nonlinear factor to the log-periodogram regression, as a way to minimize any existing bias.
- Agostinelli & Bisaglia (2003) proposed the use of a weighted maximum likelihood approach as a modification of the the estimator proposed by Beran(1994).

The impact of outliers in stationary processes

Let $\{X_t\}_{t\in\mathbb{Z}}$ be a stationary process and let $\{z_t\}_{t\in\mathbb{Z}}$ be a process contaminated by additive outliers, which is described by

$$z_t = X_t + \sum_{i=1}^{m} \omega_i I_t^{(T_i)}, \tag{1}$$

where m is the number of outliers; the unknown parameter ω_i represents the magnitude of the ith outlier at time T_i, and $I_t^{(T_i)}$ is a Bernoulli random variable with probability distribution $\Pr\left(I_t^{(T_i)} = -1\right) = \Pr\left(I_t^{(T_i)} = 1\right) = \frac{p_i}{2}$ and $\Pr\left(I_t^{(T_i)} = 0\right) = 1 - p_i$. The random variables X_t and $I_t^{(T_i)}$ are independent.

Proposition 1.

Suppose that $\{z_t\}$ follows (1) and X_t has spectral density f_X.

i. The autocovariance function (ACOVF) of $\{z_t\}$ is given by

$$\gamma_z(h) = \begin{cases} \gamma_X(0) + \sum_{i=1}^{m} \omega_i^2 p_i, & \text{if } \ h = 0, \\ \gamma_X(h), & \text{if } \ h \neq 0. \end{cases}$$

ii. The spectral density function of $\{z_t\}$ is given by

$$f_z(\lambda) = f_X(\lambda) + \frac{1}{2\pi} \sum_{i=1}^{m} \omega_i^2 p_i, \quad \lambda \in [-\pi, \pi].$$

Proposition 2.

Let $z_1, z_2, \ldots, z_n$ be a set of observations generated from model (1) with $m = 1$, and let the outlier occurs at time $t = T$. It follows that:

i. The sample ACOVF is given by

$$\widehat{\gamma}_Z(h) = \widehat{\gamma}_X(h) \pm \frac{\omega}{n}(X_{T-h} + X_{T+h} - 2\bar{y}) + \frac{\omega^2}{n}\delta(h) + o_p(n^{-1}), \tag{2}$$

where $\widehat{\gamma}_X(h) = \frac{1}{n}\sum_{t=1}^{n-h}(X_t - \bar{X})(X_{t+h} - \bar{X})$ and

$$\delta(h) = \begin{cases} 1, & \text{when } h = 0, \\ 0, & \text{otherwise.} \end{cases}$$

Proposition 2 (continuation).

ii. The periodogram is given by

$$I_z(\lambda) = I_X(\lambda) + \Delta(\omega), \quad [-\pi, \pi],$$

where

$$\Delta(\omega) = \frac{\omega^2}{2\pi n} \pm \frac{\omega}{\pi n}\left\{(X_{T} - \bar{X}) + \sum_{h=1}^{n-1}(X_{T-h} + X_{T+h} - 2\bar{X})\cos(h\lambda)\right\} \tag{3}$$

$$+ o_p(n^{-1}).$$

Proposition 3. (Chan (1992, 1995))

Suppose that $z_1, z_2, \ldots, z_n$ is a set of observations generated from model (1) and let $\widehat{\rho}_z(h) = \widehat{\gamma}_z(h)/\widehat{\gamma}_z(0)$, then

i. For $m = 1$,

$$\lim_{n\to\infty} \lim_{\omega\to\infty} \widehat{\rho}_z(h) = 0.$$

ii. For $m = 2$ and $T_2 = T_1 + l$, such that $h < T_1 < T_1 + l < n - h$, we have

$$\lim_{n\to\infty} \left\{ plim_{\substack{\omega_1\to\infty \\ \omega_2\to\pm\infty}} \widehat{\rho}_z(h) \right\} = \begin{cases} 0, & \text{if } h \neq l, \\ \pm 0.5, & \text{if } h = l. \end{cases}$$

Some specific comments

- The outliers cause an increase in the variance of process, which reduces the magnitude of the autocorrelations and introduces loss of information on the pattern of serial correlation.
- The spectral density of the process is characterized by an translation due to the contributions of magnitude of outliers.
- These results also give the evidence that an outlier can seriously affect the autocorrelation structure due to an increase in the variance.

