



HAL
open science

A dimension reduction factor approach for multivariate time series with long-memory. A robust alternative method

Valdério Anselmo Reisen, Céline Lévy-Leduc, Edson Zambon Monte, Pascal Bondon

► To cite this version:

Valdério Anselmo Reisen, Céline Lévy-Leduc, Edson Zambon Monte, Pascal Bondon. A dimension reduction factor approach for multivariate time series with long-memory. A robust alternative method. *Statistical Papers*, 2024, 65, pp.2865-2886. 10.1007/s00362-023-01504-2 . hal-04225878

HAL Id: hal-04225878

<https://agroparistech.hal.science/hal-04225878>

Submitted on 3 Oct 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A dimension reduction factor approach for multivariate time series with long-memory. A robust alternative method

Valdério Anselmo Reisen^{1,3,4}, Céline Lévy-Leduc², Edson Zambon Monte³,
and Pascal Bondon⁴

¹PPGEA, Federal University of Espírito Santo, Brazil

²Université Paris-Saclay, AgroParisTech, INRAE, UMR MIA Paris-Saclay, 91120, Palaiseau, France

³Department of Economics, PPGEco, Federal University of Espírito Santo, Brazil

⁴Université Paris-Saclay, CNRS, CentraleSupélec, Laboratoire des signaux et systèmes, 91190, Gif-sur-Yvette, France

October 3, 2023

Abstract

This paper studies factor modeling for a vector of time series with long-memory properties to investigate how outliers affect the identification of the number of factors and also proposes a robust method to reduce their impact. The number of factors is estimated using an eigenvalue analysis for a non-negative definite matrix introduced by Lam et al (2011). Two estimators are proposed; the first is based on the classical sample covariance function, and the second uses a robust covariance function estimate. In both cases, it is shown that the eigenvalues estimates have similar convergence rates. Empirical simulations support both estimators for multivariate stationary long-memory time series and show that the robust method is preferable when the data is contaminated with additive outliers. Time series of daily log returns are used as an example of application. In addition to abrupt observations, exchange rates exhibit non-stationarity properties with long memory parameters greater than one. Then we use semi-parametric long memory estimators to estimate the fractional parameters of the series. The number of factors was estimated using the classical and robust approaches. Due to the influence of the abrupt observations, these tools suggested a different number of factors to model the data. The robust method suggested two factors, while the classical approach indicated only one factor.

Keywords. Eigenvalues, factor analysis, long-memory, reduced rank, robustness.

MSC2020 subject classifications. 62M10, 62H25, 62H12

1 Introduction

Multivariate statistical techniques are standard tools for analyzing high-dimensional data. In this context, factor analysis (FA) is one of the most widely used methods to explain a considerable number k of series by a small number r of unobservable factors, to improve forecasting, and to capture cross-sectional dependence, among others. A standard assumption in FA is that the observations are not time-dependent, which is unrealistic in most application areas. The problem of simplifying multivariate time series structures dates back to the 1980s when computational advances were minimal to deal with large time-dependent data sets. For example, Geweke (1977) and Geweke and Singleton (1981) have adapted FA in the frequency-domain, see also (Priestley et al, 1974; Brillinger, 1981). Velu et al (1986) have discussed the asymptotic distribution of reduced rank matrix estimators in the multivariate auto-regressive model. Other data reduction methods were proposed in the 1980s see, for example, a review in (Fernández-Macho, 1997) and for more recent developments, see (Christou, 2020). Several recent surveys have been dedicated to dynamic factors models, including stationary and non-stationary times series processes in different areas of applications such as environmental, health, and financial sciences see, for example, (Eichler et al, 2011; Lam et al, 2011; Toman, 2014; Bai and Wang, 2016; Chen et al, 2020; Fan et al, 2021; Lin et al, 2022; Bai and Zheng, 2023) to mention a few. Peña and Box (1987) have proposed a simple model to identify hidden factors in multivariate short-memory processes. This model is widely used to deal with the identification of the number of factors, estimation, and dimension-reduction of high-dimensional time series. Lam et al (2011) and Lam and Yao (2012) have discussed estimation of the latent factors and a dimension-reduction approach for short-memory processes. They have introduced a nonnegative definite matrix that accommodates information on the covariances of the process for different lags. As an alternative method for estimating the number of factors suggested by Lam et al (2011), Reisen et al (2019) have proposed an approach for short-memory processes which is robust against additive outliers and heavy tail distributions. These authors have established asymptotic properties of the estimator of the number of factors and have discussed the method's usefulness in air pollution data. Bai and Zheng (2023) discussed the robust method introduced by Reisen et al (2019). The authors also suggested an algorithm for constructing bootstrap prediction intervals for the high-dimensional time series.

Motivated by the importance of dimension reduction in large data sets, this paper generalizes the recent works on FA by allowing the process to have long-memory properties and additive outliers, phenomena quite common in many areas of applications. The number of factors is estimated in terms of an eigenvalue analysis for the nonnegative definite matrix introduced by Lam et al (2011). Two estimators are proposed; one is based on the classical sample covariance function, and the other on a robust covariance function estimate. It is proved that the convergence rates of the eigenvalues are the same in the classical and robust approach. Since outliers are common in many areas, the methods are empirically investigated in a long-memory framework with and without outliers. Finite sample performances support the use of the approaches in practical problems. The robust

method is strongly recommended when the data is contaminated with additive outliers. Similarly to Lam and Yao (2012), the estimates of the number of factors improve as the dimension k increases, a phenomenon coined as the “blessing of dimensionality”. To show the usefulness of the proposed methodology in the dimension reduction problem, this paper analyses the daily returns of the six following foreign exchange rates against U.S. Dollar. The series display high-level observations and exhibits the non-stationary phenomenon of long memory. These features found in the data set are also important in the context of the methodology proposed here since, so far, the application in dimension reduction factor in the literature has not considered time series with the non-stationary long memory case and additive outliers. This paper is organized as follows. Section 2 discusses the model and the estimators of the rank-reduction and their asymptotic properties. Section 3 displays a simulation study to verify the finite sample performance of the methods under the scenarios of contaminated and non-contaminated multivariate time series. The real data application is presented in Section 4. A conclusion is given in Section 5 and Section 6 contains the proofs.

2 The dynamic factor model with long-memory

2.1 Model

Let $\mathbf{Y}_t = (Y_{1,t}, Y_{2,t}, \dots, Y_{k,t})'$, $t \in \mathbb{Z}$, be a zero-mean k -dimensional second-order stationary time series generated by

$$\mathbf{Y}_t = \mathbf{P}\mathbf{X}_t + \boldsymbol{\varepsilon}_t, \quad (1)$$

where \mathbf{X}_t is an unobserved zero-mean r -dimensional stationary time series of common factors ($r \leq k$), \mathbf{P} is a $k \times r$ matrix of parameters of rank r , and $\boldsymbol{\varepsilon}_t$ is a k -dimensional white-noise sequence with full-rank covariance matrix $\boldsymbol{\Sigma}_\varepsilon$. Thus, all the dynamic structure comes through the common factors \mathbf{X}_t . Model (1) was introduced by Peña and Box (1987) and it has been applied to many economic data see, for example, the review in (Lam and Yao, 2012). This model is relevant to reduce the dimensionality when $r < k$, i.e., a large number k of series can be explained by a small number r of factors. We assume that the following assumptions hold :

(A1) \mathbf{X}_t and $\boldsymbol{\varepsilon}_s$ are uncorrelated for any $(t, s) \in \mathbb{Z}^2$, and $\mathbf{P}'\mathbf{P} = \mathbf{I}_r$, where \mathbf{I}_r denotes the $r \times r$ identity matrix.

(A2) \mathbf{Y}_t is a Gaussian process satisfying

$$\gamma_{ij}(h) \sim C_{ij}h^{-(D_i+D_j)/2}, \text{ as } h \rightarrow +\infty,$$

for all $i, j \in \{1, \dots, k\}$, where $\gamma_{ij}(h) = \text{Cov}(Y_{i,t}, Y_{j,t+h})$, $C_{ij} \in \mathbb{R}$, $(D_i, D_j) \in (0, 1)^2$.

In (A1), condition $\mathbf{P}'\mathbf{P} = \mathbf{I}_r$ ensures identifiability in (1), see (Peña and Box, 1987) and (Lam and Yao, 2012) for further details. Assumption (A2) means that \mathbf{Y}_t is a stationary long-range dependent process see, for example, (Roouch et al, 2019). It follows from (1) and (A1) that

$$\Gamma_Y(h) = \mathbb{E}[\mathbf{Y}_t \mathbf{Y}'_{t+h}] = \begin{cases} \mathbf{P}\Gamma_X(0)\mathbf{P}' + \Sigma_\varepsilon & \text{when } h = 0, \\ \mathbf{P}\Gamma_X(h)\mathbf{P}' & \text{when } h \neq 0. \end{cases} \quad (2)$$

A parametric class of processes satisfying (A2) is the stationary Vector Autoregressive Fractionally Integrated Moving Average (VARFIMA) process with orders (p, d_1, \dots, d_k, q) in which $D_i = 1 - 2d_i$ see, for example, (Chung, 2002, Lemma 2). This model is defined by the difference equations

$$\phi(B)\mathbf{Z}_t = \theta(B)\boldsymbol{\eta}_t, \quad (3)$$

$$Z_{i,t} = (1 - B)^{d_i} Y_{i,t} \quad (4)$$

for $i = 1, \dots, k$, where $\boldsymbol{\eta}_t$ is a Gaussian zero-mean multivariate white noise process, B is the backward operator, *i.e.*, $BX_t = X_{t-1}$ for any process X_t . For any $d \in \mathbb{R}$, the time series $(1 - B)^d X_t$ is defined by

$$(1 - B)^d X_t = \sum_{k=0}^{\infty} b_k X_{t-k},$$

where

$$b_k = \prod_{j=1}^k \frac{j - 1 - d}{j} = \frac{\Gamma(k - d)}{\Gamma(k + 1)\Gamma(-d)}$$

are the coefficients in the Taylor series for $(1 - z)^d$ when $|z| < 1$ and $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$ is the Gamma function. It is assumed that p and q are non negative integers and $0 < d_i < 1/2$ for $i = 1, \dots, k$. The matrix-valued polynomials $\phi(\cdot)$ and $\theta(\cdot)$ satisfy that $\det(\phi(z)) \neq 0$ and $\det(\theta(z)) \neq 0$ for all $z \in \mathbb{C}$ such that $|z| \leq 1$. These two conditions are known as the causality and invertibility properties, respectively. Additional conditions have to be imposed to obtain an identifiable model, see e.g. (Brockwell and Davis, 2009, page 431) and (Reinsel, 2003, section 2.3). When all d_i 's are zero, \mathbf{Y}_t reduces to a VARMA process and has a short-memory correlation structure in the sense that the sequence of matrices $\Gamma_Y(h)$ for $h \in \mathbb{Z}$ is summable. Otherwise, \mathbf{Y}_t satisfies (A2) and has a long-memory behaviour, see (Chung, 2002).

