Applied Condition Monitoring

Fakher Chaari
Jacek Leskow
Radoslaw Zimroz
Agnieszka Wyłomańska
Anna Dudek *Editors*

Cyclostationarity: Theory and Methods – IV

Contributions to the 10th Workshop on Cyclostationary Systems and Their Applications, February 2017, Grodek, Poland

Springer

Applied Condition Monitoring

Volume 16

The book series Applied Condition Monitoring publishes the latest research and developments in the field of condition monitoring, with a special focus on industrial applications. It covers both theoretical and experimental approaches, as well as a range of monitoring conditioning techniques and new trends and challenges in the field. Topics of interest include, but are not limited to: vibration measurement and analysis; infrared thermography; oil analysis and tribology; acoustic emissions and ultrasonics; and motor current analysis. Books published in the series deal with root cause analysis, failure and degradation scenarios, proactive and predictive techniques, and many other aspects related to condition monitoring. Applications concern different industrial sectors: automotive engineering, power engineering, civil engineering, geoengineering, bioengineering, etc. The series publishes monographs, edited books, and selected conference proceedings, as well as textbooks for advanced students.** Indexing: Indexed by SCOPUS and Springerlink. The books of the series are submitted for indexing to Web of Science.

More information about this series at http://www.springer.com/series/13418

Fakher Chaari · Jacek Leskow ·
Radoslaw Zimroz · Agnieszka Wyłomańska ·
Anna Dudek
Editors

Cyclostationarity: Theory and Methods – IV

Contributions to the 10th Workshop on Cyclostationary Systems and Their Applications, February 2017, Grodek, Poland

Springer

Editors
Fakher Chaari
Department of Mechanical Engineering
National School of Engineers of Sfax
Sfax, Tunisia

Jacek Leskow
Institute of Mathematics
Cracow Technical University
Krakow, Poland

Radoslaw Zimroz
Faculty of GeoEngineering, Mining and Geology
Wrocław University of Science and Technology
Wroclaw, Poland

Agnieszka Wyłomańska
Faculty of Pure and Applied Mathematics
Wrocław University of Science and Technology
Wroclaw, Poland

Anna Dudek
AGH University of Science and Technology
Krakow, Poland

ISSN 2363-698X ISSN 2363-6998 (electronic)
Applied Condition Monitoring
ISBN 978-3-030-22528-5 ISBN 978-3-030-22529-2 (eBook)
https://doi.org/10.1007/978-3-030-22529-2

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Similar to our previous books on cyclostationarity, we present here a selection of 12 interesting papers dealing with cyclostationary and general non-stationary processes. These papers are devoted either to practical aspects or to theoretical ones. They report on interdisciplinary research and applications at the border between mathematics, statistics, and signal processing. Although the book has not been organized into different parts, we may identify three different classes of works among the ones presented here.

The first class is related to the theoretical aspects of periodically correlated (PC or cyclostationary) processes analysis. A few papers are related to the estimation problem of PC process parameters. Javors'kyj et. al describe in their chapter the coherent and component estimation of covariance invariants for multivariate periodically correlated random processes. Reisen et al. describe robust estimation functions applied to the estimation of the spectral density of univariate and periodic time series with short- and long-memory processes. A long-memory process, such as the multifractional Brownian motion with Hoelder function exponent, is also the topic of the paper of Mastalerz-Kodzis. Here, the author proposes the use of the least squares method for the estimation of the pointwise Hoelder exponent. The multifractal Brownian motion with Hoelder function exponent belongs to the class of non-stationary processes with the so-called anomalous diffusion behaviour.

The second group of papers in this volume is devoted to the general class of non-stationary processes. Grzesiek and Wyłomańska describe subordinated processes with infinite variance. Special attention is paid to the stable Levy process subordinated by different non-decreasing processes. In the area of the processes of infinite variance, we also find the paper by Gajecka-Mirek and Leśkow, where the authors consider the subsampling for heavy-tailed non-stationary and weakly dependent time series. Poczynek et al. describe the analysis of the classical Ornstein–Uhlenbeck process time changed by Gamma process and show an application to the modelling of real financial time series, especially of interest rate data. Another process that is especially useful to describe interest rates is presented in the paper by Brzozowska-Rup. Here, the author examines the Cox–Ingersoll–Ross (CIR) model in the context of spot short interest rates on the

Polish market. The paper by Doukhan and Gomez, the last of this second group of papers, is devoted to the analysis of the extreme values of stationary and weakly dependent random fields.

The last group of papers deals with the practical aspects of non-stationary and PC processes analysis. Hurd and Pipiras report on an interesting analysis of the hourly electricity volumes from Nord Pool Spot Exchange. For this, they use periodic autoregressive (PAR) time series models and consider the theoretical aspects of the modelling of PAR time series with multiple periodic effects. Michalak et al. apply the co-integration approach to vibration signal and present a complete procedure for the evaluation of the detection effectiveness with respect to changing signal-to-noise ratio. They show that co-integration is strictly related to the periodic correlation in time series models. The last paper by Wodecki et al. describes an innovative method for detecting P-wave arrival. At first, they show how to use the Kolmogorov–Smirnov statistic value to measure the non-Gaussianity of the seismic signal in the time–frequency representation. Then, they test their approach in the detection of the so-called P-wave arrival.

Contents

Modeling Periodic Autoregressive Time Series with Multiple Periodic Effects

Harry Hurd and Vladas Pipiras(✉)

Department of Statistics and Operations Research,
University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-3260, USA
hhurd1@nc.rr.com, pipiras@email.unc.edu

Abstract. Two models of periodic autoregressive time series with multiple periodic effects are introduced and studied. In the first model, the autoregression coefficients vary periodically with several dominant components associated with two or more periods (for example, day and week for hourly data). In the second model, the autoregression coefficients consist of the additive periodic effects of several nominal variables (for example, the effect of hour in a given day and the effect of day in a given week for hourly data). Truncated Fourier representations of different periods are used to parametrize the autoregression coefficients in the two models. Model estimation and inference through ordinary and weighted least squares, and model selection based on diagnostics plots, in particular, are considered for the two approaches. An application to a real time series of hourly electricity volumes from the Nord Pool Spot Exchange is also presented, where the nature and use of the two models are contrasted.

1 Introduction

Many data collected in time exhibit cyclical variations, and call for time series models with cyclical features. One class of such models consists of time series with periodically varying dependence structures. The periodicity could be in the mean, the variance, but also in the model parameters such as with periodic autoregressive (PAR) models that play a central role in this class of models. See Ghysels and Osborne [13], Franses and Paap [12], Hurd and Miamee [16].

In this work, we are interested in periodically correlated time series and, more specifically, PAR series where periodicity is driven by two or more periods. Having cyclical variations at multiple periods is expected in many data, especially when they are associated with natural cycles of 24 h, 1 week (when modeling human related activity), 4 annual quarters or seasons, and so on. We shall introduce two classes of periodically non-stationary time series that will operate at two or more periods.

To motivate briefly the construction of the models and to explain the basic ideas, suppose the goal is to model just the deterministic mean function $\mu(t)$ of the series as a function with two periodic effects. As with the application

F. Chaari et al. (Eds.): CSTA 2017, ACM 16, pp. 1–18, 2020.
https://doi.org/10.1007/978-3-030-22529-2_1

considered below in this work, suppose time t refers to hours and the two periodic effects are associated with the 24 h (1 day) and 168 h (1 week) periods. Two natural candidates for $\mu(t)$ operating at these two different periods come to mind, namely,

$$\mu(t) = \mu_{24}(t) + \mu_{168}(t), \tag{1}$$

where, for example, in the first case,

$$\mu_{24}(t) = 2 + 0.5\cos(\frac{2\pi t}{24}), \quad \mu_{168}(t) = -0.1\sin(\frac{2\pi t}{168}), \tag{2}$$

and, in the second case,

$$\begin{aligned} \mu_{24}(t) &= 1 - (0.2)1_{1\text{AM}}(t) + (0.3)1_{2\text{AM}}(t) + (0.7)1_{7\text{PM}}(t), \\ \mu_{168}(t) &= (0.3)1_{\text{Monday}}(t) - (0.1)1_{\text{Wednesday}}(t) + (4)1_{\text{Sunday}}(t), \end{aligned} \tag{3}$$

where $1_E(t)$ stands for the indicator function of "event" E, that is, it is equal to 1 if t falls into E, and 0 otherwise. The mean function $\mu(t)$ in (1) and (2) consists of two dominant components, one with period 24 and the other with period 168. The mean function $\mu(t)$ in (1) and (3), on the other hand, expresses the idea that the mean effect can be due to the hour of a given day or the day of a given week.

Our models for PAR time series with multiple periodic effects will allow for such periodic behavior for all model parameters, not just the mean function. The model extending (2) will be referred to as the model of Type A, and that extending (3) as the model of Type B. As with (2), we shall use Fourier representations of periodic model coefficients that will often require estimating fewer coefficients.

A number of other authors also considered various models exhibiting cyclical variations at several periods. For example, Gould et al. [14], De Livera et al. [6] and others consider models involving multiple periods based on exponential smoothing. The use of double seasonal ARIMA models (that is, seasonal ARIMA models with two periods) goes back at least to Box et al. [4]. Basawa et al. [3] do not quite have multiple periods but consider a hybrid model exhibiting both seasonal and periodic dependence for the same period. Neural networks in the context of multiple periods were used by Dudek [10,11] and others. Comparison of various available methods involving multiple periods can be found in Taylor et al. [20]. Applications to electricity markets dominate many of these contributions; see also Weron [22], Dannecker [5].

Our data application is also related to electricity markets. But we do not seek to provide an exhaustive comparison of our approach to other methods. The goal is to explain how one could think of periodic autoregressive time series with multiple periods at a most basic level, and how the resulting models could be estimated and manipulated in other ways. Though we also note that the introduced models do seem relevant for the considered data set.

The structure of the paper is as follows. The models of Types A and B are defined in Sect. 2 below. Estimation issues are discussed in Sect. 3, and inference, model selection and other issues in Sect. 4. A data application is considered in Sect. 5. Conclusions can be found in Sect. 6.