Lemma 1.

Let $\{X_t\}_{t\in\mathbb{Z}}$ be a stationary and invertible ARFIMA(p,d,q) process. Also, let $\{z_t\}_{t\in\mathbb{Z}}$ be such that $z_t = X_t + \sum_{i=1}^{m}\omega_i I_t^{(T)}$, where m is the number of outliers, the unknown parameter ω_i is the magnitude of the ith outlier at time T_i and $I_t^{(T_i)}$ is Bernoulli distributed: $\Pr\left(I_t^{(T_i)}=-1\right) = \Pr\left(I_t^{(T_i)}=1\right) = \frac{p_i}{2}$ and $\Pr\left(I_t^{(T_i)}=0\right) = 1-p_i$. The spectral density of $\{z_t\}$ is given by

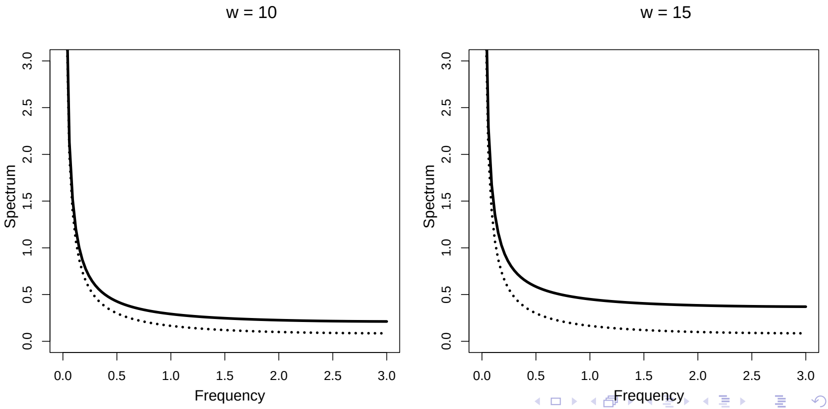
$$f_z(\lambda) = \frac{\sigma_\epsilon^2}{2\pi}\frac{|\Theta(e^{-i\lambda})|^2}{|\Phi(e^{-i\lambda})|^2}\left\{2\sin\left(\frac{\lambda}{2}\right)\right\}^{-2d} + \frac{1}{2\pi}\sum_{i=1}^{m}\omega_i^2 p_i.$$

where $\lambda\in[-\pi,\pi]$.

Spectrum of ARFIMA$(0, d, 0)$ model with $d = 0.3$

The dotted line is the spectral density of the outlier-free process and the solid line is the spectral density of the process under an additive outlier. The contaminated series is obtained by replacing 5% of the observations with additive outliers using $w = 10, 15$.

Theoretical spectrum (dot line) and contaminated spectrum (solid line)



A robust estimator of *ACF*

Rousseeuw & Croux (1993) proposed a robust scale estimator function which is based on the kth order statistic of $\binom{n}{2}$ distances $\{|y_i - y_j|, i < j\}$, and can be written as

$$Q_n(y) = c \times \{|y_i - y_j|; i < j\}_{(k)}, \tag{4}$$

where $y = (y_1, y_2, \ldots, y_n)'$, c is a constant used to guarantee consistency ($c = 2.2191$ for the normal distribution), and $k = \left\lfloor \frac{\binom{n}{2}+2}{4} \right\rfloor + 1$. The above function can be calculated using the algorithm proposed by Croux & Rousseeuw (1992), which is computationally efficient. Rousseeuw & Croux (1993) showed that the asymptotic breakdown point of $Q_n(\cdot)$ is 50%, which means that the time series can be contaminated by up to half of the observations with outliers and $Q_n(\cdot)$ will still yield sensible estimates.