One of the main tasks in factor analysis is to estimate the number of factors r in (1). Under Assumption (A1), for any prescribed positive integer h_0 , r is the number of nonzero eigenvalues of

$$\mathbf{M} = \sum_{h=1}^{h_0} \Gamma_Y(h)\Gamma_Y(h)', \quad (5)$$

and the r -dimensional linear space spanned by the columns of \mathbf{P} is spanned by the eigenvectors of \mathbf{M} with nonzero eigenvalues, see (Lam and Yao, 2012). Observe that \mathbf{M} is a symmetric and non-negative definite matrix regardless of the property of $\mathbf{\Gamma}_Y(h)$. Moreover, the use of $\mathbf{\Gamma}_Y(h)\mathbf{\Gamma}_Y(h)'$ instead of $\mathbf{\Gamma}_Y(h)'$ in the definition (5) of \mathbf{M} avoids the cancellation of the autocovariance from different lags and allows the use of an estimator of $\mathbf{\Gamma}_Y(h)$ that is not necessarily non-negative definite. This is the case of the estimator proposed here. As remarked by Lam and Yao (2012), the use of (5) in factorial analysis can be very helpful especially when the sample size is small. Lam et al (2011) reported that the estimation of the linear space spanned by the columns of \mathbf{P} and r is not sensitive to the choice of h_0 . This issue is also investigated here through simulations.

Performing an eigenvalue analysis for \mathbf{M} is easier than using the asymmetric matrix $\mathbf{\Gamma}_Y(h)$. Therefore, (5) becomes very useful in practical problems where $\mathbf{\Gamma}_Y(h)$ is replaced by an estimate $\widehat{\mathbf{\Gamma}}_Y(h)$ to compute

$$\widehat{\mathbf{M}} = \sum_{h=1}^{h_0} \widehat{\mathbf{\Gamma}}_Y(h)\widehat{\mathbf{\Gamma}}_Y(h)' \quad (6)$$

and to estimate the number of factors r by

$$\widehat{r} = \operatorname{argmin}_{1 \leq i \leq R} \widehat{\lambda}_{i+1}/\widehat{\lambda}_i, \quad (7)$$

where $r < R < k$ is a constant and $\widehat{\lambda}_1 \geq \dots \geq \widehat{\lambda}_k$ are the eigenvalues of $\widehat{\mathbf{M}}$, see (Lam et al, 2011). Lam and Yao (2012) derived the asymptotic properties of the $\widehat{\lambda}_i$'s under some assumptions and gave some practical recommendations for selecting R . Based on the fact that atypical observations (additive outliers) are quite common in time series, Reisen et al (2019) proposed to estimate r by (7) for short-range dependent processes where the $\widehat{\lambda}_i$'s are the eigenvalues of an estimate of \mathbf{M} where $\mathbf{\Gamma}_Y(h)$ is replaced by the robust sample ACF function introduced by Ma and Genton (2000), which is the robust ACF approach considered here and discussed in the next section. The theoretical properties of this robust ACF estimator have been studied by Lévy-Leduc et al (2011a) and Lévy-Leduc et al (2011b).

2.2 Estimation

Here, we propose estimates of r based on the eigenvalues computed from two estimates of \mathbf{M} for long-memory processes. The first estimate is based on the eigenvalues obtained from the sample autocovariance matrix $\widehat{\mathbf{\Gamma}}_Y(h)$, and the second estimate is built from the eigenvalues computed from the sample autocovariance matrix $\widehat{\mathbf{\Gamma}}_Y^Q(h)$ based on the scale estimator $Q_n(\cdot)$.

For a given sample $\mathbf{Y}_{1:n} = (\mathbf{Y}_1, \dots, \mathbf{Y}_n)$ of \mathbf{Y}_t , let

$$\widehat{\gamma}_{i,j}(h) = \frac{1}{n} \sum_{t=1}^{n-h} (Y_{i,t} - \bar{Y}_i)(Y_{j,t+h} - \bar{Y}_j), \quad (8)$$

where $\bar{Y}_i = \sum_{t=1}^n Y_{i,t}/n$ and $\bar{Y}_j = \sum_{t=1}^n Y_{j,t}/n$.

Proposition 1. Assume that (A2) holds. Let h be a fixed positive integer and $\widehat{\mathbf{\Gamma}}_Y(h) = (\widehat{\gamma}_{i,j}(h))_{1 \leq i,j \leq k}$, where $\widehat{\gamma}_{i,j}(h)$ is defined in (8). Then,

(i) If, for all i in $\{1, \dots, k\}$, $D_i > 1/2$,

$$\sqrt{n} \sup_{1 \leq j \leq k} \left| \widehat{\lambda}_j - \lambda_j \right| = O_p(1), \text{ as } n \rightarrow \infty,$$

(ii) If, there exists i_0 in $\{1, \dots, k\}$ such that $D_{i_0} < 1/2$,

$$n^{D_{i_0}} \sup_{1 \leq j \leq k} \left| \widehat{\lambda}_j - \lambda_j \right| = O_p(1), \text{ as } n \rightarrow \infty,$$

where $(\widehat{\lambda}_j)_{1 \leq j \leq k}$ and $(\lambda_j)_{1 \leq j \leq k}$ denote respectively the eigenvalues of $\left(\sum_{h=1}^{h_0} \widehat{\mathbf{\Gamma}}_Y(h) \widehat{\mathbf{\Gamma}}_Y(h)' \right)$ and $\left(\sum_{h=1}^{h_0} \mathbf{\Gamma}_Y(h) \mathbf{\Gamma}_Y(h)' \right)$, $\mathbf{\Gamma}_Y(h) = (\gamma_{i,j}(h))_{1 \leq i,j \leq k}$, and h_0 is a fixed positive integer.

Proof of Proposition 1. The proof follows directly from Lemmas 1, 2 and 4 in the Appendix. \square

Given the set of observation X_1, \dots, X_n from a Gaussian univariate stationary process X_t , the $Q_n(\cdot)$ estimator of the standard deviation of X_1 proposed by Rousseeuw and Croux (1993) is the τ th order statistic defined by

$$Q_n(X) = c \{ |X_i - X_j|; i < j \}_{(\tau)}, \quad i, j = 1, \dots, n, \quad (9)$$

where $c = 2.2191$ is a constant, $\tau = \lfloor ((\binom{n}{2} + 2)/4) \rfloor + 1$ and $\lfloor x \rfloor$ is the largest integer smaller than x . Following Ma and Genton (2000), for a sample $(\mathbf{Y}_1, \dots, \mathbf{Y}_n)$, the alternative autocovariance estimator of $\gamma_{i,j}(h) = \text{Cov}(Y_{i,t}, Y_{j,t+h})$ for all $i, j = 1, \dots, k$, is given by

$$\widehat{\gamma}_{i,j}^Q(h) = \frac{1}{4} \left[Q_{n-h}^2(Y_{i,1:n-h} + Y_{j,h+1:n}) - Q_{n-h}^2(Y_{i,1:n-h} - Y_{j,h+1:n}) \right], \quad (10)$$

where $Y_{i,1:n-h} = (Y_{i,1}, \dots, Y_{i,n-h})$ and $Y_{j,h+1:n} = (Y_{j,h+1}, \dots, Y_{j,n})$. Let $\widehat{\mathbf{\Gamma}}_Y^Q(h)$ be the matrix with entries $\widehat{\gamma}_{i,j}^Q(h)$. Note that, $\widehat{\gamma}_{i,j}^Q(h)$, for each $h \geq 0$, does not necessarily have the non-negative definite property, nevertheless, $\widehat{\mathbf{M}}^Q$ defined by

$$\widehat{\mathbf{M}}^Q = \sum_{h=1}^{h_0} \widehat{\mathbf{\Gamma}}_Y^Q(h) \widehat{\mathbf{\Gamma}}_Y^Q(h)' \quad (11)$$

has this property and the estimator \widehat{r}^Q of r is obtained from (7) where the $\widehat{\lambda}_i$'s are replaced by the eigenvalues $\widehat{\lambda}_i^Q$'s of $\widehat{\mathbf{M}}^Q$.