2 PAR Models with Multiple Periodic Effects

For the sake of clarity, we focus on PAR models with two periodic effects and comment on the case of multiple periodic effects in Remarks 2 and 3 below. The two periodic effects will be associated with two periods that are denoted s_1, s_2. We shall suppose that $s_1 < s_2$ and s_2/s_1 is an integer. For example, in the application in Sect. 5 below, $s_1 = 24\,\mathrm{h}$ (1 day) and $s_2 = 24 \cdot 7 = 168\,\mathrm{h}$ (1 week).

2.1 Model A

To introduce our first model with two periodic effects, we need several preliminary observations and definitions. A function $f(t)$ is s-periodic if $f(t+s) = f(t)$ for all $t \in \mathbb{Z}$. Note that an s_1-periodic function is also s_2-periodic (with the assumptions on s_1, s_2 stated above). An s_2-periodic function $f(t)$ can always be expressed through a Fourier representation as

$$f(t) = f_0 + \sum_{m=1}^{\lfloor s_2/2 \rfloor} \Big(f_{1,m} \cos(\frac{2\pi m t}{s_2}) + f_{2,m} \sin(\frac{2\pi m t}{s_2}) \Big), \tag{4}$$

where $f_0, f_{1,m}, f_{2,m} \in \mathbb{R}$. It can then also be expressed (uniquely) as

$$f(t) = f_0 + f_1(t) + f_2(t), \tag{5}$$

where

$$\begin{aligned} f_1(t) &= \sum_{m_1=1}^{\lfloor s_1/2 \rfloor} \Big(f_{1,(s_2/s_1)m_1} \cos(\frac{2\pi (s_2/s_1) m_1 t}{s_2}) + f_{2,(s_2/s_1)m_1} \sin(\frac{2\pi (s_2/s_1) m_1 t}{s_2}) \Big) \\ &= \sum_{m_1=1}^{\lfloor s_1/2 \rfloor} \Big(f_{1,(s_2/s_1)m_1} \cos(\frac{2\pi m_1 t}{s_1}) + f_{2,(s_2/s_1)m_1} \sin(\frac{2\pi m_1 t}{s_1}) \Big) \end{aligned} \tag{6}$$

and

$$f_2(t) = \sum_{m=1,\ldots,\lfloor s_2/2 \rfloor; m/s_1 \notin \mathbb{Z}} \Big(f_{1,m} \cos(\frac{2\pi m t}{s_2}) + f_{2,m} \sin(\frac{2\pi m t}{s_2}) \Big). \tag{7}$$

We shall refer to $f_j(t)$ as the s_j-periodic component of $f(t)$, $j = 1, 2$.

The following definition concerns our first model with two periodic effects.

Definition 1. A time series $\{X_t\}_{t\in\mathbb{Z}}$ is *type A periodic autoregressive of order p* (A–PAR(p)) *with two periodic effects* if

$$X_t = \mu(t) + Y_t, \tag{8}$$

$$Y_t = \phi_1(t) Y_{t-1} + \ldots + \phi_p(t) Y_{t-p} + \sigma(t)\epsilon_t \tag{9}$$

with $\{\epsilon_t\}_{t\in\mathbb{Z}} \sim \mathrm{WN}(0,1)$ (that is, a white noise series with $\mathbb{E}\epsilon_t = 0$ and $\mathbb{E}\epsilon_t^2 = 1$) and s_2-periodic $\mu(t)$, $\sigma(t)^2$ and $\phi_r(t)$, $r = 1,\dots,p$, with the decompositions

$$\begin{aligned} \mu(t) &= \mu_0 + \mu_1(t) + \mu_2(t), \\ \sigma(t)^2 &= \sigma_0^2 + \sigma_1^{(2)}(t) + \sigma_2^{(2)}(t), \\ \phi_r(t) &= \phi_{r,0} + \phi_{r,1}(t) + \phi_{r,2}(t), \quad r = 1,\dots,p, \end{aligned} \tag{10}$$

as in (5), where at least one of the s_1-periodic components $\mu_1(t), \sigma_1^{(2)}(t), \phi_{r,1}(t)$, $r = 1,\dots,p$, is non-zero.

In practice, motivated by the representations (5)–(7), we shall model the coefficients $\phi_r(t)$ and their components $\phi_{r,1}(t)$ and $\phi_{r,2}(t)$ as

$$\phi_{r,j}(t) = \sum_{m_j=1}^{H_j} \Big(a_{r,m_j}^{(j)} \cos(\frac{2\pi m_j t}{s_j}) + b_{r,m_j}^{(j)} \sin(\frac{2\pi m_j t}{s_j}) \Big), \quad j = 1,2, \tag{11}$$

assuming $H_2 < s_2/s_1$ (which ensures that indices m_2 in (11) are not multiples of s_2/s_1). The indices $j = 1$ and $j = 2$ in (11) correspond to s_1-periodic and s_2-periodic components, respectively. Modeling periodic time series through the (reduced) Fourier representations of their coefficients goes back at least to Jones and Brelsford [17]. See also Dudek et al. [8] and references therein.

The parameters μ_0, $\mu_1(t)$, $\mu_2(t)$, σ_0^2, $\sigma_1^{(2)}(t)$, $\sigma_2^{(2)}(t)$, on the other hand, will be estimated in a nonparametric fashion, though a parametric route analogous to (11) is also a possibility. Note also that $\sigma_1^{(2)}(t)$, $\sigma_2^{(2)}(t)$ are not necessarily positive.

Remark 1. By the discussion above, the series $\{X_t\}_{t\in\mathbb{Z}}$ in Definition 1 is also PAR(p) with the larger period s_2. We also note that our main interest here is in such series $\{X_t\}_{t\in\mathbb{Z}}$ which are stable, that is, for which the multivariate VAR representation of the s_2-vector series $\{(X_{s_2(\tilde{t}-1)+1}, X_{s_2(\tilde{t}-1)+2}, \dots, X_{s_2\tilde{t}})'\}_{\tilde{t}\in\mathbb{Z}}$ is stable. Here and throughout, a prime indicates a vector or matrix transpose. Conditions for the latter are well-known in the literature; see, for example, Lütkepohl [19].

Remark 2. The framework described above can be extended straightforwardly to the case of multiple periods $s_1, s_2, \dots, s_K$, assuming that $s_1 < s_2 < \dots < s_K$ and s_K/s_j are integers. Though some caution would need to be exercised in how many terms in the Fourier representations are included when some multiples of two periods s_{j_1} and s_{j_2} are the same (and smaller than s_K).

2.2 Model B

We now turn to a different PAR model that builds on the idea behind the model (1) and (3) for the mean discussed in Sect. 1. We adopt the following quite general framework concerning two periodic effects.

We think of each time t and observation X_t as associated with two nominal variables, that vary periodically in time, and are interested to model their effects. We assume that the two variables have k_1 and k_2 levels, respectively. We shall represent the two nominal variables by two functions $g_1(t)$ and $g_2(t)$, assuming that they are s_1-periodic and s_2-periodic, respectively, and take values $\{1,\dots,k_1\}$ and $\{1,\dots,k_2\}$, respectively, that are associated with respective levels. As above, we assume that $s_1 < s_2$ and s_2/s_1 is an integer. It is not necessarily the case that $s_j = k_j$, as the following examples illustrate.

Example 1. In the application to hourly data in Sect. 5 below, the two periodic effects will be the effect of the hour in a day and the effect of the day in a week. For hourly data, these effects are periodic with periods $s_1 = 24\,\mathrm{h}$ (1 day) and $s_2 = 24 \cdot 7 = 168\,\mathrm{h}$ (1 week), respectively. The corresponding nominal variables have $k_1 = 24$ (hours 1 through 24) and $k_2 = 7$ (Monday through Sunday) levels, respectively. The effects can be captured through the two corresponding functions $g_1(t)$ and $g_2(t)$ with the properties described above. They can also be represented as

$$g_1(t) = t, \quad t = 1,\dots,24, \quad g_2(t) = \lceil \frac{t}{24} \rceil, \quad t = 1,\dots,168, \tag{12}$$

where $\lceil x \rceil$ denotes the ceiling integer part of x, and then extended periodically with their respective periods.

Example 2. One could have the second variable (function) in Example 1 having only $k_2 = 2$ levels (values), for workdays and weekends. Similarly, the first variable (function) in Example 1 could have $k_2 = 4$ levels (values), for night hours (1–6AM), morning hours (6AM–12PM), afternoon hours (12–6PM) and evening hours (6PM–12AM).

Definition 2. A time series $\{X_t\}_{t\in\mathbb{Z}}$ is *type B periodic autoregressive of order p* (B–PAR(p)) *with two periodic effects* if

$$X_t = \mu(t) + Y_t, \tag{13}$$
$$Y_t = \phi_1(t)Y_{t-1} + \ldots + \phi_p(t)Y_{t-p} + \sigma(t)\epsilon_t \tag{14}$$

with $\{\epsilon_t\}_{t\in\mathbb{Z}} \sim \mathrm{WN}(0,1)$ and

$$\begin{aligned} \mu(t) &= \mu_0 + \mu_1(g_1(t)) + \mu_2(g_2(t)), \\ \sigma(t)^2 &= \sigma_0^2 + \sigma_1^{(2)}(g_1(t)) + \sigma_2^{(2)}(g_2(t)), \\ \phi_r(t) &= \phi_{r,0} + \phi_{r,1}(g_1(t)) + \phi_{r,2}(g_2(t)), \quad r = 1,\dots,p, \end{aligned} \tag{15}$$

where the functions $g_1(t)$ and $g_2(t)$ are defined before Example 1, are associated with two nominal variables and are s_1-periodic and s_2-periodic, respectively.

Definition 2 requires further clarification. With $f(t)$ denoting $\mu(t)$, $\sigma(t)^2$ or $\phi_r(t)$, let

$$f(t) = f_0 + f_1(g_1(t)) + f_2(g_2(t)) \tag{16}$$

be the decomposition analogous to those in (15). Recall from above that $g_j(t)$ takes an integer value from 1 to k_j, which we shall denote by u_j. Thus, f_j acts on a value u_j as $f_j(u_j)$, where $u_j = g_j(t)$. For identifiability purposes, we assume that

$$\sum_{u_j=1}^{k_j} f_j(u_j) = 0, \quad j = 1, 2. \tag{17}$$

We also note that the function $f_j(g_j(t)))$ is s_j-periodic, $j = 1, 2$, and hence, with our assumptions on s_1, s_2, the function $f(t)$ is s_2-periodic with the larger s_2.