A robust estimator of ACF- continuation

$Q(\cdot)$, Ma & Genton (2000) proposed a highly robust estimator for the ACOVF:

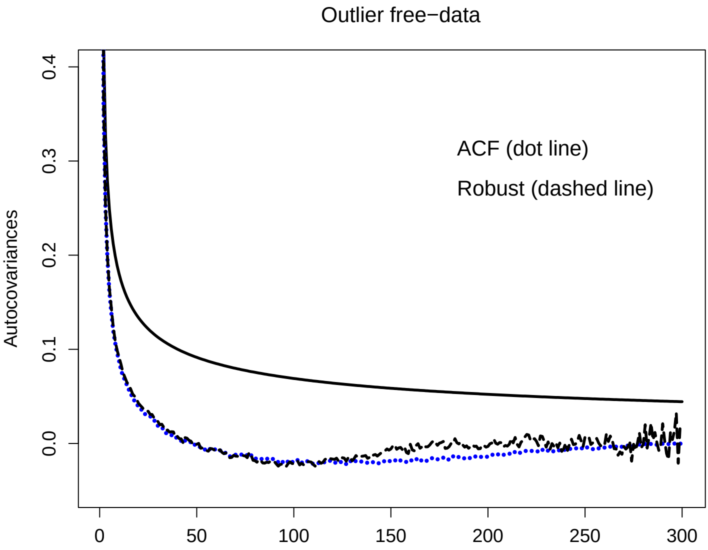
$$\widetilde{\gamma}(h) = \frac{1}{4}\left[Q^2_{n-h}(u+v) - Q^2_{n-h}(u-v)\right], \tag{5}$$

where u and v are vectors containing the initial $n-h$ and the final $n-h$ observations, respectively. The robust estimator for the autocorrelation function is

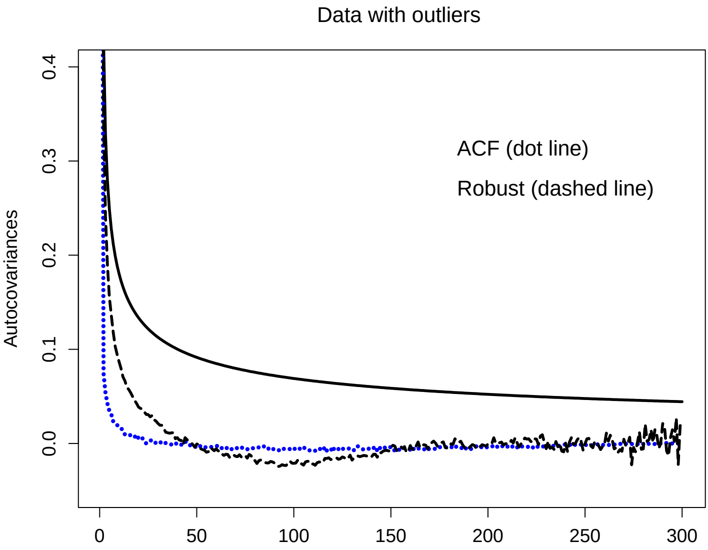
$$\widetilde{\rho}(h) = \frac{Q^2_{n-h}(u+v) - Q^2_{n-h}(u-v)}{Q^2_{n-h}(u+v) + Q^2_{n-h}(u-v)}.$$

It can be shown that $|\widetilde{\rho}(h)| \leq 1$ for all h.

ACF ARFIMA$(0, d, 0)$ model, $d = 0.3$, $n = 300$



ACF ARFIMA$(0, d, 0)$ model, $d = 0.3$, $n = 300$



Theoretical Results-Short-memory and long-memory cases

It supposes that the empirical c.d.f. F_n, adequately normalized, converges. Let us first define the Influence Function. Following Huber (1981), the influence function $x \mapsto \mathrm{IF}(x, T, F)$ is defined for a functional T at a distribution F at point x as the limit

$$\mathrm{IF}(x, T, F) = \lim_{\varepsilon \to 0+} \varepsilon^{-1}\{T(F + \varepsilon(\delta_x - F)) - T(F)\} ,$$

where δ_x is the Dirac distribution at x. Influence functions are a classical tool in robust statistics used to understand the effect of a small contamination at the point x on the estimator.