Proposition 2. Let h be a fixed positive integer and $\widehat{\Gamma}_Y^Q(h) = \left(\widehat{\gamma}_{i,j}^Q(h)\right)_{1 \leq i,j \leq k}$, where $\widehat{\gamma}_{i,j}^Q(h)$ is defined in (10). Assume that (A2) holds, then

(i) If, for all i in $\{1, \dots, k\}$, $D_i > 1/2$,

$$\sqrt{n} \sup_{1 \leq j \leq k} \left| \widehat{\lambda}_j^Q - \lambda_j \right| = O_p(1), \text{ as } n \rightarrow \infty,$$

(ii) If, there exists i_0 in $\{1, \dots, k\}$ such that $D_{i_0} < 1/2$,

$$n^{D_{i_0}} \sup_{1 \leq j \leq k} \left| \widehat{\lambda}_j^Q - \lambda_j \right| = O_p(1), \text{ as } n \rightarrow \infty,$$

where $(\widehat{\lambda}_j^Q)_{1 \leq j \leq k}$ and $(\lambda_j)_{1 \leq j \leq k}$ denote respectively the eigenvalues of $\left(\sum_{h=1}^{h_0} \widehat{\Gamma}_Y^Q(h) \widehat{\Gamma}_Y^Q(h)'\right)$ and $\left(\sum_{h=1}^{h_0} \Gamma_Y(h) \Gamma_Y(h)'\right)$, $\Gamma_Y(h) = (\gamma_{i,j}(h))_{1 \leq i,j \leq k}$, and h_0 is a fixed positive integer.

Proof of Proposition 2. The proof follows directly from Lemmas 1, 2 and 3 in the Appendix. \square

Proposition 3. Assume that (A2) holds. Let $(\widehat{\lambda}_j)_{1 \leq j \leq k}$, $(\widehat{\lambda}_j^Q)_{1 \leq j \leq k}$ and denote respectively the eigenvalues of $\left(\sum_{h=1}^{h_0} \widehat{\Gamma}_Y(h) \widehat{\Gamma}_Y(h)'\right)$, $\left(\sum_{h=1}^{h_0} \widehat{\Gamma}_Y^Q(h) \widehat{\Gamma}_Y^Q(h)'\right)$ and $\left(\sum_{h=1}^{h_0} \Gamma_Y(h) \Gamma_Y(h)'\right)$, where h_0 is a fixed positive integer and $\widehat{\Gamma}_Y(h)$, $\widehat{\Gamma}_Y^Q(h)$ and $\Gamma_Y(h)$ are defined in Propositions 1 and 2. Let \widehat{r} and \widehat{r}^Q be defined by

$$\widehat{r} = \operatorname{argmin}_{1 \leq i \leq R} \widehat{\lambda}_{i+1} / \widehat{\lambda}_i \quad \text{and} \quad \widehat{r}^Q = \operatorname{argmin}_{1 \leq i \leq R} \widehat{\lambda}_{i+1}^Q / \widehat{\lambda}_i^Q,$$

where $r < R < k$ is a constant. Then, \widehat{r} and \widehat{r}^Q are consistent estimators of r .

Proof of Proposition 3. By Propositions 1 and 2, $(\widehat{\lambda}_i)_{1 \leq i \leq k}$ and $(\widehat{\lambda}_i^Q)_{1 \leq i \leq k}$ are consistent estimators of $(\lambda_i)_{1 \leq i \leq k}$. Thus, the ratios $(\widehat{\lambda}_{i+1} / \widehat{\lambda}_i)_{1 \leq i < k}$ and $(\widehat{\lambda}_{i+1}^Q / \widehat{\lambda}_i^Q)_{1 \leq i < k}$ are consistent estimators of $(\lambda_{i+1} / \lambda_i)_{1 \leq i < k}$, with the convention $0/0 = 0$. Since, with this convention,

$$r = \operatorname{argmin}_{1 \leq i \leq R} \lambda_{i+1} / \lambda_i,$$

the conclusion follows by using the uniform convergence proved in Proposition 1. \square

Propositions 1 and 2 show that the eigenvalues of the two estimators of matrix \mathbf{M} have the same convergence rates governed by the fractional parameter D_i . Proposition 3 shows that the two estimators of r are consistent. Note that, as is well-known, $\widehat{\Gamma}_Y(h)$ is sensitive to outliers whereas, in general, $\widehat{\Gamma}_Y^Q(h)$ is almost unaffected. This feature suggests that the estimator \widehat{r}^Q is a choice to estimate r for long-memory time series with additive outliers. This issue is addressed in the simulation study.

3 Monte Carlo simulations

This section discusses the empirical performances of \hat{r} and \hat{r}^Q under the scenarios of uncontaminated and contaminated long-memory time series with additive outliers. We set $r = 3$ in (1) and we generate factors \mathbf{X}_t , $t = 1, \dots, n$, from VARFIMA model with order $(1, d_1, d_2, d_3, 0)$, namely

$$\begin{aligned}\phi(B)\mathbf{X}_t &= \boldsymbol{\eta}_t, \\ X_{i,t} &= (1 - B)^{d_i} X_{i,t}, \quad i = 1, \dots, 3,\end{aligned}$$

where the coefficient matrix ϕ is diagonal with 0.6, -0.5, and 0.3 as the main diagonal elements, $\boldsymbol{\eta}_t$ are independent zero-mean Gaussian vectors with identity covariance matrix, and $(d_1, d_2, d_3) = (0.1, 0.2, 0.1), (0.3, 0.2, 0.4)$ which corresponds to low and strong positive correlations, respectively. These choices are intentional in the sense of considering series with parameter vectors satisfying the two regions of the asymptotic properties of the estimators discussed in Proposition 1 and 2. The sample sizes are $n = 100, 200, 400, 800, 1600$, $k = 0.2n, 0.5n, 0.8n, 1.2n$, and $h_0 = 1, 3$. The elements of \mathbf{P} in (1) are realizations of independent random variables with the uniform distribution on $[-1, 1]$ and the random variables $\boldsymbol{\varepsilon}_t$ are independent zero-mean Gaussian vectors with identity covariance matrix. A similar simulation process for \mathbf{P} and $\boldsymbol{\varepsilon}_t$ was considered by Lam and Yao (2012). The Monte Carlo study is based on 1000 replications. The contaminated data is generated as in (Reisen et al, 2019). Only the first component $X_{1,t}$ of \mathbf{X}_t is contaminated by additive outliers with the probability of occurrence $p = 0.05$ and magnitude $\omega = 15$.

The frequencies of occurrence of $\hat{r} = 3$ and $\hat{r}^Q = 3$ over the 1000 replications are reported in Tables 1 and 2 in the uncontaminated long-memory time series, and Table 3 displays the contaminated and uncontaminated cases. In these tables, $h_0 = 1$ and $(d_1, d_2, d_3) = (0.3, 0.2, 0.4)$. Similar results are obtained when $h_0 = 3$ and $(d_1, d_2, d_3) = (0.1, 0.2, 0.1)$. The results in Table 1 show that \hat{r} performs better as n and k increase. Also, for each fixed n , \hat{r} performs better as k increases (“blessing of dimensionality”). However, for large n ($n \geq 800$), the estimator \hat{r} reaches 100% accuracy.

These results are similar to the ones in (Lam and Yao, 2012, Table 1). Table 2 leads to similar conclusions, although \hat{r}^Q slightly underperforms \hat{r} for $n \leq 400$. Similar to the performance of \hat{r} , \hat{r}^Q reaches 100% accuracy for $n \geq 800$. Since financial data generally have many large number of observations, both estimators can be used to identify the number of factors.

The picture of the performance of \hat{r} changes dramatically in the context of multivariate time series with additive outliers. The superiority of \hat{r}^Q over \hat{r} appears in Table 3 where \hat{r}^Q shows to be very resistant against additive outliers whereas \hat{r} is totally affected, as was expected. Table 3 also displays the results for $p = 0$ (uncontaminated data) to show how the frequencies of occurrences are distributed across \hat{r} -values. In order to verify the effect of the percentage and magnitude of outliers on the estimates of the number of factors r , additional simulations were carried out with different values of p and ω . Tables 4 and 5 display the performance of \hat{r} and \hat{r}^Q when $p = 0.01, 0.10$, $\omega = 15$ and $p = 0.05$ and $\omega = 10$,

respectively. Other cases are available upon request. The estimates displayed in these tables lead to similar conclusions as in the previous cases, *i.e.*, \hat{r} is much more affected by the outliers than \hat{r}^Q .

The number of nonzero eigenvalues r of \mathbf{M} is defined by (5) and this is independent of the choice of $h_0 > 0$. However, in practice when \mathbf{M} is estimated by $\widehat{\mathbf{M}}$ or $\widehat{\mathbf{M}}^Q$, it is recommended to choose a small value of h_0 since the strongest correlations are, in general, at small values of the time lag h , and for large h , the estimate of $\Gamma_Y(h)$ is not very accurate see, also, (Lam and Yao, 2012). In Tables 6 and 7 we report the relative frequencies for $\hat{r} = 3$ and $\hat{r}^Q = 3$ when $h_0 = 1$ and $h_0 = 3$ for $n = 100, 200, 400$, $k = 0.8n$ and $(d_1, d_2, d_3) = (0.3, 0.2, 0.4), (0.1, 0.2, 0.1)$, respectively.

From both tables, we observe that for $n = 400$, quite a standard sample size in real problems, the choice of h_0 does not greatly affect the estimates. On the other hand, we can see two interesting features when $n = 100, 200$. In all cases, $h_0 = 3$ suggests that the relative frequency underestimates the true value more significantly than for $h_0 = 1$. This may be justified by fact the estimates of $\Gamma_Y(h)$ are not very much accurate in the long-memory series for not large n . In this scenario, the magnitude of the fractional parameters also influences the estimation. In the case of Table 7, the empirical rates are higher than the ones from Table 6, which is an expected result according to the rates of convergence of the estimators given in Proposition 1 and 2. Thus, this empirical investigation corroborates the use of $h_0 = 1$ in the estimation of the numbers of factors and, for large n , both estimation methods \hat{r}^Q and \hat{r} perform similarly in the long-memory time series, independently of the magnitude of the long-memory parameters. The scenario changes dramatically when additive outliers are suspected in the data. In this situation, the robust approach proposed here is strongly recommended.