The function $f_j(u_j)$, $j = 1, 2$, $u_j = 1, \ldots, k_j$, can be expressed through a Fourier representation as

$$f_j(u_j) = \sum_{m_j=1}^{\lfloor k_j/2 \rfloor} \Big(f_{1,m_j}^{(j)} \cos(\frac{2\pi m_j u_j}{k_j}) + f_{2,m_j}^{(j)} \sin(\frac{2\pi m_j u_j}{k_j}) \Big). \tag{18}$$

In practice, to have fewer coefficients to estimate, we shall model the coefficients $\phi_r(t)$ and their components as

$$\phi_{r,j}(u_j) = \sum_{m_j=1}^{H_j} \Big(a_{r,m_j}^{(j)} \cos(\frac{2\pi m_j u_j}{k_j}) + b_{r,m_j}^{(j)} \sin(\frac{2\pi m_j u_j}{k_j}) \Big), \tag{19}$$

where $H_j \le \lfloor k_j/2 \rfloor$. The parameters $\mu_j(u_j)$, $\sigma_j^{(2)}(u_j)$, $j = 1, 2$, on the other hand, will be estimated in a nonparametric fashion, though again a parametric route analogous to (19) is also a possibility.

Example 3. We continue with the setting of Example 1. In this example, by combining (12) and (19), the functions $\phi_{r,j}(g_j(t))$ are modeled as

$$\phi_{r,1}(g_1(t)) = \sum_{m_1=1}^{H_1} \Big(a_{r,m_1}^{(1)} \cos(\frac{2\pi m_1 t}{24}) + b_{r,m_1}^{(1)} \sin(\frac{2\pi m_1 t}{24}) \Big) \tag{20}$$

and

$$\phi_{r,2}(g_2(t)) = \sum_{m_2=1}^{H_2} \Big(a_{r,m_2}^{(2)} \cos(\frac{2\pi m_2 \lceil t/24 \rceil}{7}) + b_{r,m_2}^{(2)} \sin(\frac{2\pi m_2 \lceil t/24 \rceil}{7}) \Big). \tag{21}$$

We note again that the function $\phi_{r,1}(g_1(t))$ is 24-periodic, and that $\phi_{r,2}(g_2(t))$ is 168-periodic but also constant over successive intervals of length 24.

Remark 3. The framework described above can be extended straightforwardly to the case of multiple periodic effects, by introducing additional functions $g_j(t)$ associated with these effects.

Remark 4. As A–PAR(p) models discussed in Remark 1, B–PAR(p) models are also PAR(p) models with the larger period s_2. It is instructive here to contrast

the two introduced models from the perspective of these standard PAR models. A PAR(p) model with period s_2 has its coefficients vary periodically with period s_2. These coefficients can always be expressed through a Fourier representation. In the applications of the A–PAR model, only a small number of these Fourier coefficients are assumed to be non-zero, more specifically, the first few in the Fourier representation and also the first few in the component of the representation that is s_1-periodic. The B–PAR model, on the other hand, assumes that the periodicity consist of two additive effects associated with two periodic nominal variables. The latter effects do not need to be components of the Fourier representation of the model coefficients (as, for example, the coefficients (21) above).

Remark 5. The preceding remark also suggests that A–PAR and B–PAR models might serve quite different purposes. By increasing the number of non-zero coefficients in the A–PAR model Fourier representation, one could effectively get any PAR model with period s_2. From this perspective, the A–PAR model is quite flexible. With the B–PAR model, on the other hand, one might be more interested in which effects and which of their levels are more pronounced in the dynamics of the PAR process. This is illustrated further in our application to a real data set in Sect. 5.

3 Estimation Procedure

We discuss here estimation of the parameters $\mu(t)$, $\sigma(t)^2$ and $\phi_r(t)$ of the A–PAR and B–PAR models, using the Fourier representations (11) and (19) of the parameters. The way the A–PAR and B–PAR models were introduced allows us to present essentially a unified estimation framework. We suppose that the observed data consist of observations $X_1, \ldots, X_T$, where the sample size T is a multiple of both s_1 and s_2 for simplicity.

3.1 Estimation of Mean

For an A–PAR model, we estimate the means as $\widehat{\mu}_0 = \overline{X}$ (the overall mean),

$$\widehat{\mu}_1(t) = \frac{1}{(T/s_1)} \sum_{n=1}^{T/s_1} (X_{t+s_1(n-1)} - \overline{X}), \quad t = 1, \ldots, s_1, \tag{22}$$

and extended periodically with period s_1 for other t's, and

$$\widehat{\mu}_2(t) = \frac{1}{(T/s_2)} \sum_{n=1}^{T/s_2} (X_{t+s_2(n-1)} - \widehat{\mu}_1(t)), \quad t = 1, \ldots, s_2, \tag{23}$$

and extended periodically with period s_2 for other t's. Once can check that $\widehat{\mu}(t) = \widehat{\mu}_0 + \widehat{\mu}_1(t) + \widehat{\mu}_2(t)$ is just the periodic mean at period s_2. For a B–PAR model, the mean effects are estimated through a least squares regression of X_t on the two nominal variables described in the beginning of Sect. 2.2. Again, let $\widehat{\mu}(t)$ be the overall estimated mean which is generally different from that for the A–PAR model (see Fig. 1 in Sect. 5 for an illustration).

3.2 OLS Estimation

Let $\widehat{Y}_t = X_t - \widehat{\mu}(t)$. In applying the ordinary least squares (OLS), the model parameters are estimated as

$$\left\{\widetilde{\phi}_{r,0}, \widetilde{a}^{(j)}_{r,m_j}, \widetilde{b}^{(j)}_{r,m_j}\right\}_{r=1,\ldots,p,m_j=1,\ldots,H_j,j=1,2}$$
$$= \underset{\phi_{r,0},,a^{(j)}_{r,m_j},b^{(j)}_{r,m_j}}{\operatorname{argmin}} \sum_t (\widehat{Y}_t - \phi_1(t)\widehat{Y}_{t-1} - \ldots - \phi_p(t)\widehat{Y}_{t-p})^2, \tag{24}$$

where $\phi_r(t) = \phi_{r,0} + \phi_{r,1}(t) + \phi_{r,2}(t)$ and $\phi_{r,j}(t)$ are given in (11) or (19), depending on the type of the model. Let $\widetilde{\phi}_r(t)$ be the resulting OLS parameter estimators. Consider also the errors

$$\widetilde{\eta}_t = \widehat{Y}_t - \widetilde{\phi}_1(t)\widehat{Y}_{t-1} - \ldots - \widetilde{\phi}_p(t)\widehat{Y}_{t-p}. \tag{25}$$

The model parameter $\sigma(t)^2$ and its components σ_0^2, $\sigma_1^{(2)}(t)$, $\sigma_2^{(2)}(t)$ could then be estimated analogously to the mean $\mu(t)$ and its three components as in Sect. 3.1 but replacing X_t with $\widetilde{\eta}_t^2$. We shall refer to $\widetilde{\eta}_t/\widetilde{\sigma}(t)$ as the residuals from the OLS estimation.

Remark 6. There are several potential issues with the suggested estimation of $\sigma(t)^2$ that, in particular, are encountered in the application in Sect. 5. When T/s_2 is small (e.g. $T/s_2 = 6$ in the application considered below) and $\widetilde{\sigma}(t)^2$ is computed as the s_2-periodic sample mean, note that the estimation of each $\sigma(t)^2$ involves just T/s_2 error terms (e.g. 6 in the application below). The quality of estimation of $\sigma(t)^2$ is then dubious, and we try to rectify this by slightly smoothing the estimates over time. This procedure does have some minor effect on the estimates and their standard errors, and might call for further investigation in the future. (We do not perform smoothing when estimating the mean $\mu(t)$ since we expect these estimates to be already quite smooth.) On the other hand, for Model B, we also note that the suggested procedure is not guaranteed to yield nonnegative estimates of $\sigma(t)^2$, which also happens in our application. In this case, we use the estimates of $\sigma(t)^2$ obtained for Model A.

3.3 WLS Estimation

Having the OLS estimate $\widetilde{\sigma}(t)^2$ of the variance of the error terms, the model parameters could be reestimated by using the weighted least squares (WLS) as

$$\left\{\widehat{\phi}_{r,0}, \widehat{a}^{(j)}_{r,m_j}, \widehat{b}^{(j)}_{r,m_j}\right\}_{r=1,\ldots,p,m_j=1,\ldots,H_j,j=1,2}$$
$$= \underset{\phi_{r,0},,a^{(j)}_{r,m_j},b^{(j)}_{r,m_j}}{\operatorname{argmin}} \sum_t (\widehat{Y}_t - \phi_1(t)\widehat{Y}_{t-1} - \ldots - \phi_p(t)\widehat{Y}_{t-p})^2/\widetilde{\sigma}(t)^2. \tag{26}$$

Likewise, the variance $\sigma(t)^2$ could be reestimated as $\widehat{\sigma}(t)^2$ by using the model errors based on the WLS estimates (and this process could be iterated till convergence occurs), with possible modifications discussed in Remark 6 above. Letting $\widehat{\eta}_t$ be the error terms from the WLS estimation, defined similarly to (25), the WLS residuals are defined as $\widehat{\eta}_t/\widehat{\sigma}(t)$.