Theoretical Results-Short-memory case

$(X_i)_{i\geq 1}$ is a stationary mean-zero Gaussian process with auto-covariance sequence $\gamma(h) = \mathbb{E}(X_1 X_{h+1})$ satisfying:

$$\sum_{h\geq 1} |\gamma(h)| < \infty\ .$$

Theorem

Under some assumption $Q_n(X_{1:n})$ satisfies the following central limit theorem:

$$\sqrt{n}(Q_n(X_{1:n}) - \sigma) \to \mathcal{N}(0, \widetilde{\sigma}^2) \ ,$$

where $\sigma = \sqrt{\gamma(0)}$ and the limiting variance $\widetilde{\sigma}^2$ is given by

$$\gamma(0)\mathbb{E}[\mathrm{IF}^2(X_1/\sigma, Q, \Phi)] + 2\gamma(0) \sum_{k \geq 1} \mathbb{E}[\mathrm{IF}(X_1/\sigma, Q, \Phi)\mathrm{IF}(X_{k+1}/\sigma, Q, \Phi)]$$

$\mathrm{IF}(\cdot, Q, \Phi)$ is the Influence Function defined previously.

Theorem

Let h be a non negative integer. Under some assumptions the autocovariance estimator $\widehat{\gamma}_Q(h, X_{1:n}, \Phi)$ *satisfies the following Central Limit Theorem:*

$$\sqrt{n}\,(\widehat{\gamma}_Q(h, X_{1:n}, \Phi) - \gamma(h)) \longrightarrow \mathcal{N}(0, \breve{\sigma}_h^2)\ ,$$

where

$$\breve{\sigma}^2(h) = \mathbb{E}[\psi^2(X_1, X_{1+h})] + 2\sum_{k\geq 1} \mathbb{E}[\psi(X_1, X_{1+h})\psi(X_{k+1}, X_{k+1+h})] \tag{6}$$

where ψ *is a function of* $\gamma(h)$ *and IF. (See, Theorem 4 in Leduc, Boistard, Moulines, Taqqu and Reisen (2011)).*

Main theoretical Results-Long-memory case

Now, let $(X_i)_{i\geq 1}$ be a stationary mean-zero Gaussian process with autocovariance $\gamma(h) = \mathbb{E}(X_1 X_{h+1})$ satisfying:

$$\gamma(h) = h^{-D}L(h),\ 0 < D < 1\ ,$$

where L is slowly varying at infinity and is positive for large h.A classical model for long memory process is the so-called ARFIMA(p, d, q), which is a natural generalization of standard ARIMA(p, d, q) models. By allowing d to assume any value in $(-1/2, 1/2)$. $D = 1 - 2d$ in above.

Theorem

(Theorem 8 in Leduc et all (2011)) Let h be a non negative integer and under some assumptions the robust autovariance of $(X_i)_{i\geq 1}$, $\widehat{\gamma}_Q(h, X_{1:n}, \Phi)$, satisfies the following limit theorems as n tends to infinity.

(i) *If $D > 1/2 (d < 1/4)$,*

$$\sqrt{n}\,(\widehat{\gamma}_Q(h, X_{1:n}, \Phi) - \gamma(h)) \xrightarrow{d} \mathcal{N}(0, \check{\sigma}^2(h)) \ ,$$

where

$$\check{\sigma}^2(h) = \mathbb{E}[\psi^2(X_1, X_{1+h})] + 2\sum_{k\geq 1} \mathbb{E}[\psi(X_1, X_{1+h})\psi(X_{k+1}, X_{k+1+h})] \ ,$$

ψ being defined previously.

(ii) *If $D < 1/2$,*

$$\beta(D)\frac{n^D}{\widetilde{L}(n)}\,(\widehat{\gamma}_Q(h, X_{1:n}, \Phi) - \gamma(h)) \xrightarrow{d} \frac{\gamma(0) + \gamma(h)}{2}(Z_2 - Z_1^2)$$

In the theorem, $\beta(D) = \mathrm{B}((1-D)/2, D)$, B denotes the Beta function, the processes $Z_{1,D}(\cdot)$ (the standard fractional Brownian motion) and $Z_{2,D}(\cdot)$ (the Rosenblatt process) are defined in the Levy-Leduc et al (2011), and

$$\widetilde{L}(n) = 2L(n) + L(n+h)(1+h/n)^{-D} + L(n-h)(1-h/n)^{-D}$$

Proposition

Under Assumption and $D < 1/2$ for the process $(X_i)_{i\geq 1}$, the robust autocovariance estimator $\widehat{\gamma}_Q(h, X_{1:n}, \Phi)$ has the same asymptotic behavior as the classical autocovariance estimator. There is no loss of efficiency.