To complete this empirical investigation, Figures 1 and 2 show the box-plots of the ratios $\hat{\lambda}_{j+1}/\hat{\lambda}_j$ and $\hat{\lambda}_{j+1}^Q/\hat{\lambda}_j^Q$ when $h_0 = 1$, $n = 100, 400$, $k = 0.8n$, $(d_1, d_2, d_3) = (0.3, 0.2, 0.4)$. Similar to (Lam and Yao, 2012) and (Reisen et al, 2019)'s findings for short-memory processes, the ratios are also close to 1 when $j > r$ in the case of long-memory time series. As expected from Proposition 3, the estimates becomes more accurate as n increases, that is, the consistency property of the estimators.

To clarify the effect of additive outliers on the sample ACF, we generated an ARFIMA model with $d = 0.3$ and $n = 200$ without and with outliers ($p = 0.05$, $\omega = 15$). The sample ACFs of the uncontaminated and contaminated data are in Figures 3 and 4, respectively. Figure 3 shows that in the absence of outliers, the robust ACF presents similar behavior as the classical one. However, the performance of the classical estimator changes significantly when outliers are introduced in the series (Figure 4). This estimator lacks its robustness, while the robust one displays similar behavior to the uncontaminated case.

Finally, this paper's results discussed theoretically and empirically contribute to using the Q_n scale estimator in the dimension reduction factor approach for high-dimensional time series with long memory, with and without outliers. More generally, the robust estimation of eigenvalues based on the Q_n estimator can also be applied in other multivariate techniques in the context of long-memory time series with and without outliers.

Table 1: Relative frequencies for $\hat{r} = 3$ when $h_0 = 1$, $(d_1, d_2, d_3) = (0.3, 0.2, 0.4)$ (uncontaminated long-memory time series)

n	100	200	400	800	1600
$k = 0.2n$	0.507	0.796	0.967	1	1
$k = 0.5n$	0.720	0.904	0.991	1	1
$k = 0.8n$	0.768	0.927	0.984	1	1
$k = 1.2n$	0.822	0.940	0.995	1	1

Table 2: Relative frequencies for $\hat{r}^Q = 3$ when $h_0 = 1$, $(d_1, d_2, d_3) = (0.3, 0.2, 0.4)$ (uncontaminated long-memory time series).

n	100	200	400	800	1600
$k = 0.2n$	0.351	0.711	0.900	1	1
$k = 0.5n$	0.554	0.874	0.959	1	1
$k = 0.8n$	0.649	0.898	0.966	1	1
$k = 1.2n$	0.727	0.918	0.975	1	1

Table 3: Relative frequencies for \hat{r} and \hat{r}^Q when $n = 200$, $h_0 = 1$, $(d_1, d_2, d_3) = (0.3, 0.2, 0.4)$ (contaminated and uncontaminated long-memory time series).

	$p = 0$			$p = 0.05$ and $\omega = 15$			$p = 0$			$p = 0.05$ and $\omega = 15$		
	$\hat{r} = 1$	$\hat{r} = 2$	$\hat{r} = 3$	$\hat{r} = 1$	$\hat{r} = 2$	$\hat{r} = 3$	$\hat{r}^Q = 1$	$\hat{r}^Q = 2$	$\hat{r}^Q = 3$	$\hat{r}^Q = 1$	$\hat{r}^Q = 2$	$\hat{r}^Q = 3$
$k = 0.2n$	0.028	0.176	0.796	0.151	0.312	0.502	0.040	0.249	0.711	0.054	0.271	0.675
$k = 0.5n$	0.003	0.093	0.904	0.085	0.250	0.600	0.005	0.121	0.874	0.013	0.143	0.844
$k = 0.8n$	0.001	0.072	0.927	0.056	0.234	0.614	0.001	0.072	0.898	0.003	0.127	0.870
$k = 1.2n$	0.001	0.059	0.940	0.056	0.226	0.588	0.002	0.080	0.918	0.003	0.118	0.879

Table 4: Relative frequencies for \hat{r} and \hat{r}^Q when $n = 200$, $h_0 = 1$, $(d_1, d_2, d_3) = (0.3, 0.2, 0.4)$, $p = 0.01$, $p = 0.10$ and $\omega = 15$ (contaminated long-memory time series).

	$p = 0.01$ and 15						$p = 0.10$ and $\omega = 15$					
	$\hat{r} = 1$	$\hat{r} = 2$	$\hat{r} = 3$	$\hat{r}^Q = 1$	$\hat{r}^Q = 2$	$\hat{r}^Q = 3$	$\hat{r} = 1$	$\hat{r} = 2$	$\hat{r} = 3$	$\hat{r}^Q = 1$	$\hat{r}^Q = 2$	$\hat{r}^Q = 3$
$k = 0.2n$	0.008	0.212	0.776	0.016	0.272	0.732	0.246	0.336	0.368	0.068	0.344	0.588
$k = 0.5n$	0.016	0.140	0.840	0.008	0.092	0.900	0.232	0.292	0.412	0.024	0.204	0.780
$k = 0.8n$	0.008	0.080	0.900	0.008	0.090	0.902	0.176	0.196	0.512	0.016	0.184	0.805
$k = 1.2n$	0.004	0.076	0.915	0.000	0.080	0.925	0.148	0.124	0.494	0.010	0.160	0.830

Table 5: Relative frequencies for \hat{r} and \hat{r}^Q when $k = 0.8n$, $n = 200$, $h_0 = 1$, $(d_1, d_2, d_3) = (0.3, 0.2, 0.4)$, $p = 0.05$ and $\omega = 10$ (contaminated long-memory time series).

	$p = 0.05$ and $\omega = 10$					
	$\hat{r} = 1$	$\hat{r} = 2$	$\hat{r} = 3$	$\hat{r}^Q = 1$	$\hat{r}^Q = 2$	$\hat{r}^Q = 3$
$k = 0.8n$	0.024	0.110	0.840	0.000	0.100	0.900

Remark 1. The effect of the percentage and magnitude of outliers on the autocovariance and autocorrelation functions, among other statistical functions of time series, is discussed

Table 6: Relative frequency estimates for dimensional reduction for $k = 0.8n$, $h_0 = 1$ and $h_0 = 3$, $d_1 = 0.3, d_2 = 0.2, d_3 = 0.4$

n	$h_0 = 1$		$h_0 = 3$	
	$\hat{r} = 3$	$\hat{r}^Q = 3$	$\hat{r} = 3$	$\hat{r}^Q = 3$
100	0.768	0.649	0.435	0.370
200	0.927	0.898	0.630	0.555
400	0.984	0.979	0.800	0.750

Table 7: Relative frequency estimates for dimensional reduction for $k = 0.8n$, $h_0 = 1$ and $h_0 = 3$, considering $d_1 = 0.1, d_2 = 0.2, d_3 = 0.1$

n	$h_0 = 1$		$h_0 = 3$	
	$\hat{r} = 3$	$\hat{r}^Q = 3$	$\hat{r} = 3$	$\hat{r}^Q = 3$
100	0.790	0.740	0.500	0.450
200	0.985	0.960	0.760	0.695
400	1	1	0.960	0.930

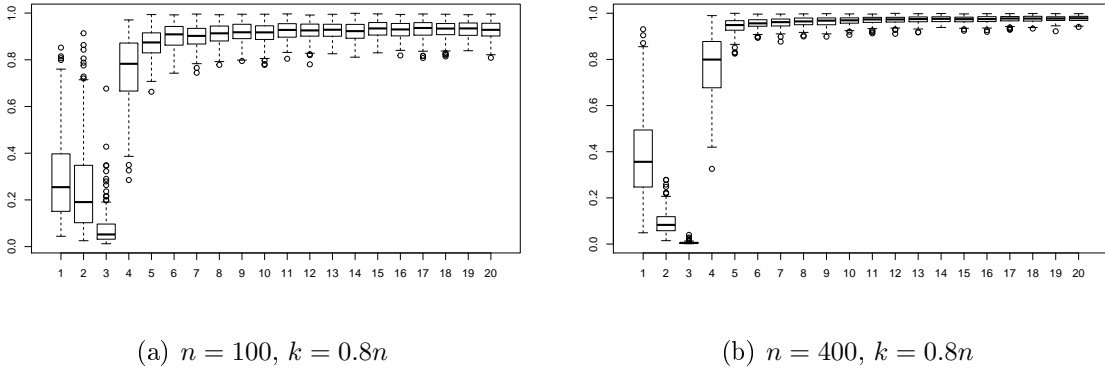


Figure 1: Box-plots of $\hat{\lambda}_{j+1}/\hat{\lambda}_j$ when $h_0 = 1, r = 3, (d_1, d_2, d_3) = (0.3, 0.2, 0.4)$.

technically and empirically in Molinares et al (2009), Lévy-Leduc et al (2011b), Reisen et al (2020) among others. Cotta et al (2023) extended these works to the multivariate case. From the results displayed in Tables 4 and 5 and the ones discussed by these authors, it appears that the percentage and magnitude of additive outliers are the main characteristics responsible for the effect of these observations on the sampling functions mentioned above. Since the Factorial Analysis, among other multivariate statistical tools, depends on the autocovariance and autocorrelation structures of stationary time series models, the impact of outliers on these functions is translated in some way to the other statistical quantities which are derived from them, such as \hat{r} and \hat{r}^Q discussed here.