4 Inference and Other Tasks

In the implementation of the OLS and WLS estimation, a PAR(p) model is expressed in the form of a linear regression as

$$Y = R\alpha + Z. \tag{27}$$

For example, for an A–PAR(p) model, $Y = (\widehat{Y}_{p+1}, \ldots, \widehat{Y}_T)'$ is a $(T-p)$–vector of periodically demeaned observations $\widehat{Y}_t$, $\alpha = (\alpha_1' \ldots \alpha_p')'$ is a $((1+2H_1+2H_2)p)$–vector of parameters with

$$\alpha_r = (\phi_{r,0}, a_{r,1}^{(1)}, \ldots, a_{r,H_1}^{(1)}, b_{r,1}^{(1)}, \ldots, b_{r,H_1}^{(1)}, a_{r,1}^{(2)}, \ldots, a_{r,H_2}^{(2)}, b_{r,1}^{(2)}, \ldots, b_{r,H_2}^{(2)})',$$

the regressors R can be expressed as a $(T-p) \times ((1+2H_1+2H_2)p)$ matrix $(R_{p+1} \ldots R_T)'$ with $R_t = (I_p \otimes B_t)Y_{t,lags}$, $Y_{t,lags} = (Y_{t-1}, \ldots, Y_{t-p})'$,

$$B_t = \Big(1, \cos(\frac{2\pi t}{s_1}), \ldots, \cos(\frac{2\pi H_1 t}{s_1}), \sin(\frac{2\pi t}{s_1}), \ldots, \sin(\frac{2\pi H_1 t}{s_1}),$$
$$\cos(\frac{2\pi t}{s_2}), \ldots, \cos(\frac{2\pi H_2 t}{s_2}), \sin(\frac{2\pi t}{s_2}), \ldots, \sin(\frac{2\pi H_2 t}{s_2})\Big)'$$

and Z refers to the error terms. Within the linear formulation (27), the OLS and WLS parameter estimators and their standard errors have well-known expressions in terms of R and Y, which we use here as well but omit for the shortness sake.

In addition to the OLS and WLS estimation as outlined above, we also use their counterparts when some of the coefficients are set to 0. We shall refer to the corresponding models as *restricted* PAR models. Estimation and computing standard errors for restricted PAR models are carried out in a standard way by expressing zero constraints through

$$\alpha = C\gamma, \tag{28}$$

where γ is a k–vector of non-zero coefficients and C is a $((1+2H_1+2H_2)p) \times k$ restriction matrix, with rows of zeros corresponding to the zero elements of α, and rows with a single entry of 1 corresponding to non-zero elements of α. The OLS and WLS estimation and inference are then performed essentially by replacing R by RC.

If needed, model selection can be guided by some information criterion, such as BIC and AIC defined in the usual way as (-2) multiplied by the log-likelihood, with an appropriate penalty. In the data application below, we shall be guided by looking at parameter "significance" and suitable diagnostics plots of model residuals. Similarly, the introduced PAR models can be used in forecasting in a straightforward way as with standard AR models and their PAR extensions. Out-of-sample forecasting performance could also be employed as another tool for selecting a model.

Remark 7. Under mild assumptions on the residuals $\{\epsilon_t\}$ in the A–PAR and B–PAR models (with typical assumptions being the i.i.d. property and finiteness of the 4th moment), the parameter estimators $\{\widetilde{\phi}_{r,0}, \widetilde{a}^{(j)}_{r,m_j}, \widetilde{b}^{(j)}_{r,m_j}\}$ in (24) and $\{\widehat{\phi}_{r,0}, \widehat{a}^{(j)}_{r,m_j}, \widehat{b}^{(j)}_{r,m_j}\}$ in (26) (assuming the true variance $\sigma^2(t)$ is used in estimation) are expected to be asymptotically normal. Indeed, these estimators are linear transformations of the analogous PAR model parameter estimators $\{\widetilde{\phi}_r(t)\}$ and $\{\widehat{\phi}_r(t)\}$. The asymptotic normality of the latter under mild assumptions is proved in Basawa and Lund [2], Anderson and Meerschaert [1]. The analogous linear transformation argument to establish the asymptotic normality of the coefficient estimators in the Fourier representation of the parameters is also used in Tesfaye et al. [21].

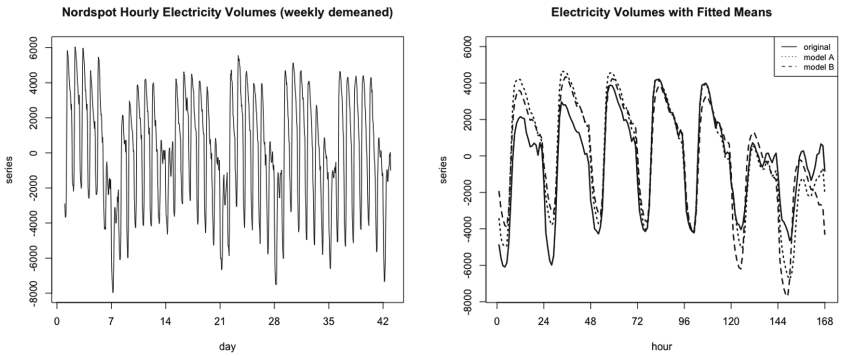


Fig. 1. Left: Weekly demeaned energy volume series for 6 weeks. Right: The volume series for week 2 with estimated means according to Models A and B.

5 Data Application

To illustrate our proposed models, we shall work with a time series of hourly electricity volumes from Nord Pool Spot Exchange.[1] This data was considered in a number of other works related to periodically correlated series, for example, Dudek et al. [7]. We consider the series for 6 weeks in 2008, and remove the weekly mean from the data. The length of the series is thus $T = 1,008$. Note that 6 weeks (1 week being the period of the underlying PAR model) are sufficient for our modeling purposes since the number of parameters is reduced considerably through the Fourier representations. For example, a small number of non-zero coefficients in the Fourier representation could be estimated, in principle, even from the data covering just one period. The resulting series is presented in Fig. 1, left plot. The right plot of the figure presents one week of the series with the mean effects estimated according to Models A and B. In the rest of the section, we shall fit Models A and B to the periodically demeaned series, that is, the difference between the observed and fitted values in Fig. 1, right plot.

[1] http://www.npspot.com.

5.1 Fitting Model A

Figure 2 depicts the periodically demeaned series according to Model A, and its sample PACF. The sample PACF suggests including lags 1, 2 and 24 into an autoregressive model. Figure 3 presents two commonly used plots to detect periodic correlations: the spectral coherence plot according to Hurd and Gerr [15] (left plot of the figure), and a related test statistic with a critical value line from Lund [18] (right plot; with a tuning parameter $M = 10$ in Lund [18]). See also Hurd and Miamee [16], Sects. 10.4 and 10.5 The spectral coherence is plotted using the R package perARMA [9].

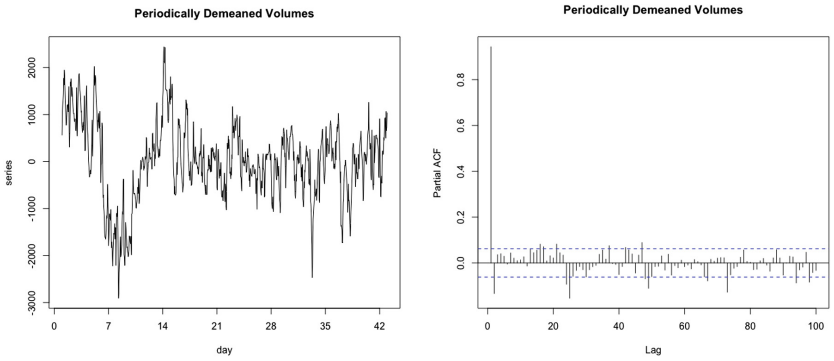


Fig. 2. Left: Periodically demeaned volume series for 6 weeks (Model A). Right: The corresponding sample PACF.

If a series exhibits periodic correlations at period s, the spectral coherence plot should have diagonal lines emerging at multiples of the index T/s. Here, $T/s = 1,008/s$. The plot in Fig. 3 suggests the first major diagonal line around the index 40. In fact, it corresponds to the period $s_1 = 24$ with $T/s_1 = 42$. There are also traces of diagonal lines at indices smaller than 42 but it is difficult to say for sure what these indices are. The latter could be determined easier from the Lund test statistic plot, which essentially averages the spectral coherence statistic at different indices along the corresponding diagonals, and also provides a critical value (the horizontal dashed line in the plot). As expected, the Lund test statistic has a large value at index 42. But note also that the values are larger, some above the critical values, at multiples of the index 6. This index corresponds to the period $s_2 = 168$ (1 week) since $T/s_2 = 6$. We thus conclude from these plots that periodic correlations are present in the periodically demeaned series at both periods $s_1 = 24$ and $s_2 = 168$.

We also see the presence of periodic correlations at the two periods $s_1 = 24$ and $s_2 = 168$ when fitting Model A. We shall report here on our fitting attempts for A–PAR(p) models of orders $p = 2$ and $p = 26$, to accomodate the partial autocorrelations seen at these lags in Fig. 2. Experimenting with various restricted A–PAR(2) models, we settled on the model with the following

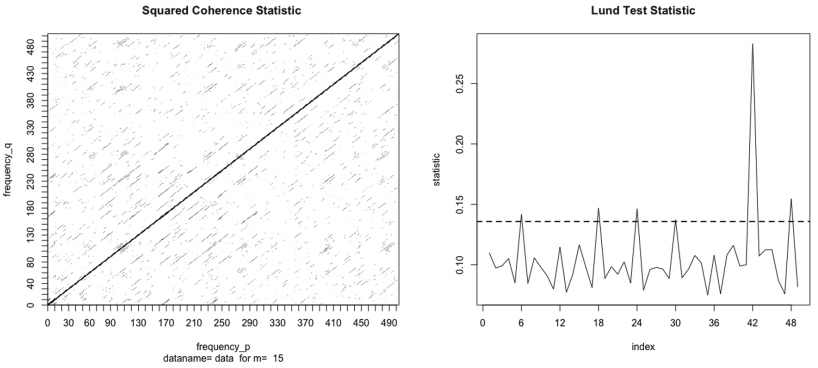


Fig. 3. Left: The spectral coherence plot for periodically demeaned volume series for 6 weeks (Model A). Right: The Lund test statistic for the same series with a horizontal dashed line indicating the critical value.

non-zero WLS estimated coefficients, with the standard errors indicated in the parentheses: at lag 1,

$$\begin{aligned}
&\widehat{\phi}_{1,0} = 1.104\,(0.025),\\
&\widehat{a}^{(1)}_{1,5} = -0.291\,(0.038), \quad \widehat{a}^{(1)}_{1,10} = -0.102\,(0.037),\\
&\widehat{b}^{(1)}_{1,7} = 0.202\,(0.036), \quad \widehat{b}^{(1)}_{1,9} = 0.081\,(0.041),\\
&\widehat{a}^{(2)}_{1,1} = 0.023\,(0.012)
\end{aligned}$$

and at lag 2,

$$\begin{aligned}
&\widehat{\phi}_{2,0} = -0.178\,(0.025),\\
&\widehat{a}^{(1)}_{2,5} = 0.245\,(0.038), \quad \widehat{a}^{(1)}_{2,10} = 0.084\,(0.037),\\
&\widehat{b}^{(1)}_{2,7} = -0.195\,(0.036), \quad \widehat{b}^{(1)}_{2,9} = -0.082\,(0.040).
\end{aligned}$$

Note that only one non-zero coefficient, namely $\widehat{a}^{(2)}_{1,1}$, is included in the component for period $s_2 = 168$. The resulting WLS estimated parameter functions $\widehat{\phi}_1(t)$ and $\widehat{\phi}_2(t)$ are plotted in Fig. 4. The component of the mean with the non-zero coefficient $\widehat{a}^{(2)}_{1,1}$ at period $s_2 = 168$ produces a "global" trend in the coefficients $\widehat{\phi}_1(t)$ over the 168 h, which is clearly visible in the left plot. Without this global trend, the coefficients can be checked to be close to what one would get from fitting a standard PAR(2) model with period $s_1 = 24$.