An Application: Long-memory parameter estimators-d estimators

The GPH estimator (Geweke and Porter-Hudak (1983)) is given by

$$d_{GPH} = -\frac{\sum_{j=1}^{g(n)} (x_j - \bar{x}) \log I(\lambda_j)}{\sum_{j=1}^{g(n)} (x_j - \bar{x})^2}, \tag{7}$$

where $x_j = \log \left\{ 2 \sin \left(\frac{\lambda_j}{2} \right) \right\}^2$, $g(n)$ being the bandwidth in the regression equation which has to satisfy $g(n) \to \infty$, $n \to \infty$, with $\frac{g(n)}{n} \to 0$.

The GPH estimator

Hurvich, Deo, Brodsky (1998) proved that, under some regularity conditions on the choice of the bandwidth, the GPH estimator is consistent for the memory parameter and is asymptotically normal when the time series is Gaussian. The authors also established that the optimal $g(n)$ is of order $o(n^{4/5})$. They showed that if $g(n) \to \infty, n \to \infty$ with $\frac{g(n)}{n} \to 0$ and $\frac{g(n)}{n} \log g(n) \to 0$, then, under some conditions on $0 < f_u(\lambda_j) < \infty$, the GPH estimator is a consistent estimator of $d \in (-0.5, 0.5)$ with variance $\mathrm{var}(d_{GPH}) = \frac{\pi^2}{24g(n)} + o(g(n)^{-1})$.

A robust estimator of d

Assumption: Let $M = \min\{h', n^{\beta}\}$ with $0 < \beta < 1$, where

$$h' = \min\left\{0 < h < n : \varepsilon_n^{temp}\left(\widehat{\gamma}_Q(h)\right) \leq \frac{m}{n}\right\} - 1,$$

m and n are the numbers of outliers and the sample size, respectively.

A robust estimator of d

Let $\widetilde{I}(\lambda)$ be given by

$$\widetilde{I}(\lambda) = \frac{1}{2\pi} \sum_{s=-(n-1)}^{n-1} \kappa(s)\widetilde{R}(s)\cos(s\lambda), \qquad (8)$$

where $\widetilde{R}(s)$ is the sample autocovariance function in (5) and $\kappa(s)$ is defined as

$$\kappa(s) = \begin{cases} 1, & |s| \leq M, \\ 0, & |s| > M. \end{cases}$$

$\kappa(s)$ is called *truncated periodogram lag window* see, e.g., Priestley (1981, p. 433-437). We shall call the estimator in (8) *robust truncated pseudo-periodogram*, since it does not have the same finite-sample properties as the periodogram, with $M = n^{\beta}$, $0 < \beta < 1$.

A robust estimator of d

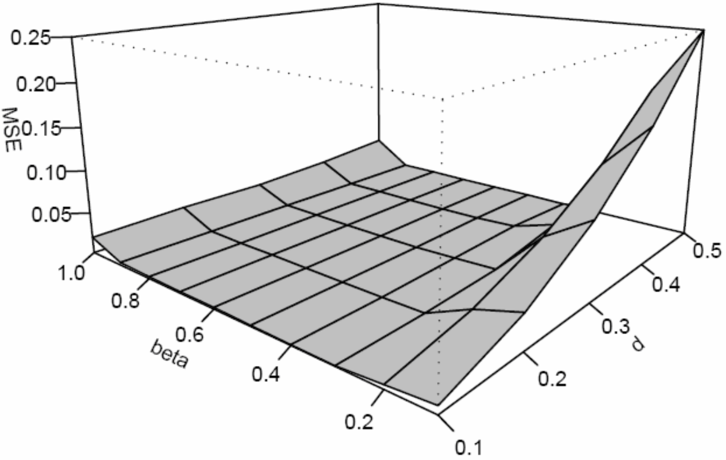
The robust GPH estimator we propose is

$$d_{GPHR} = -\frac{\sum_{i=1}^{g(n)} (x_i - \bar{x}) \log \widetilde{I}(\lambda_i)}{\sum_{i=1}^{g(n)} (x_i - \bar{x})^2}, \tag{9}$$

where $x_i = \log \left\{ 2 \sin \left(\frac{\lambda_i}{2} \right) \right\}^2$ and $g(n)$ is as before.