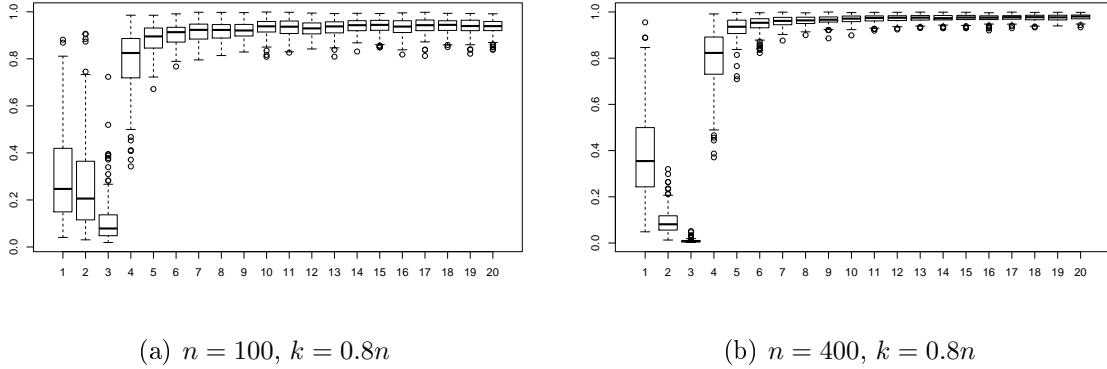


Figure 2: Box-plots of $\widehat{\lambda}_{j+1}^Q/\widehat{\lambda}_j^Q$ when $h_0 = 1, r = 3, (d_1, d_2, d_3) = (0.3, 0.2, 0.4)$.

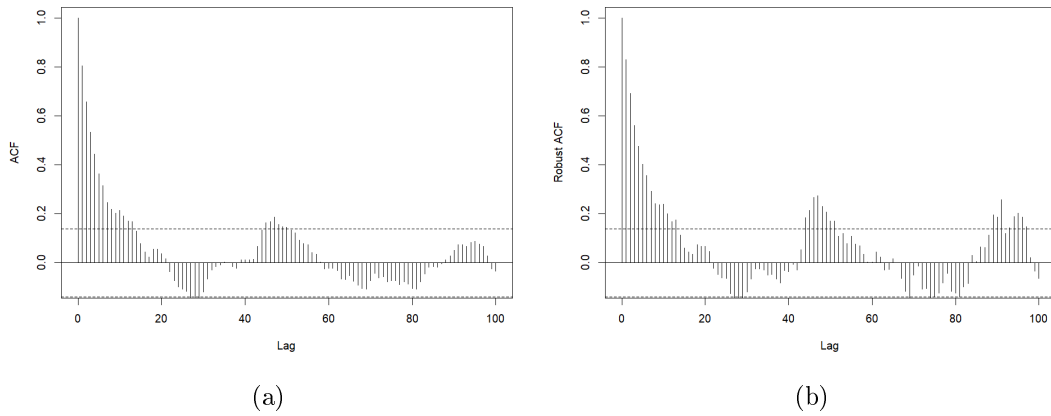


Figure 3: Classical (a) and robust (b) ACFs from the series with $d = 0.3, n = 200$, without outliers.

4 Real data example. FA for foreign exchange rates

In this section, we analyse the daily returns of the six following foreign exchange rates against U.S. Dollar: Australian Dollar (AUD), Canadian Dollar (CAD), British Pound (GBP), Norwegian kroner (NOK), Swedish Kroner (SEK), and Singapore (SGD). All the series are observed from January 2010 to December 2019 ($n = 2608$). We use adjusted daily opening values (end-of-session quotations).

Firstly, we estimate the fractional parameters for each series, assuming that they are long-memory processes. The estimates and their standard deviations are displayed in Table 8 where \widehat{d}_{GPH} is the standard semi-parametric GPH estimator and $\widehat{d}_{\text{RGPH}}$ is its robust version, see (Molinares et al, 2009). Following these authors, the number of observations in the regression equation (the bandwidth) was set to $m = n^\alpha$ with $\alpha = 0.5$; similar estimates

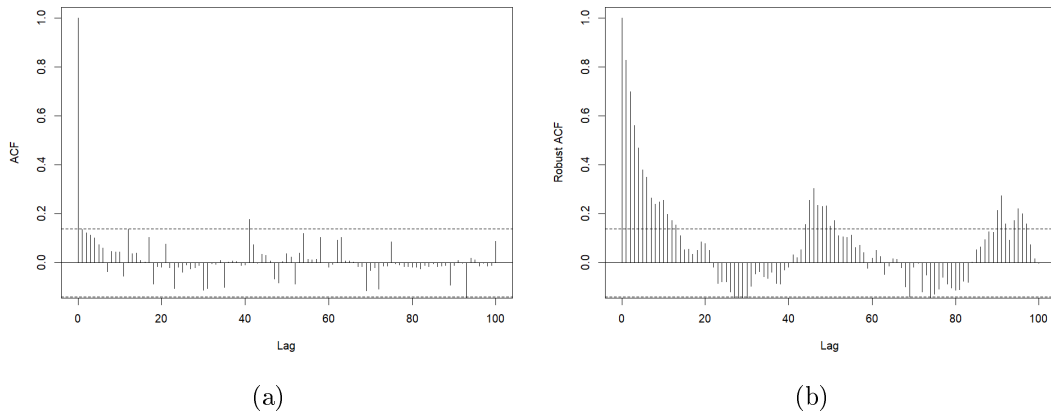


Figure 4: Classical (a) and robust (b) ACFs from the series with $d = 0.3$, $n = 200$, with outliers ($p = 0.05$, $\omega = 15$).

were obtained using $\alpha = 0.7, 0.8$. These estimators were computed using the package `tsqn` in R, see (Cotta et al, 2017). Both estimators strongly suggest that each exchange rate series is non-stationary, being either a random walk ($d = 1$) or a non-stationary long-memory process with $d > 1$. This phenomenon is quite common with financial time series. For each exchange rate, the standard and the robust estimators are different and $\hat{d}_{\text{RGPH}} > \hat{d}_{\text{GPH}}$, the former having a smaller standard deviation ($s.d$). This empirical feature indicates that there are possibly atypical observations that caused additive outlier effects on the estimates of d_{GPH} .

Table 8: Estimates d_{GPH} and d_{RGPH} with the standard deviations ($s.d$) in parenthesis for the exchange rates

Series	\hat{d}_{GPH}	\hat{d}_{RGPH}
AUD	1.130 (0.101)	1.257 (0.061)
CAD	1.043 (0.078)	1.186 (0.055)
GBP	0.987 (0.093)	1.087 (0.053)
NOK	0.996 (0.079)	1.182 (0.056)
SEK	1.125 (0.101)	1.256 (0.067)
SGD	1.036 (0.085)	1.097 (0.057)

Since each exchange rate series is non-stationary, the difference in the logarithm of each series was taken, that is for each series Z_t , we define the log returns by $Y_t = \ln(Z_t) - \ln(Z_{t-1})$. Figure 4 plots the log returns of the exchange rates and exhibits some observations with extreme values compared to the average level. As well-known, the sample mean, the sample variance and autocovariance function are very sensitive to extreme values in the sense that these type of observations increase the sample variance and reduce the sample

autocorrelation and, consequently, lead to a reduction of the estimate of the long-memory parameter. The effect of additive outliers on the long-memory estimators is well-addressed in the literature see, for instance, (Molinares et al, 2009). For example, in the case of log returns of SEK, the sample autocorrelation estimates for the lags $h = 1, 3, 5, 10$, using (8) and (10) are $-0.051, -0.003, -0.009, -0.0002$ and $-0.073, 0.033, 0.0152, -0.009$, respectively.

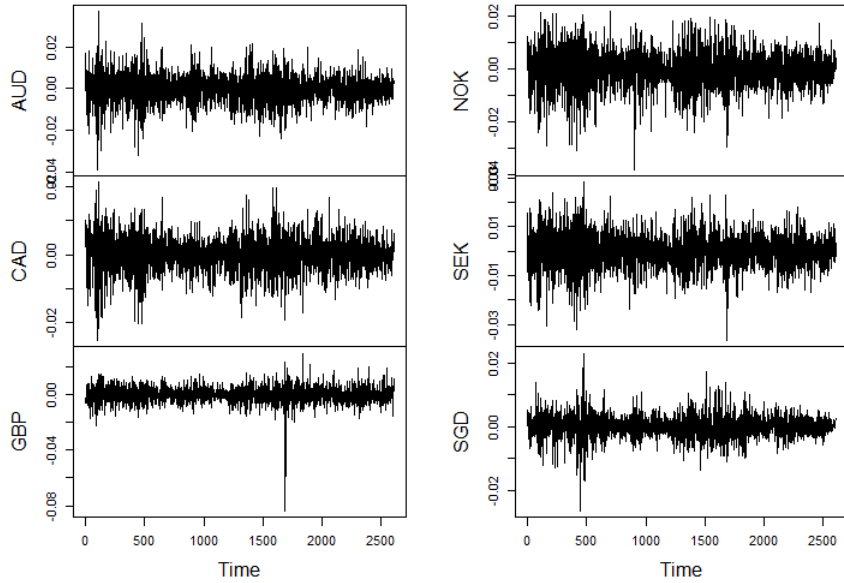


Figure 5: Log returns of the exchange rates.

The long-memory parameter estimators of the log returns of the exchange rates and their standard deviations are displayed in Table 9. The hypothesis test $H_0 : d = 0$ versus $H_1 : d \neq 0$ was tested using test statistics built from \hat{d}_{GPH} and \hat{d}_{RGPH} and by using the standard normal approximation coming from the central limit theorems established by Reisen et al (2017). The cases where H_0 has been rejected are with the symbol (*) in Table 9. We can see from this table that the null hypothesis of the test is rejected for almost all time series when the test statistic based on the robust estimator is used, which is not the case when the test statistic is based on the GPH estimator. As previously, the bandwidths in the regression equations of the estimators \hat{d}_{GPH} and \hat{d}_{RGPH} were set to $m = n^{1/2}$. As in Table 8, \hat{d}_{RGPH} is quite different from \hat{d}_{GPH} , and the former generally shows more significant estimates. This fact gives strong evidence of the presence of additive outliers in the data, since this kind of observation decreases the correlation structure of the data and, consequently, the sample ACF is completely crushed and the standard long-memory estimator underestimates the true parameter (Molinares et al, 2009).