Figure 5 depicts the sample ACF and the Lund test statistic for the WLS residuals of the fitted A–PAR(2) model. Note some remaining autocorrelations around lag 24, which should not be surprising since we fitted a PAR model of order $p = 2$. The plot with the Lund test statistic is depicted using the same vertical scale as in Fig. 3: the peaks at dominant indices have become smaller in general but are not completely negligible.

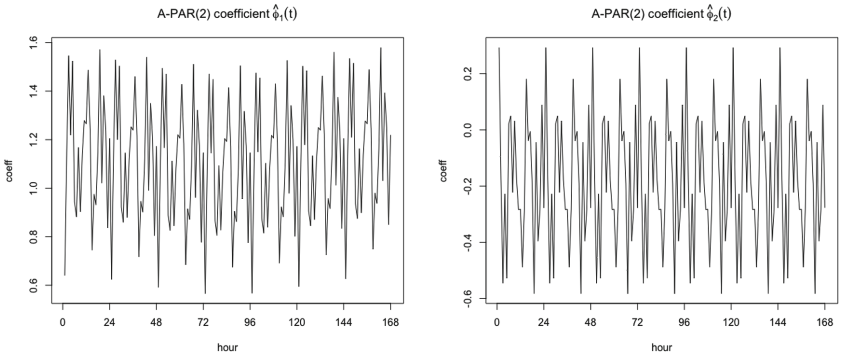


Fig. 4. The WLS estimated parameter functions $\widehat{\phi}_1(t)$ and $\widehat{\phi}_2(t)$ of the fitted A–PAR(2) model.

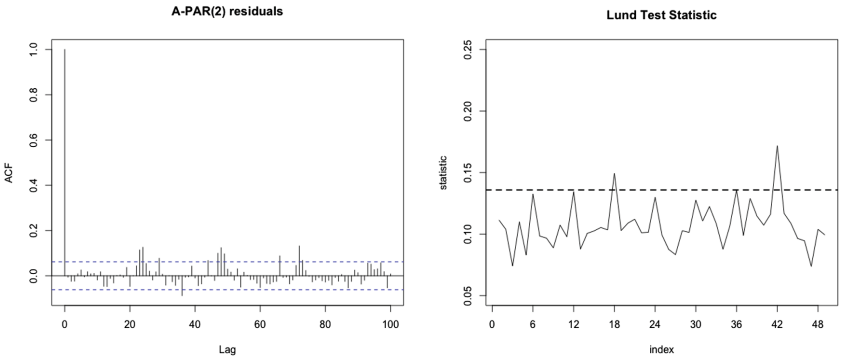


Fig. 5. The sample ACF and the Lund test statistic for the WLS residuals of the fitted A–PAR(2) model.

To remove the remaining autocorrelations in the residuals, one could fit an A–PAR(p) model of higher order p. (Another possibility would be to use a seasonal PAR model as in Basawa et al. [3].) In analogy to non-periodic seasonal models, we have experimented with fitting restricted A–PAR(26), by allowing some of the coefficients at lags 24, 25 and 26 to be non-zero. We shall not report here the fitted models but rather indicate several key observations. We found significant periodicity in the coefficients $\phi_{24}(t)$, $\phi_{25}(t)$ and $\phi_{26}(t)$, but also only in the component with period $s_1 = 24$. Typical sample ACF and Lund statistic plots for the WLS residuals of a fitted restricted A–PAR(26) are presented in Fig. 6. Note the smaller autocorrelations around multiples of lag 24 compared to those in Fig. 5. The Lund statistic plot continues having several peaks above the critical value line but their locations are no longer multiples of 6. (For example, the largest peak is no longer at 42.) It remains to clarify what might cause this shift in indices where peaks are present.

5.2 Fitting Model B

We now turn to fitting Model B and follow a similar presentation structure as for Model A in the previous section. Figure 7 presents similarly the periodically demeaned volume series according to Model B and its sample PACF. Figure 8 depicts the spectral coherence and Lund statistic plots. Note that the diagonal lines at the multiples of the indices 6 and 42 in the coherence plot, as well as the peaks at these indices in the Lund statistic plot, are much more pronounced compared to those in Fig. 3. This interesting difference is due to the way the mean effect is computed in Model B.

When fitting a B–PAR(2) model with $H_1 = 10$ and $H_2 = 3$ in the representations (20) and (21), and then reestimating it through a restricted B–PAR(2) model when including only the significant coefficients from the non-restricted model, leads to the following significant non-zero coefficients: at lag 1, $\phi_{1,0}$,

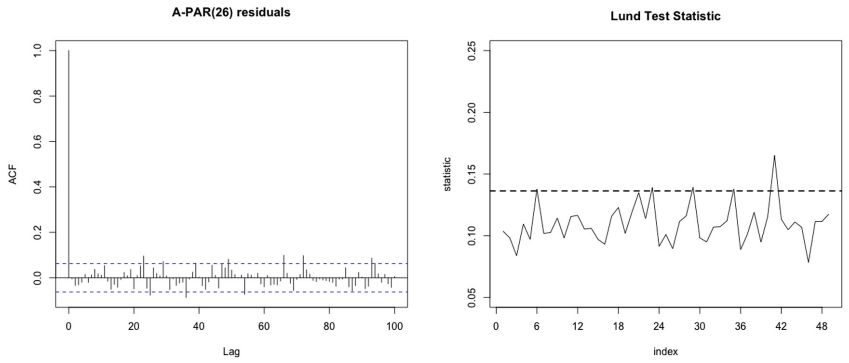


Fig. 6. The sample ACF and the Lund test statistic for the residuals of the fitted restricted A–PAR(26) model.

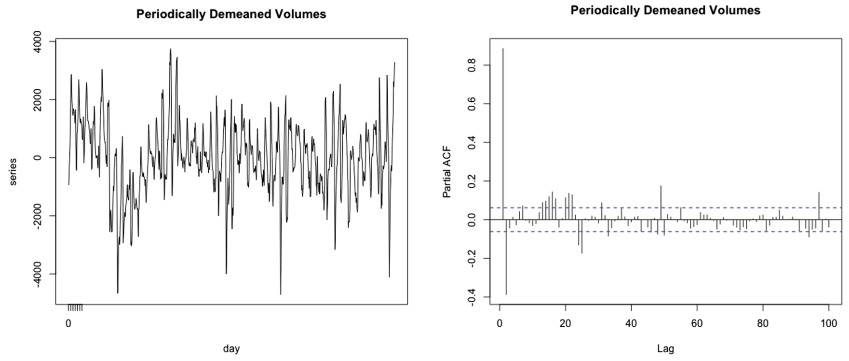


Fig. 7. Left: Periodically demeaned volume series for 6 weeks (Model B). Right: The corresponding sample PACF.

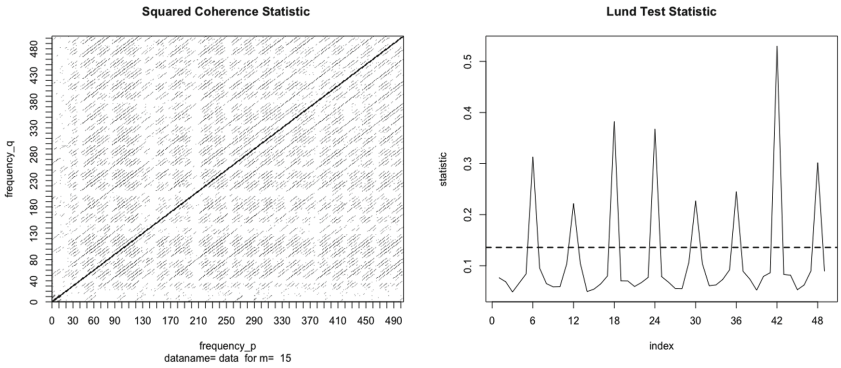


Fig. 8. Left: The spectral coherence plot for periodically demeaned volume series for 6 weeks (Model B). Right: The Lund test statistic for the same series with a horizontal dashed line indicating the critical value.