A robust estimator of d

The value of β, in $M = n^{\beta}$, was selected empirically by minimizing the MSE of the long-memory parameter estimates. The Figure presents simulation results for a free-outliers ARFIMA process generated with $n = 800$ and 10000 Monte Carlo experi-



Numerical results: ARFIMA$(0, d, 0)$ with $d = 0.3$

			$g(n) = n^{0.7}$		$M = n^{0.7}$	
d	n		$\hat{d}_{GPH}$	$\hat{d}_{GPH_c}$	$\hat{d}_{GPHR}$	$\hat{d}_{GPHR\,c}$
0.30	100	mean	0.2988	0.1134	0.2584	0.2449
		sd	0.1735	0.1619	0.1558	0.1556
		bias	-0.0012	-0.1866	-0.0416	-0.0551
		MSE	0.0301	0.0610	0.0260	0.0272
	300	mean	0.3062	**0.1007**	0.2907	**0.2837**
		sd	0.1005	0.0978	0.0926	0.0960
		bias	0.0062	-0.1993	-0.0093	-0.0163
		MSE	0.0101	0.0493	0.0087	0.0095
	800	mean	0.3003	0.1184	0.2949	0.2869
		sd	0.0679	0.0715	0.0573	0.0610
		bias	0.0003	-0.1816	-0.0051	-0.0131
		MSE	0.0046	0.0381	0.0033	0.0039

ω = 10, outliers = 5% (of sample)

Numerical results: ARFIMA$(0, d, 0)$ with $d = 0.45$

			$g(n) = n^{0.7}$		$M = n^{0.7}$	
d	n		$\hat{d}_{GPH}$	$\hat{d}_{GPH_c}$	$\hat{d}_{GPHR}$	$\hat{d}_{GPHR_c}$
0.45	100	mean	0.4561	0.1923	0.3975	0.3778
		sd	0.1722	0.1727	0.1506	0.1433
		bias	0.0061	-0.2577	-0.0525	-0.0722
		MSE	0.0297	0.0962	0.0254	0.0258
	300	mean	0.4594	0.2015	0.4329	0.4233
		sd	0.0986	0.0976	0.1041	0.1013
		bias	0.0094	-0.2485	-0.0171	-0.0267
		MSE	0.0098	0.0713	0.0111	0.0110
	800	mean	0.4620	**0.2306**	0.4457	**0.4349**
		sd	0.0688	0.0809	0.0562	0.0576
		bias	0.0121	-0.2194	-0.0043	-0.0151
		MSE	0.0049	0.0547	0.0032	0.0035

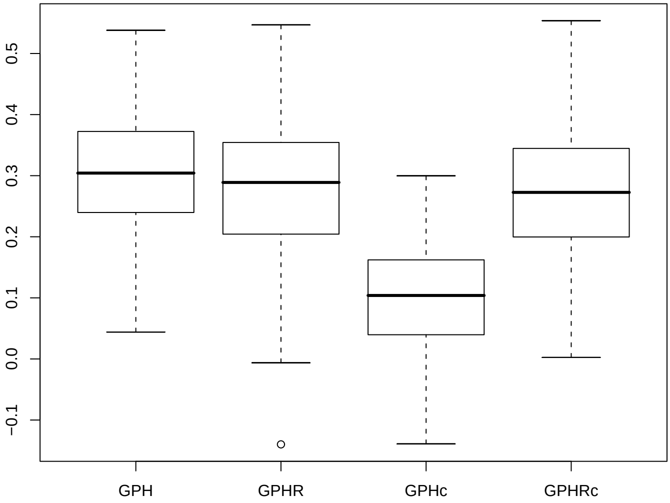
ω = 10, outliers = 5% (of sample)