The above results indicate that the standard and robust estimators of r may present different conclusions in identifying the number of the factors, *i.e.* it is expected that the

Table 9: Values of \widehat{d}_{GPH} and $\widehat{d}_{\text{RGPH}}$ with their standard deviations in parenthesis for the log returns of the exchange rates. The values with (*) corresponds to the rejection of the hypothesis test $H_0 : d = 0$ versus $H_1 : d \neq 0$ at the significance level of 5% of $\mathcal{N}(0, 1)$ where the test statistics are based either on \widehat{d}_{GPH} or $\widehat{d}_{\text{RGPH}}$.

Series	\widehat{d}_{GPH}	$\widehat{d}_{\text{RGPH}}$
AUD	0.073 (0.101)	0.024 (0.012)*
CAD	0.061 (0.109)	0.021 (0.015)
GBP	-0.090 (0.085)	0.064 (0.021)*
NOK	0.049 (0.102)	0.079 (0.010)*
SEK	0.091 (0.117)	0.113 (0.011)*
SGD	0.079 (0.095)	0.147 (0.012)*

number of factors obtained by the classical method will tend to be smaller than the robust one. This issue is discussed as follows.

We estimate r by \widehat{r} and \widehat{r}^Q taking $h_0 = 1$ in (6) and (11), respectively. Similar results were obtained with $h_0 = 3$. The eigenvalues $\widehat{\lambda}_1 \geq \dots \geq \widehat{\lambda}_6$ of $\widehat{\mathbf{M}}$ and their ratios $\widehat{\lambda}_{j+1}/\widehat{\lambda}_j$ are displayed in Figure 6. The robust versions obtained from $\widehat{\mathbf{M}}^Q$ are shown in Figure 7. Figure 6b gives $\widehat{r} = 1$ while Figure 7b gives $\widehat{r}^Q = 2$. Then, the standard and robust FA do not select the same number of factors. The estimates of the fractional parameters indicate the presence of additive outliers effects. These affect the estimates of the eigenvalues and, consequently, the estimates of the number of factors, as discussed in the Simulation Section.

Since the results using the autocovariance matrices may be affected by the distinct variability of the returns, the estimates of r were also performed by replacing the covariance matrices in (6) and (11) by the corresponding autocorrelation matrices. The estimates for the number of factors were the same.

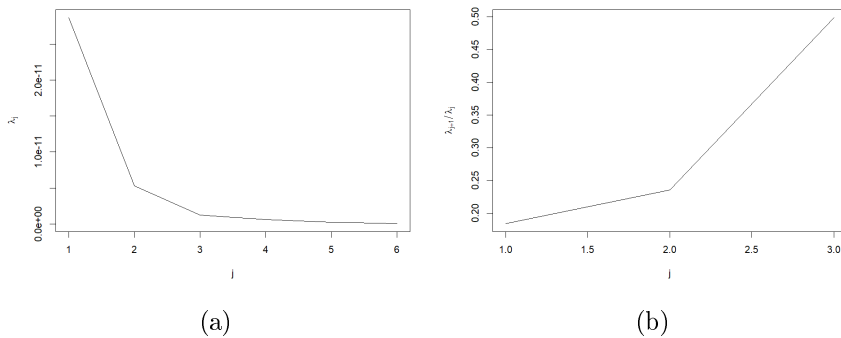


Figure 6: Eigenvalues $\widehat{\lambda}_j$ of $\widehat{\mathbf{M}}$ and their ratios $\widehat{\lambda}_{j+1}/\widehat{\lambda}_j$.

To check whether or not the abrupt observations appearing in the data (Figure 4) affect the estimated number of factors, the series were modified by replacing the abrupt

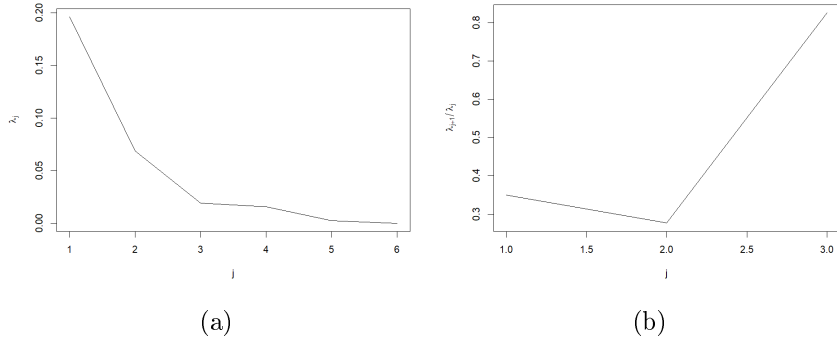


Figure 7: Eigenvalues $\hat{\lambda}_j^Q$ of $\widehat{\mathbf{M}}^Q$ and their ratios $\hat{\lambda}_{j+1}^Q/\hat{\lambda}_j^Q$.

observations by their sample mean. The ratios are given in Figures 8 and 9. From these we can see that the classical method (Figure 8) gives the same rates as the robust (Figure 9) one. This simple example corroborates the fact that \hat{r} lacks robustness toward additive outliers in the original data.

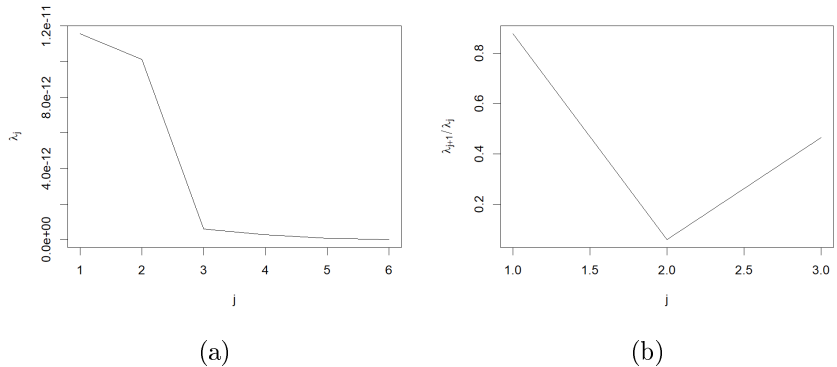


Figure 8: Eigenvalues $\hat{\lambda}_j$ of $\widehat{\mathbf{M}}$ and their ratios $\hat{\lambda}_{j+1}/\hat{\lambda}_j$, for real data without abrupt observations.

The FA can be used for forecasting purpose. Indeed, this is much simpler to use (1) than fitting a multivariate stationary time series model with dimension k to \mathbf{Y}_t . For instance, the h -step ahead linear forecast $\widehat{\mathbf{Y}}_{n+h}^{(h)}$ of \mathbf{Y}_n can be obtained by $\widehat{\mathbf{Y}}_{n+h}^{(h)} = \widehat{\mathbf{P}}\widehat{\mathbf{X}}_{n+h}^{(h)}$, where the columns of the estimated factor loading matrix $\widehat{\mathbf{P}}$ are the \hat{r} orthonormal eigenvectors of $\widehat{\mathbf{M}}^Q$ and $\widehat{\mathbf{X}}_{n+h}^{(h)}$ is the h -step ahead linear forecast of \mathbf{X}_n , based on the estimated past values $\widehat{X}_1, \dots, \widehat{X}_n$, see (Lam et al, 2011).

Furthermore, according to Hallin and Liška (2007), factor models with large dimensions are attracting increasing attention in economics, emphasizing finance and macroeconomic applications. They argue that since the classical multivariate time series tech-

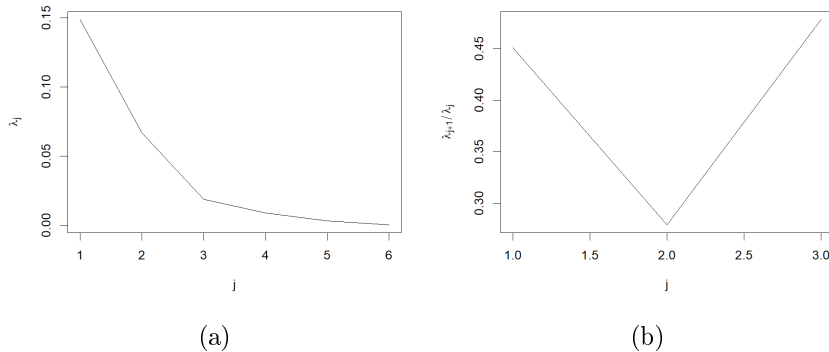


Figure 9: Eigenvalues $\widehat{\lambda}_j^Q$ of $\widehat{\mathbf{M}}^Q$ and their ratios $\widehat{\lambda}_{j+1}^Q/\widehat{\lambda}_j^Q$, for real data without abrupt observations.

niques are helpless, factor models have a natural role in economic applications with large databases. Determining the number of factors is a critical step in this context. For example, under the assumption of the APT theory given in (Ross, 1976), there are common risk factors across assets. The classical method indicated one common factor in the exchange rate data, while the robust method gave two factors. Thus, investors may view market risks incorrectly based on the classic model. Measuring and managing financial market risks is critical for market participants. Investors, especially traders, know of possible losses due to market fluctuations.

5 Conclusion

This paper extends the results on FA given by Reisen et al (2019) by allowing the time series to be a long-memory process. Asymptotic properties of the eigenvalues of the standard and robust covariance matrices estimates were established. The finite sample size investigation indicated that both estimators of the number of factors perform similarly in the uncontaminated data scenario. In contrast, the robust estimator is strongly recommended when the data contain additive outliers. In this context, classical methodology leads to a spurious choice of the number of factors. Thus, based on the technical and empirical results, the proposed robust estimator of the number of factors is suggested here to be applied in the context of long-memory time series with or without additive outliers. An application to exchange rates series was discussed and revealed that the robust approach suggested the number of factors equals two while the standard method indicated only one. The influence of outliers observations can justify this. As future work, one could consider robust alternative approaches to estimating $\mathbf{\Gamma}_Y(h)$ in (5), such as the one based on in the frequency domain, recently proposed by Lévy-Leduc et al (2022), and on ranking-based multivariate statistical tools, e.g, (He et al, 2022) and references therein.