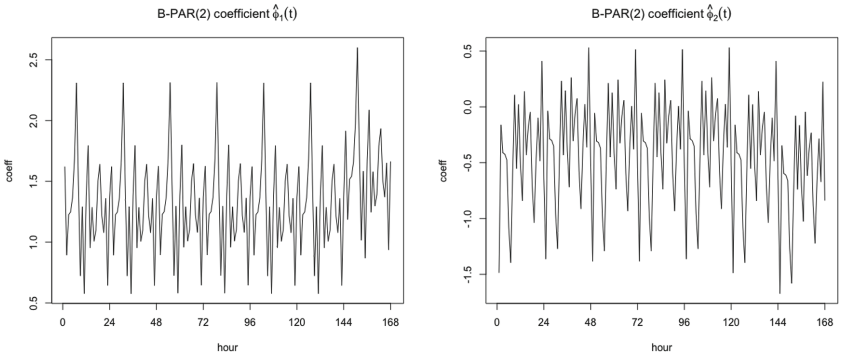


Fig. 9. The estimated parameter functions $\widehat{\phi}_1(t)$ and $\widehat{\phi}_2(t)$ of the fitted B–PAR(2) model.

$$a^{(1)}_{1,m_1} : m_1 = 2,3,6,7,8,9, \quad b^{(1)}_{1,m_1} : m_1 = 3,4,6,9,10,$$
$$a^{(2)}_{1,m_2} : m_2 = 1,3, \quad b^{(2)}_{1,m_2} : m_2 = 2,$$

and at lag 2, $\phi_{2,0}$,

$$a^{(1)}_{2,m_1} : m_1 = 1,2,3,7,8,10, \quad b^{(1)}_{2,m_1} : m_1 = 4,6,10, \quad a^{(2)}_{2,m_2} : m_2 = 1,3.$$

We shall not indicate here the values and standard errors of the corresponding WLS estimates but rather present a few revealing plots of the coefficient functions. More specifically, Fig. 10 shows the WLS estimated parameter functions $\widehat{\phi}_1(t)$ and $\widehat{\phi}_2(t)$ of the fitted B–PAR(2) model. Note that the effect of the day of a week, especially that of Sunday, is more apparent in the figure when compared to Fig. 4. This can also be seen clearer through the two components $\widehat{\phi}_{r,1}(g_1(t))$ and $\widehat{\phi}_{r,2}(g_2(t))$ depicted in Fig. 9, where the effects of the day (solid line) is

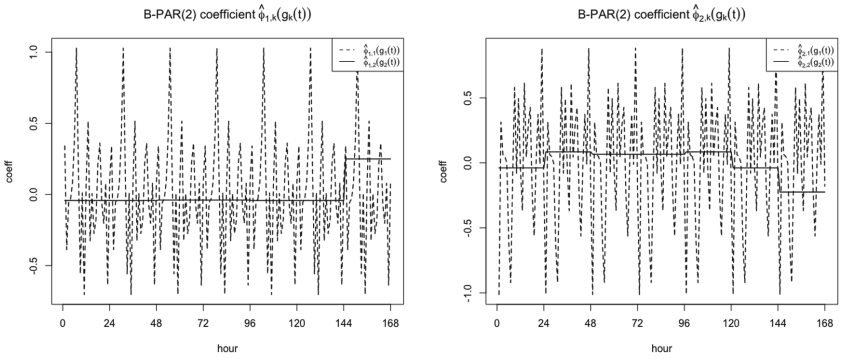


Fig. 10. The estimated parameter functions $\widehat{\phi}_{1,k}(g_k(t))$ and $\widehat{\phi}_{2,k}(g_k(t))$ of the fitted B–PAR(2) model.

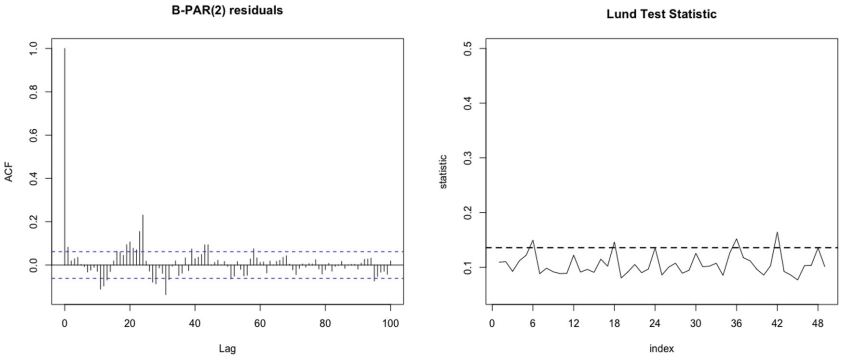


Fig. 11. The sample ACF and the Lund test statistic for the WLS residuals of the fitted B–PAR(2) model.

more pronounced towards Sunday for lag 1 and Saturday through Monday for lag 2 coefficients.

Figure 11 depicts the sample ACF and the Lund test statistic for the WLS residuals of the fitted B–PAR(2) model. The conclusions are not very different from those for the A–PAR(2) model from Fig. 5. In particular, as with Model A above, one could fit a B–PAR(p) model with higher order p to remove the remaining autocorrelations around lag 24 in the WLS residuals.

6 Conclusions

In this work, we introduced two periodic autoregression models with two or more periodic effects, discussed their inference and presented an application, showing their relevance for real data. Some of the issues that can be explored in the future include: incorporating moving average components into our models, comparing out-of-sample forecasting performance between the introduced and

among competing models, applications to other data sets, clarifying the role of the used estimation methods for error variances, and others.

Acknowledgements. The second author was supported in part by NSF Grant DMS-17-12966 at the University of North Carolina.

References

1. Anderson PL, Meerschaert MM (2005) Parameter estimation for periodically stationary time series. J Time Ser Anal 26(4):489–518
2. Basawa IV, Lund R (2001) Large sample properties of parameter estimates for periodic ARMA models. J Time Ser Anal 22(6):651–663
3. Basawa IV, Lund R, Shao Q (2004) First-order seasonal autoregressive processes with periodically varying parameters. Stat Probab Lett 67(4):299–306
4. Box GEP, Jenkins GM, Reinsel GC (1994) Time series analysis: forecasting and control, 3rd edn. Prentice Hall Inc., Englewood Cliffs
5. Dannecker L (2015) Energy time series forecasting: efficient and accurate forecasting of evolving time series from the energy domain. Springer, Heidelberg
6. De Livera AM, Hyndman RJ, Snyder RD (2011) Forecasting time series with complex seasonal patterns using exponential smoothing. J Am Stat Assoc 106(496):1513–1527
7. Dudek AE, Hurd H, Wójtowicz W (2015) PARMA models with applications in R. In: Cyclostationarity: theory and methods-II. Springer, Cham, pp 131–153
8. Dudek AE, Hurd H, Wójtowicz W (2016) Periodic autoregressive moving average methods based on Fourier representation of periodic coefficients. Wiley Interdisc Rev Comput Stat 8(3):130–149
9. Dudek AE, Hurd HL, Wójtowicz W (2016) perARMA: periodic time series analysis. R package version 1.6
10. Dudek G (2013) Forecasting time series with multiple seasonal cycles using neural networks with local learning. In: International conference on artificial intelligence and soft computing. Springer, Heidelberg, pp 52–63
11. Dudek G (2015) Generalized regression neural network for forecasting time series with multiple seasonal cycles. In: Intelligent systems 2014. Springer, Cham, pp 839–846
12. Franses PH, Paap R (2004) Periodic time series models. Advanced texts in econometrics. Oxford University Press, Oxford
13. Ghysels E, Osborn DR (2001) The econometric analysis of seasonal time series. Themes in modern econometrics. Cambridge University Press, Cambridge With a foreword by Thomas J. Sargent
14. Gould PG, Koehler AB, Ord JK, Snyder RD, Hyndman RJ, Vahid-Araghi F (2008) Forecasting time series with multiple seasonal patterns. Eur J Oper Res 191(1):207–222
15. Hurd HL, Gerr NL (1991) Graphical methods for determining the presence of periodic correlation. J Time Ser Anal 12(4):337–350
16. Hurd HL, Miamee A (2007) Periodically correlated random sequences: spectral theory and practice. Wiley series in probability and statistics. Wiley-Interscience [Wiley], Hoboken
17. Jones RH, Brelsford WM (1967) Time series with periodic structure. Biometrika 54:403–408

18. Lund RB (2011) Choosing seasonal autocovariance structures: PARMA or SARMA. In: Bell WR, Holan SH, McElroy TS (eds) Economic time series: modelling and seasonality. Chapman and Hall, London, pp 63–80
19. Lütkepohl H (2005) New introduction to multiple time series analysis. Springer, Berlin
20. Taylor JW, De Menezes LM, McSharry PE (2006) A comparison of univariate methods for forecasting electricity demand up to a day ahead. Int J Forecast 22(1):1–16
21. Tesfaye YG, Anderson PL, Meerschaert MM (2011) Asymptotic results for Fourier-PARMA time series. J Time Ser Anal 32(2):157–174
22. Weron R (2007) Modeling and forecasting electricity loads and prices: a statistical approach. Wiley, Hoboken

Subsampling for Heavy Tailed, Nonstationary and Weakly Dependent Time Series

Elżbieta Gajecka-Mirek[1(✉)] and Jacek Leśkow[2]

[1] Departament of Economics, State University of Applied Sciences in Nowy Sącz, Nowy Sącz, Poland
egajecka@gmail.com
[2] Cracow University of Technology, Kraków, Poland
jacek.leskow@pk.edu.pl

Abstract. We present new results on estimation of the periodic mean function of the periodically correlated time series that exhibits heavy tails and long memory under a weak dependence condition. In our model that is a generalization of the work of McElroy and Politis [35,42] we show that the estimator of the periodic mean function has an asymptotic distribution that depends on the degree of heavy tails and the degree of the long memory. Such an asymptotic distribution clearly poses a problem while trying to build the confidence intervals. Thus the main point of this research is to establish the consistency of one of the resampling methods - the subsampling procedure - in the considered model. We obtain such consistency under relatively mild conditions on time series at hand. The selection of the block length plays an important role in the resmapling methodology. In the article we discuss as well one of the possible ways of selection the length the subsampling window. We illustrate our results with simulated data as well as with real data set corresponding to Nord Spool data. For such data we consider practical issues of constructing the confidence band for the periodic mean function and the choice of the subsampling window.

Keywords: Periodically correlated time series · Weak dependence · Consistency of subsampling · Heavy tails · Long range dependence

1 Introduction

Our approach to model long memory is motivated by the definition of weak dependence introduced in [17] which is different from the popular $\alpha-$mixing assumption [48]. The latter involves the distributions of the time series while the former makes the reference to moments of functions of time series. While modeling nonstationary phenomena with time series techniques, it is usually much easier to verify assumptions regarding the moments than assumptions regarding the distributions.

The heavy tails that we work with will be modeled with the use of stable distributions. The stable model is quite general allowing situations when there is no variance. Such generality, however, comes with the inconvenience of not having a closed form of the pdf.