Numerical results: ARFIMA$(0, d, 0)$ with $d = 0.45$

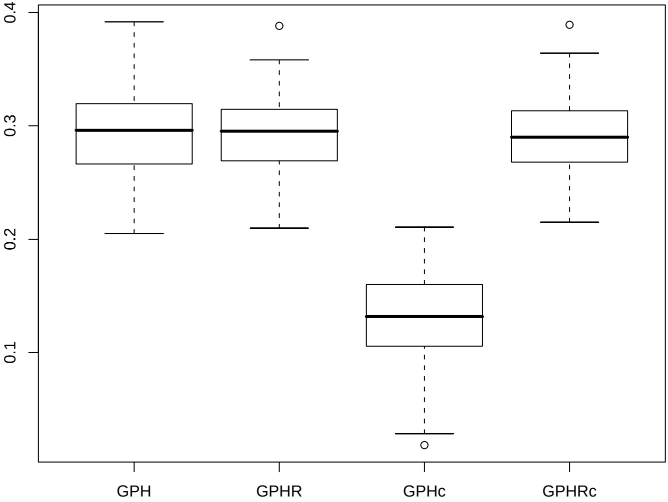
ω	n		d_{GPH_c}	d_{GPHR_c}
3	100	mean	0.3747	0.3799
		sd	0.1953	0.1513
		bias	-0.0753	-0.0701
		MSE	0.0438	0.0278
	800	mean	0.4080	0.4309
		sd	0.0679	0.0576
		bias	-0.0419	-0.0191
		MSE	0.0064	0.0037
5	100	mean	0.3108	0.3741
		sd	0.1934	0.1452
		bias	-0.1392	-0.0759
		MSE	0.0567	0.0268
	800	mean	0.3526	0.4270
		sd	0.0846	0.0568
		bias	-0.0974	-0.0229
		MSE	0.0166	0.0038
10	100	mean	0.1923	0.3778
		sd	0.1727	0.1433
		bias	-0.2577	-0.0722
		MSE	0.0962	0.0258
	800	mean	0.2306	0.4349
		sd	0.0809	0.0576
		bias	-0.2194	-0.0151
		MSE	0.0547	0.0035

outliers = 5% (of sample)

Numerical results: ARFIMA$(0, d, 0)$ $n = 300$



Numerical results: ARFIMA$(0, d, 0)$ $n = 3000$

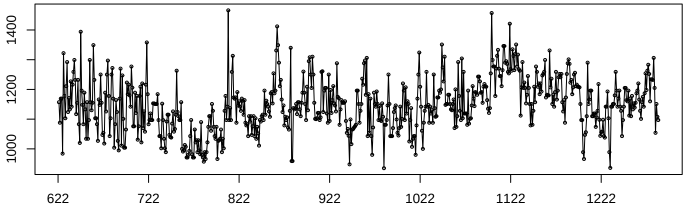


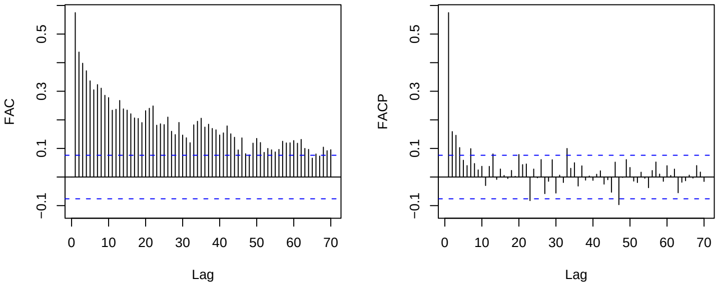
Applications: Nile river 622 - 1281 D.C.

We have applied the methodology proposed in previous section to the annual minimum water levels of the Nile river measured at the Roda Gorge near Cairo. This data set has been widely used as to illustrate long-memory memory modeling strategies; see Beran (1992), Reisen, Abraham & Toscano (2002), Robinson (1995), among others. The period analyzed ranges from 622 A.D. to 1284 A.D. (663 observations).