6 Appendix: Technical lemmas

Lemmas 1 and 2 were stated and proved in Reisen et al (2019) but are recalled here for the reader convenience.

Lemma 1. Let $\widehat{\mathbf{A}}_n$ be a sequence of $p \times p$ symmetric matrices and \mathbf{A} a $p \times p$ symmetric matrix such that $u_n(\widehat{\mathbf{A}}_n - \mathbf{A}) = O_p(1)$, where u_n is a sequence of positive numbers tending to infinity as n tends to infinity, then

$$u_n \sup_{1 \leq j \leq p} |\lambda_j(\widehat{\mathbf{A}}) - \lambda_j(\mathbf{A})| = O_p(1), \quad \text{as } n \rightarrow \infty,$$

where $(\lambda_j(\widehat{\mathbf{A}}))_{1 \leq j \leq p}$ and $(\lambda_j(\mathbf{A}))_{1 \leq j \leq p}$ are the eigenvalues of $\widehat{\mathbf{A}}_n$ and \mathbf{A} , respectively.

Lemma 2. Let $\widehat{\mathbf{A}}_n(h)$ be a sequence of $p \times p$ symmetric matrices and $\mathbf{A}(h)$ a $p \times p$ symmetric matrix such that $u_n(\widehat{\mathbf{A}}_n(h) - \mathbf{A}(h)) = O_p(1)$, for each fixed $h \in \{1, \dots, h_{max}\}$, where u_n is a sequence of positive numbers tending to infinity as n tends to infinity, then

$$u_n \left(\sum_{h=1}^{h_{max}} \widehat{\mathbf{A}}_n(h) \widehat{\mathbf{A}}_n(h)' - \sum_{h=1}^{h_{max}} \mathbf{A}(h) \mathbf{A}(h)' \right) = O_p(1),$$

as n tends to infinity.

Lemma 3. Let h be a non negative integer and i and j two integers in $\{1, \dots, k\}$. Assume that (A2) holds, then the autocovariance estimator $\widehat{\gamma}_{i,j}^Q(h)$ defined in (10) satisfies the following limit theorems as n tends to infinity.

(i) If, for all i in $\{1, \dots, k\}$, $D_i > 1/2$,

$$\sqrt{n}(\widehat{\gamma}_{i,j}^Q(h) - \gamma_{ij}(h)) \xrightarrow{d} \mathcal{N}(0, \widetilde{\sigma}_{i,j}^2(h)), \quad \text{as } n \rightarrow \infty,$$

where

$$\widetilde{\sigma}_{i,j}^2(h) = [\psi(Y_{i,1}, Y_{j,1+h})^2] + 2 \sum_{k \geq 1} \mathbb{E}[\psi(Y_{i,1}, Y_{j,1+h}) \psi(Y_{i,k+1}, Y_{j,k+1+h})],$$

where ψ is

$$\begin{aligned} \psi(x, y) = & \frac{1}{2} (\gamma_{i,i}(0) + \gamma_{j,j}(0) + 2\gamma_{i,j}(h)) \text{IF} \left(\frac{x+y}{\sqrt{\gamma_{i,i}(0) + \gamma_{j,j}(0) + 2\gamma_{i,j}(h)}}, Q, \Phi \right) \\ & - \frac{1}{2} (\gamma_{i,i}(0) + \gamma_{j,j}(0) - 2\gamma_{i,j}(h)) \text{IF} \left(\frac{x-y}{\sqrt{\gamma_{i,i}(0) + \gamma_{j,j}(0) - 2\gamma_{i,j}(h)}}, Q, \Phi \right), \end{aligned} \quad (12)$$

and IF is defined in (Lévy-Leduc et al, 2011b, Equation (20)).

(ii) If, there exists i_0 in $\{1, \dots, k\}$ such that $D_{i_0} < 1/2$,

$$n^{D_{i_0} \wedge D_j} (\widehat{\gamma}_{i_0, j}^Q(h) - \gamma_{i_0, j}(h)) = O_P(1), \text{ as } n \rightarrow \infty.$$

Proof of Lemma 3. Observe that the autocovariance $\gamma_{i, j}^{(+)}(\ell)$ of the process $(Y_{i, t} + Y_{j, t+h})_{t \geq 1}$ is equal to

$$\gamma_{i, j}^{(+)}(\ell) = \text{Cov}(Y_{i, t} + Y_{j, t+h}, Y_{i, t+\ell} + Y_{j, t+h+\ell}) = \gamma_{i, i}(\ell) + \gamma_{i, j}(h + \ell) + \gamma_{i, j}(\ell - h) + \gamma_{j, j}(\ell).$$

By (A2) and by using a Taylor expansion, $\gamma_{i, j}^{(+)}(\ell)$ is proportional to $\ell^{D_i \wedge D_j}$. Hence, the process $(Y_{i, t} + Y_{j, t+h})_{t \geq 1}$ satisfies (Lévy-Leduc et al, 2011b, Assumption (A2)) with $D = D_i \wedge D_j$. Since the autocovariance $\gamma_{i, j}^{(-)}(\ell)$ of the process $(Y_{i, t} - Y_{j, t+h})_{t \geq 1}$ is equal to

$$\gamma_{i, j}^{(-)}(\ell) = \text{Cov}(Y_{i, t} - Y_{j, t+h}, Y_{i, t+\ell} - Y_{j, t+h+\ell}) = \gamma_{i, i}(\ell) - \gamma_{i, j}(h + \ell) - \gamma_{i, j}(\ell - h) + \gamma_{j, j}(\ell),$$

by following the same lines, the process $(Y_{i, t} - Y_{j, t+h})_{t \geq 1}$ also satisfies (Lévy-Leduc et al, 2011b, Assumption (A2)) with $D = D_i \wedge D_j$. In the case (i), the proof follows the same lines as the ones of the proof of (i) in (Lévy-Leduc et al, 2011b, Theorem 4). In the case (ii), by applying the Delta method to (Lévy-Leduc et al, 2011b, Equation (74)), we get

$$n^{D_{i_0} \wedge D_j} (Q_{n-h}(Y_{i_0, 1:n-h} + Y_{j, h+1:n})^2 - \text{Var}(Y_{i_0, t} + Y_{j, t+h})) = O_P(1).$$

Similarly, we have that

$$n^{D_{i_0} \wedge D_j} (Q_{n-h}(Y_{i_0, 1:n-h} - Y_{j, h+1:n})^2 - \text{Var}(Y_{i_0, t} - Y_{j, t+h})) = O_P(1),$$

which gives the result. \square

Lemma 4. Let h be a non negative integer and i and j two integers in $\{1, \dots, k\}$. Assume that (A2) holds, then the autocovariance estimator $\widehat{\gamma}_{i, j}(h)$ defined in (8) satisfies the following limit theorems as n tends to infinity.

(i) If, for all i in $\{1, \dots, k\}$, $D_i > 1/2$,

$$\sqrt{n}(\widehat{\gamma}_{i, j}(h) - \gamma_{i, j}(h)) \xrightarrow{d} \mathcal{N}(0, \check{\sigma}_{i, j}^2(h)), \text{ as } n \rightarrow \infty,$$

where

$$\begin{aligned} \check{\sigma}_{i, j}^2(h) &= \mathbb{E} [(Y_{i, 1} Y_{j, h+1} - \gamma_{i, j}(h))^2] \\ &+ 2 \sum_{k \geq 1} \mathbb{E} [(Y_{i, 1} Y_{j, h+1} - \gamma_{i, j}(h)) (Y_{i, 1+k} Y_{j, 1+h+k} - \gamma_{i, j}(h))]. \end{aligned}$$

(ii) If, there exists i_0 in $\{1, \dots, k\}$ such that $D_{i_0} < 1/2$,

$$n^{D_{i_0} \wedge D_j} (\widehat{\gamma}_{i_0, j}(h) - \gamma_{i_0, j}(h)) = O_P(1), \text{ as } n \rightarrow \infty.$$

Proof of Lemma 4. (i) Note that

$$\widehat{\gamma}_{i,j}(h) = \frac{1}{n} \sum_{t=1}^{n-h} Y_{i,t} Y_{j,t+h} - \bar{Y}_i \bar{Y}_j.$$

By Theorem 5.1 of Taqqu (1975),

$$\bar{Y}_i = O_P(n^{-D_i/2}) \text{ and } \bar{Y}_j = O_P(n^{-D_j/2}).$$

Let $\mathcal{Y}_t = (Y_{i,t}, Y_{j,t+h})$ and $f : (x, y) \mapsto xy$ then, by Theorem 4 of Arcones (1994), we get that

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{n-h} (f(\mathcal{Y}_t) - \mathbb{E}(f(\mathcal{Y}_t))) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n-h} (Y_{i,t} Y_{j,t+h} - \gamma_{ij}(h)) \xrightarrow{d} \mathcal{N}(0, \check{\sigma}_{i,j}^2(h)) \quad (13)$$

since f is of Hermite rank 2, $r^{(1,2)}(h) = \mathbb{E}[Y_{i,t} Y_{j,t+h}] = \gamma_{i,j}(h)$, $r^{(1,1)}(h) = \mathbb{E}[Y_{i,t} Y_{i,t+h}] = \gamma_{i,i}(h)$, $r^{(2,2)}(h) = \mathbb{E}[Y_{j,t} Y_{j,t+h}] = \gamma_{j,j}(h)$ and $D_i > 1/2$, for all i . In (13),

$$\begin{aligned} \check{\sigma}_{i,j}^2(h) &= \mathbb{E} [(Y_{i,1} Y_{j,h+1} - \gamma_{i,j}(h))^2] \\ &\quad + 2 \sum_{k \geq 1} \mathbb{E} [(Y_{i,1} Y_{j,h+1} - \gamma_{i,j}(h)) (Y_{i,1+k} Y_{j,1+h+k} - \gamma_{i,j}(h))]. \end{aligned}$$