F. Chaari et al. (Eds.): CSTA 2017, ACM 16, pp. 19–40, 2020.
https://doi.org/10.1007/978-3-030-22529-2_2

Finally, the nonstationarity that we consider is of periodic type. While the formal results will be presented and developed in the subsequent sections, here we would like to make an emphasis on wide applicability of such approach. The time series and signals with periodic type of behavior were very intensively studied in the last decade. Over 2000 research papers were dedicated to such models in topics ranging from climatology to communication signals to mechanical systems monitoring (see [23]). In the context of mechanical systems, an applied study of wheel bearings diagnostics was developed (see [9]) where the presence of periodic behavior was detected together with nongaussian and heavy tail properties.

This provides a motivation to develop models where dependence structure is described via moments rather than via distributions and where heavy tails are also present. The convenient model that is combining all these features is a Gegenbauer process to be constructed in the subsequent section.

Many real life phenomena are characterized by a seasonal behavior. Seasonal data appear for example in economics, biology, climatology, telecommunications. Recently, a considerable research [12, 13] was devoted to time series with second order type periodicity, for example periodically correlated time series. For such time series we have periodicity both in the mean and in the variance and covariance. In scientific literature, such time series are referred to as periodically non stationary, periodically stationary or cyclostationary. For a review the reader can refer to Dehay and Hurd [11], Hurd et al. [32].

Very often periodicity of time series is concurrent with long range dependence [30, 45]. When seasonality is not difficult to remove, we can model the data by seasonal fractional models [20, 25]. Hui and Li in [30] consider the use of fractional differencing in modeling persistence phenomenon in a periodic process. They combine periodicity and long-memory by proposing a process consisting two independent fractional long-memory components.

The main purpose of this paper is to present results regarding estimating the mean function for a class of nonstationary time series that are heavy tailed and exhibit long memory together with periodic behavior in their mean and variance.

Our paper is organized as follows. Sections 2 and 3 contain accordingly the fundamental definitions, the model we created and tools to be used to present our results. In Sect. 3.1.1 we present main theorems that is a limit theorem for the synchronous estimator of the mean and the theorem on subsampling consistency. Section 4 contains the simulation results while Sects. 5, acknowledgments and appendix are dedicated to conclusions, acknowledgments and proofs of the theorems.

2 Fundamental Definitions and the Model

In order to better express our concept of periodicity in time series, we recall the following popular definition of stationarity of order r. For a more detailed discussion, we refer the reader to the text [31].

Definition 1 ([31], p. 3). The time series $\{X_t\}_{t\in Z}$ is called strictly stationary if for each $t_1, t_2, t_3, ..., t_n \in Z$ we have

$$(X_{t_1}, X_{t_2}, ..., X_{t_n}) \stackrel{d}{=} (X_{t_1+1}, X_{t_2+1}, ..., X_{t_n+1}).$$

Definition 2. The time series $\{X_t\}_{t\in Z}$ is called weakly stationary of order r, $(WS(r))$, if $E|X_t|^r < \infty$ and for each $t, \tau_1, \tau_2, ..., \tau_{r-1} \in Z$ and $h \in Z$,

$$E(X_t X_{t+\tau_1} ... X_{t+\tau_{r-1}}) = E(X_{t+h} X_{t+\tau_1+h} ... X_{t+\tau_{r-1}+h}).$$

Definition 3. A time series $\{X_t\}_{t\in Z}$ is called (strictly) periodically stationary PS with period T if, for every n, any collection of times $t_1, ..., t_n \in Z$, and Borel sets $A_1, ..., A_n \subset R$,

$$P_{t_1+T,...,t_n+T}(A_1, ..., A_n) = P_{t_1,...,t_n}(A_1, ..., A_n),$$

and there are no smaller values of $T > 0$ for which above equation holds. $\tau \in Z$. P is the Distribution Function of Time Series $\{X_t\}_{t\in Z}$ and is defined as follow:
let $\mathscr{Z}$ be the set of all vectors $\{\mathbf{t} = (t_1, ...t_n)' \in Z^n : t_1 < t_2 < ... < t_n, n = 1, 2, ...\}$. Then the (finite-dimensional) distribution functions of $\{X_t\}_{t\in Z}$ are the functions $\{F_t(\cdot), t \in \mathscr{Z}\}$ defined for $\mathbf{t} = (t_1, t_2, ..., t_n)'$ by

$$F_t(x) = P(X_{t_1+1} \le x_1, ..., X_{t_n+1} \le x_n), \quad x = (x_1, ..., x_n)' \in R^n.$$

Definition 4. ([31], p. 3). Time series $\{X_t\}_{t\in Z}$ is periodically correlated (PC) in the Gladyshev sense, if the mean $\mu_X(t)$ is periodic $(\mu_X(t) = \mu_X(t+T))$ and the autocovariance function $B_X(t, \tau)$ is periodic in t for all $\tau \in Z$.

There is an obvious extension of the long memory time series to the case of periodically stationary processes PS or periodically correlated processes PC.

Definition 5. A PC or PS time series $\{X_t\}_{t\in Z}$ has a long memory if the autocovariance function $\gamma(s)(h) = Cov(X_{s+qT}, X_{s+(q+h)T})$ for each $q \in Z$ satisfies the following formula

$$\sum_{0<|h|<n} \gamma(s)(h) \sim C(s) n^{\beta}, \quad s \in \{1, \ldots, T\}$$

as $n \to \infty$, where $\beta \in [0, 1)$. For each $s \in \{0, \ldots, T-1\}$ $C(s)$ is the finite constant such that

$$C(s) = 2 \cdot \lim_{n\to\infty} \frac{\sum_{h=1}^{n-1} \gamma(s)(h)}{n^{\beta}} > 0.$$

In 1999 Doukhan and Louhichi [17] and simultaneously Bickel and Bühlmann [6] introduced the concept of the weak dependence. The weak dependence is expressed in terms of covariance of functions and exactly this framework is more convenient in applications than the condition on distributions.

For the convenience of the readers, we are going to present the definition of weak dependence that is establishing a relationship between the covariance and memory for time series. It is important to note that this concept works also for non-Gaussian processes.

Let $(E, \|\cdot\|)$ be a normed space and $u \in N^*$. Here, the symbol N^* denotes the positive integers without zero. Define also a class $\mathscr{L}$ of functions in the following way:

$$\mathscr{L} = \{h : E^u \to R, \ \| h \|_\infty \le 1, Lip(h) < \infty\},$$

where $Lip(h) = \sup_{x\neq y} \frac{|h(x)-h(y)|}{\|x-y\|_1}$ and $\| x \|_1 = \sum_{i=1}^{u} \| x_i \|$. Please note that $\mathscr{L}$ is a class of bounded and Lipschitz functions defined on E^u with values on the real line. For details, we refer the reader to [10].

Definition 6. A sequence $\{X_t\}_{t\in Z}$ of random variables taking values in $E = R^d$ ($d \in N^*$) is $(\varepsilon, \mathscr{L}, \Psi)-$weakly dependent if there exists $\Psi : \mathscr{L} \times \mathscr{L} \times N^* \times N^* \to R$ and a sequence $\{\varepsilon_r\}_{r\in N}$ ($\varepsilon_r \to 0$) such that for any $(f,g) \in \mathscr{L} \times \mathscr{L}$, and $u \in N^*$, $v \in N^*$, $r \in N$

$$|Cov(f(X_{i_1},...,X_{i_u}), g(X_{j_1},...,X_{j_v}))| \leq \Psi(f,g,u,v)\varepsilon_r$$

whenever $i_1 < i_2 < \ldots < i_u \leq r + i_u \leq j_1 < j_2 < ... < j_v$.

There are several versions of weak dependence and those versions depend on the form of the function Ψ. The choice of Ψ function defines a particular type of weak dependence. For example, we call a sequence $\{X_t\}_{t\in Z}$ λ-weakly dependent, if the function Ψ is of the following form:

$$\Psi(f,g,u,v) = uvLip(f)Lip(g) + uLip(f) + vLip(g).$$

This means, that the covariance in the definition of weak mixing is dominated by a sequence ε_r tending to zero and a function that has a mixed coefficient depending on $u \cdot v$ and also a linear term in u and v.

Another examples are η-dependence and θ-dependence. We call a sequence $\{X_t\}_{t\in Z}$ η-dependent if the function Ψ is of the following form:

$$\Psi(f,g,u,v) = uLip(f) + vLip(g).$$

On the other hand, we call a sequence $\{X_t\}_{t\in Z}$ θ-dependent if the function Ψ is of the form:

$$\Psi(f,g,u,v) = uLip(f).$$

For η and θ weakly dependent sequences the covariance in the definition is dominated by a sequence ε_r tending to zero and a function that has only linear terms in u and v. For a more detailed discussion, the reader is refereed to the paper of [10].

Below we give a standard examples of weakly dependent time series.

Example 1 ([10], p. 8)**.** *Let us define the Bernoulli shift as a sequence* $X_n = H(\xi_n, \xi_{n-1}, ...)$, *where* $H(x) = \sum_{k=0}^{\infty} 2^{-(k+1)} x_k$. *Then the random sequence* X_n *is not mixing but is weakly dependent.*

Let us give the stationary representation of X_n as $X_n = \sum_{k=0}^{\infty} 2^{-(k+1)} \xi_{n-k}$. Where ξ_{n-k} is the k-th digit in the binary representation of the number coming randomly from the interval $[0,1]$.

Such X_n is a deterministic function of X_0, so the event $A = (X_0 \leq \frac{1}{2})$ belongs to the $\sigma-$algebras: $\sigma(X_t, t \leq 0)$ and $\sigma(X_t, t \geq n)$. From definition of α-mixing we get that

$$\alpha(n) \geq | P(A \cap A) - P(A)P(A) | = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}.$$

So, obviously, this sequence can not be α-mixing.

On the other hand the sequence $\{X_n\}$ is η-weakly dependent (see [10], p. 25).

It is also possible to construct AR(1) model, that will be weakly dependent but it will not be α-mixing.

Example 2 ([14], p. 3)**.** *Define AR(1) model as:* $X_t = aX_{t-1} + \varepsilon_t$, *where* $|a| < 1$ *and innovations* $\{\varepsilon_t\}$ *are i.i.d. Bernoulli variables with parameter* $p = P(\varepsilon_t = 1) = 1 - P(\varepsilon_t = 0)$. *Such AR(1) model is not* α*-mixing but is* $\eta-$*weakly dependent.*

It is relatively easy to construct stationary time series that is weakly dependent. The following result proves the point. For details, see [15].