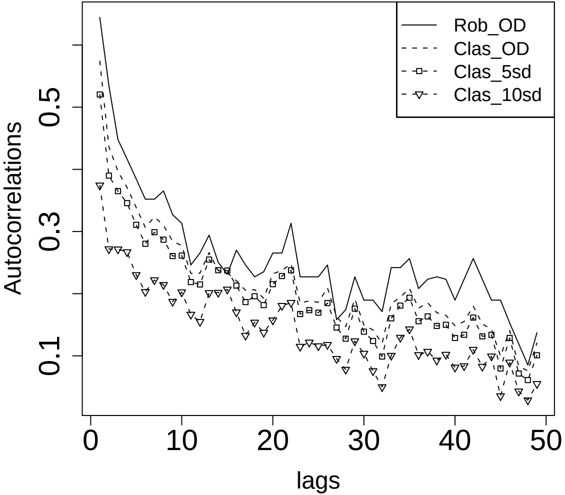
Various conclusions have been reached as to whether or not this series contains outliers. For example, Chareka, Matarise & Turner(2006) developed a test to identify outliers and ran it on the Nile data. Their test located two outliers at 646 A.D. and at 809 A.D.

Applications: Nile river 622 - 1281 D.C.





Applications: Nile river 622 - 1281 D.C.



Applications: Nile river 622 - 1281 D.C.

Bandwidth	GPH		GPHR	
	$\hat{d}$	s.e.	$\hat{d}$	s.e.
$g(n) = 25$	0.503	0.142	0.459	0.057
$g(n) = 49$	0.537	0.117	0.475	0.045
$g(n) = 94$	0.396	0.079	0.416	0.040
$g(n) = 180$	0.386	0.054	0.460	0.039

Tabela: Estimated values of d using the Nile data.

Based on a slight modification of the robust estimator proposed by Beran (1994), Agostinelli & Bisaglia (2004) found 0.412 as the estimate of d which is very close to the GPHR estimate when $g(n) = 94$.

Concluding remarks

The simulation results showed that the GPH estimator of the fractional differencing parameter can be considerably biased when the data contain atypical observations, and that the robust estimator we propose displays good finite-sample performance even when the data contain highly atypical observations. Future research should address the important issue of establishing the asymptotic properties for the proposed estimator.

THANK YOU

References

Priestley, M.
Spectral analysis and time series.
Academic Press, 1983.

Agostinelli, C. & Bisaglia, L.
Robust estimation of ARFIMA process.
Technical Report Università Cà Foscari di Venezia, 2003.

Beran, J.
On class of M-estimators for Gaussian long-memory models.
Biometrika, 81:755-766, 1994.

References

Chan, W.
A note on time series model specification in the presence outliers.
Journal of Applied Statistics, 19:117–124, 1992.

Chan, W.
Outliers and Financial Time Series Modelling: A Cautionary Note.
Mathematics and Computers in Simulation, 39:425–430, 1995.

Croux, C. & Rousseeaw P.
Time-efficient algorithms for two highly robust estimators of scale.
Computational Statistics, 1:1–18, 1992.

References

Geweke, J. & Porter-Hudak, S.
The estimation and application of long memory time series model.
Journal of time serie analysis, 1:221–238, 1983.

Haldrup, N. & Nielsen, M.
Estimation of fractional integration in the presence of data noise.
Computational statistical and data analysis, 51:3100-3114, 2007.

Hosking, J.
Fractional differencing.
Biometrika, 68:165-176, 1981.

References

Hurvich C. M. & Beltaão K.
Asymptotics for low-frequency ordinates of the periodogram of a long-memory time series.
Journal of Time Series Analysis, 14:455–472, 1993.

Hurvich C. M., Deo R. & Brodsky J.
The mean square error of Geweke and Porter-Hudak's estimator of the memory parameter of a long-memory time series.
Journal of Time Series Analysis, 19:19–46, 1998.

Künsch H. R.
Discrimination between monotonic trends and long-range dependence.
Journal of Applied Probability, 23:1025–1030, 1989.

References

Ma, Y. & Genton, M.
Highly robust estimation of the autocovariance function.
Journal of time series analysis, 21:663–684, 2000.

Rousseeaw P. & Croux, C.
Alternatives to the median absolute deviation.
Journal of the American Statistical Association, 88:1273–1283, 1993.