(ii) Observe that

$$\widehat{\gamma}_{i,j}(h) = \frac{1}{4} \left(\widehat{\sigma}_{n-h, Y_{i,1:n-h} + Y_{j,h+1:n}}^2 - \widehat{\sigma}_{n-h, Y_{i,1:n-h} - Y_{j,h+1:n}}^2 \right) \quad (14)$$

where

$$\begin{aligned} \widehat{\sigma}_{n-h, Y_{i,1:n-h} + Y_{j,h+1:n}}^2 &=: \widehat{\sigma}_{+,i,j}^2 = \frac{1}{n} \sum_{t=1}^{n-h} (Y_{i,t} + Y_{j,t+h})^2 - (\bar{Y}_i + \bar{Y}_j)^2 \\ &\quad \text{with } \bar{Y}_i = \frac{1}{n} \sum_{t=1}^n Y_{i,t} \text{ and } \bar{Y}_j = \frac{1}{n} \sum_{t=1}^n Y_{j,t} \end{aligned}$$

and

$$\widehat{\sigma}_{n-h, Y_{i,1:n-h} - Y_{j,h+1:n}}^2 =: \widehat{\sigma}_{-,i,j}^2 = \frac{1}{n} \sum_{t=1}^{n-h} (Y_{i,t} - Y_{j,t+h})^2 - (\bar{Y}_i - \bar{Y}_j)^2.$$

By using the same arguments as those used in the proof of Lemma 3, we get that $(Y_{i,t} + Y_{j,t+h})_{t \geq 1}$ and $(Y_{i,t} - Y_{j,t+h})_{t \geq 1}$ satisfy (Lévy-Leduc et al, 2011b, Assumption (A2)) with $D = D_i \wedge D_j$. Since $D_{i_0} < 1/2$, $D_{i_0} \wedge D_j < 1/2$ for all j in $\{1, \dots, k\}$. Thus, $(Y_{i_0,t} + Y_{j,t+h})_{t \geq 1}$ and $(Y_{i_0,t} - Y_{j,t+h})_{t \geq 1}$ satisfy Assumption (A2) with $D < 1/2$. Hence, by (Lévy-Leduc et al, 2011b, Proposition 3(b)) and the Delta method,

$$n^{D_{i_0} \wedge D_j} (\widehat{\sigma}_{+,i_0,j}^2 - \text{Var}(Y_{i_0,t} + Y_{j,t+h})) = O_p(1), \text{ as } n \rightarrow \infty$$

and

$$n^{D_{i_0} \wedge D_j} (\hat{\sigma}_{-,i_0,j}^2 - \text{Var}(Y_{i_0,t} - Y_{j,t+h})) = O_p(1), \text{ as } n \rightarrow \infty,$$

which concludes the proof by (14). □

Acknowledgments

The authors gratefully acknowledge CNPq, CAPES and FAPES for their financial support. This research was also partially supported by DATAIA Convergence Institute as part of the “Programme d’Investissement d’Avenir, (ANR17-CONV-0003) operated by Centrale-Supélec”. The authors thank the referees for their valuable suggestions that have led to clarifying and substantially improving the quality of the paper.

References

- Arcones MA (1994) Limit theorems for nonlinear functionals of a stationary gaussian sequence of vectors. *Ann Probab* 22(4):2242–2274. <https://doi.org/10.1214/aop/1176988503>
- Bai J, Wang P (2016) Econometric analysis of large factor models. *Annu Rev Econom* 8(1):53–80
- Bai X, Zheng L (2023) Robust factor models for high-dimensional time series and their forecasting. *Commun Stat* 52(19):1–14. <https://doi.org/10.1080/03610926.2022.2033777>
- Brillinger DR (1981) *Time series, data analysis and theory*. Cambridge Series in Statistical and Probabilistic
- Brockwell P, Davis R (2009) *Time Series: Theory and Methods*. Springer Series in Statistics, Springer
- Chen EY, Tsay RS, Chen R (2020) Constrained factor models for high-dimensional matrix-variate time series. *J Am Stat Assoc* 115(530):775–793
- Christou E (2020) Robust dimension reduction using sliced inverse median regression. *Stat Pap* 61:1799–1818. <https://doi.org/10.1007/s00362-018-1007-z>
- Chung C (2002) Sample means, sample autocovariances, and linear regression of stationary multivariate long memory processes. *Econ Theory* 18:51–78
- Cotta H, Reisen V, Bondon P, et al (2017) *tsqn: Applications of the qn estimator to time series (univariate and multivariate)*. URL <https://CRAN.R-project.org/package=tsqn>, r package version 1.0.0
- Cotta H, Reisen VA, Bondon P (2023) A robust alternative method for the estimation of the covariance and the correlation matrices for multivariate time series. Unpublished <https://doi.org/10.13140/RG.2.2.24313.65129>
- Eichler M, Motta G, von Sachs R (2011) Fitting dynamic factor models to non-stationary time series. *J Econometrics* 163(1):51–70
- Fan J, Wang K, Zhong Y, et al (2021) Robust high dimensional factor models with applications to statistical machine learning. *Stat Sci* 36(2):303–327
- Fernández-Macho FJ (1997) A dynamic factor model for economic time series. *Kybernetika* 33(6):583–606
- Geweke J (1977) The dynamic factor analysis of economic time series. In: Aigner DJ, Goldberger AS (eds) *Latent variables in socio-economic models*. North-Holland, Amsterdam, p 365–383

- Geweke JF, Singleton KJ (1981) Latent variable models for time series: A frequency domain approach with an application to the permanent income hypothesis. *J Econometrics* 17(3):287–304
- Hallin M, Liška R (2007) Determining the number of factors in the general dynamic factor model. *J Am Stat Assoc* 102(478):603–617
- He Y, Wang Y, Yu L, et al (2022) Matrix kendall’s tau in high-dimensions: A robust statistic for matrix factor model
- Lam C, Yao Q (2012) Factor modeling for high-dimensional time series: Inference for the number of factors. *Ann Stat* 40(2):694–726. <https://doi.org/10.1214/12-AOS970>, URL <http://dx.doi.org/10.1214/12-AOS970>
- Lam C, Yao Q, Bathia N (2011) Estimation of latent factors for high-dimensional time series. *Biometrika* 98(4):901–918
- Lévy-Leduc C, Boistard H, Moulines E, et al (2011a) Large sample behavior of some well-known robust estimators under long-range dependence. *Stat* 45(1):59–71
- Lévy-Leduc C, Boistard H, Moulines E, et al (2011b) Robust estimation of the scale and of the autocovariance function of gaussian short-and long-range dependent processes. *J Time Ser Anal* 32(2):135–156
- Lin TI, Chen IA, Wang WL (2022) A robust factor analysis model based on the canonical fundamental skew-t distribution. *Stat Pap* pp 1–27
- Lévy-Leduc C, Bondon P, Reisen VA (2022) Robust autocovariance estimation from the frequency domain for univariate stationary time series. *J Stat Plan Inference* 221:281–298
- Ma Y, Genton MG (2000) Highly robust estimation of the autocovariance function. *J Time Ser Anal* 21:663–684
- Molinares FF, Reisen VA, Cribari-Neto F (2009) Robust estimation in long-memory processes under additive outliers. *J Stat Plan Inference* 139(8):2511–2525
- Peña D, Box GEP (1987) Identifying a simplifying structure in time series. *J Am Stat Assoc* 82(399):pp. 836–843. URL <http://www.jstor.org/stable/2288794>
- Priestley M, Rao T, Tong H (1974) Applications of principal component analysis and factor analysis in the identification of multivariable systems. *IEEE Trans Automat Contr* 19(6):730–734
- Reinsel G (2003) *Elements of Multivariate Time Series Analysis*. Springer Series in Statistics, Springer New York
- Reisen VA, Lévy-Leduc C, Taqqu MS (2017) An M-estimator for the long-memory parameter. *J Stat Plan Inference* 187:44–55

- Reisen VA, Sgrâncio AM, Lévy-Leduc C, et al (2019) Robust factor modeling for high-dimensional time series: an application to air pollution data. *Appl Math Comput* 346:842–852
- Reisen VA, Lévy-Leduc C, Bondon P, et al (2020) An overview of robust spectral estimators. In: Chaari F., Leskow J., Zimroz R., Wyłomańska A., Dudek A. (eds) *Cyclostationarity: Theory and Methods – IV. CSTA 2017. Applied Condition Monitoring Time Series and Cyclostationary process*, vol 16. Springer, Cham., https://doi.org/https://doi.org/10.1007/978-3-030-22529-2_12
- Rooch A, Zelo I, Fried R (2019) Estimation methods for the LRD parameter under a change in the mean. *Stat Pap* 60:313–347. <https://doi.org/10.1007/s00362-016-0839-7>
- Ross SA (1976) The arbitrage theory of capital asset pricing. *J Econ Theory* 13(3):341–360
- Rousseeuw PJ, Croux C (1993) Alternatives to the median absolute deviation. *J Am Stat Assoc* 88(424):1273–1283
- Taqqu MS (1975) Weak convergence to fractional Brownian motion and to the Rosenblatt process. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 31(4):287–302. <https://doi.org/https://doi.org/10.1007/BF00532868>
- Toman A (2014) Robust confirmatory factor analysis based on the forward search algorithm. *Stat Pap* 55:233–252
- Velu RP, Reinsel GC, W. WD (1986) Reduced rank models for multiple time series. *Biometrika* 73(1):105–118