Fact 1 ([15])**.** *Assume that* $\{\varepsilon_t\}_{t\in Z}$ *are i.i.d centered and unit variance innovations and define a linear process* X_t *as*

$$X_t = \sum_{k=0}^{\infty} b_k \varepsilon_{t-k},$$

where $k \in Z$ *and the series* b_k *is square summable. Then* X_t *is* $\theta-$*weakly dependent, where* $\theta_{2r}^2 = \sum_{k>r} b_k^2$.

For more examples of weakly dependent sequences see [10] and [19].

Another important feature of time series is long memory. The investigation of this topic was started by Lawrance and Kottegoda [39], McLeod and Hipel [43], and also by Hosking [29]. In the seminal papers of Granger and Joyeux [24] and Hosking [28] it was proposed to use the fractional differencing in modeling this kind of data.

We would like now to present a convenient model that will combine long memory with weak dependence and periodicity. Such a model is based on the Gegenbauer process (see [28]) since it can combine long memory and a seasonal behavior. For the sake of clarity, we give below the definition of the Gegenbauer process. The details can be found in [25].

Definition 7. The process $\{G_t\}_{t\in Z}$ defined by the equation:

$$\Pi_{1\le i\le k}(I - 2\nu_i B + B^2)^{d_i} G_t = \varepsilon_t, \tag{1}$$

is the k-factor Gegenbauer process. Here, $0 < d_i < 1/2$ if $|\nu_i| < 1$ or $0 < d_i < 1/4$ if $|\nu_i| = 1$ for $i = 1,...,k$. Moreover, I is the identity operator and B is the backshift operator and $\{\varepsilon_t\}_{t\in Z}$ is the Gaussian white noise.

The fact below comes from [19].

Fact 2. *Process defined by the Eq. (1) is long memory, stationary, causal and invertible and has a moving average representation:*

$$G_t = \sum_{j\ge 0} \psi_j(d,\nu)\varepsilon_{t-j},$$

with $\sum_{j=0}^{\infty} \psi_j^2(d,\nu) < \infty$, *where* $\psi_j(d,\nu), j \ge 0$, *is defined by:*

$$\psi_j(d,\nu) = \sum_{\substack{0\le l_1,...,l_n\le j \\ l_1+...+l_n=j}} C_{l_1}(d_1,\nu_1)\cdot...\cdot C_{l_k}(d_k,\nu_k),$$

where $C_{l_i}(d_i,\nu_i)$ *are the Gegenbauer polynomials defined as follows:*

$$(1 - 2\nu z + z^2)^{-d} = \sum_{j\ge 0} C_j(d,\nu) z^j,\ |z| \le 1,\ |\nu| \le 1.$$

The Gegenbauer processes are stationary, seasonal fractional models (see [25], [26]). It is enough to take $\nu = cos\varpi_t$, with $\varpi_t = 2\pi/T$, where T is the length of the season. For the sake of simplicity, in our paper we will assume that in the definition of the Gegenbauer process the parameter $k = 1$.

Having defined the Gegenbauer process that combines long memory and seasonal behavior we now will model the heavy tailed part of our data. For that, we will use the stable distributions. Let us recall the definition of the stable distribution. For details, we refer the reader to [49].

Definition 8. A random variable X has a stable distribution if there exist parameters: $0 < \alpha \leq 2,\ -1 \leq sk \leq 1,\ \sigma > 0$ and $\mu \in R$ such that the characteristic function of X has a form:

$$\varphi(t) = \begin{cases} \exp\{-\tau^{\alpha}|t|^{\alpha}(1 - i \cdot sk \cdot sgn(t)tg\frac{\pi\alpha}{2}) + i\mu t\}, & \alpha \neq 1 \\ \exp\{-\tau|t|(1 + i \cdot sk \cdot \frac{2}{\pi}sgn(t)ln|t|) + i\mu t\}, & \alpha = 1 \end{cases}$$

The index $\alpha \in (0,2]$ is frequently called the stability index, $sk \in [-1,1]$ is the skewness parameter, $\tau > 0$ is the scale parameter and $\mu \in R$ is the location parameter.

The stable variables are very convenient to use in many applications, especially in modeling heavy tailed phenomena. However, the inconvenience of working with α-stable distributions for $\alpha < 2$ is that they will not have second order moments.

2.1 The Model

We build a long-memory periodic process with period T by adjusting a long memory stationary model to the T-variate process. The model will be simultaneously dealing with three features of time series: periodicity, long memory and heavy tails.

Let the time series $\{X_t\}_{t \in Z}$ be defined as:

$$X_t = \sigma_t GG_t + \eta(t), \tag{2}$$

where

- **A1** The volatility time series σ_t and the Gaussian-Gegenbauer time series GG_t are independent
- **A2** The sequence of random variables σ_t is i.i.d coming from α-stable family.
- **A3** GG_t is periodic Gaussian-Gegenbauer time series defined as

  $$GG_t = f(t) \cdot G_t\ .$$

  In the above equality, G_t is Gaussian-Gegenbauer mean zero time series with $k = 1, \mid \nu \mid \leq 1, LM(\beta)$ with $\beta \in [0,1)$. The function $f(t)$ is periodic, deterministic, bounded with a known period T. The autocovariance of G_t is denoted as γ_G.
- **A4** The deterministic function $\eta(t)$ is periodic with the same period T as $f(t)$.

Comment 1. *Note that $\beta = 2d$, where d is the parameter from the definition of the model (2) and β is the long memory parameter from Definition 5.*

A similar model to (2) was considered by Politis and McElroy (see [42]). However, the Authors did not consider the Gegenbauer processes as the representation for long memory.

2.2 Properties of the Model

Fact 3. *Process X_t defined by Eq. (2) has long memory in the sense of the Definition 5, with $\beta \in [0,1)$.*

The proof of the Fact follows directly from the Definition 5. □

Theorem 1. *For each $s = 1,...,T$ the Gaussian - Gegenbauer process GG_{s+pT} is not strong mixing, but it has weak dependence properties. This means that X_t defined by the Eq. (2) is also weakly dependent. More precisely, the model defined by the Eq. (2) is $\lambda-$weakly dependent, for each $s = 1,...,T$.*

Proof of the above theorem is in the Appendix.

The properties above provide a clear motivation to study heavy-tailed and weakly dependent structures. In the subsection below we provide a study of volatility with heavy tails.

Define the volatility process in the model (2) as follows:

$$\sigma_t = \sqrt{\varepsilon_t},$$

where ε_t are i.i.d. $\alpha/2$-stable with skewness parameter equal to one, scale parameter $(cos(\pi\alpha/4))^{2/\alpha}$, and a location parameter $\mu = 0$. From the research of Taqqu ([49], Propositions 1.2.16 and 1.2.17), we know that $E(\sigma) = E(\sigma_t)$ is finite and bigger than zero. The construction of X_t is based on sub-Gaussian process ([49], p. 142).

We also have the following.

Theorem 2. *The marginal distributions of X_t defined by the Eq. (2) are $S\alpha S$ with the scale parameter*

$$\tau(X_t) = |f_t|\sqrt{\gamma_G(0)/2}$$

and the mean $f_t \cdot E(\sigma)$, $t = 1,...,NT$ where NT is the number of observation and T is the length of the period.

The proof of the above fact follows from [49], Propositions 1.2.3 and 1.3.1.

If the volatility in model (2) has a stable distribution then the second moment of process X_t is infinite, so X_t is not second-order. Thus the model (2) is not periodically correlated in the sense of Gladyshev, since it does not have variance. However, it is periodically stationary (*PS* for short). In such a way we have succeeded in constructing a heavy tailed, long memory, weakly dependent and not α-mixing time series that has infinite variance, finite covariances that is also periodically stationary in the sense of Definition 4. This statement is explained by the following

Comment 2. *The X_t defined by the Eq. (2) does not have variance, but it has periodic autocovariance function $\gamma(t+T,h) < \infty$ for $h \neq 0$.*

Indeed:

$$\begin{aligned}\gamma(t+T,h) &= f_{t+T}f_{t+T+h}(E\sigma)^2\gamma_G(t+T,h)\\ &= f_{t+T}f_{t+T+h}(E\sigma)^2\gamma_G(h) = f_t f_{t+h}(E\sigma)^2\gamma_G(t,h) = \gamma(t,h).\end{aligned}$$

The following section will be dedicated to statistical inference in our model (2). We will focus on constructing the estimator of the mean function η_t, proving the central limit theorem for the estimator under weak dependence, heavy tails and long memory. Finally, we will prove that subsampling of such estimator is consistent.

3 Main Results

3.1 The Estimator and Its Properties

One of the main objectives of this paper is to study the subsampling procedure for the estimator of the mean function $\eta(s)$ in model (2). In order to do that, we will have to define the estimator $\hat{\eta}_N(s)$ for $s = 1,\dots,T$. We will also have to study the asymptotic properties of such estimator. Let us assume, that we have a given sample $(X_1, X_2, X_3, ..., X_n)$ from a series $\{X_n\}$ with properties such as (2). Without a loss of generality, let us assume that our sample size n is equal $n = NT$, where T is the known period. This essentially means, that we assume that the sample size is proportional to the length T of the period. This notation reflects how many periods are there in n observations.

We start with the definition.

Definition 9. We define the estimator $\hat{\eta}_N(s)$ of $\eta(s)$ as:

$$\hat{\eta}_N(s) = \frac{1}{N}\sum_{p=0}^{N-1} X_{s+pT}, \;\; s = 1,2,\dots,T, \tag{3}$$

where T is the known period.

Define

$$\zeta = max\{1/\alpha, (\beta+1)/2\},$$

where α is the heavy tails parameter and β is a long memory parameter.

Let us also define $PCF_N(s)$ - the modified sample partial autocorrelation function that estimates the long memory (see also [42]):

Definition 10

$$PCF_N(s) = |\sum_{|h|=1}^{[N^{1/2}]} \frac{1}{N-|h|}\sum_{p=0}^{N-|h|}(X_{s+pT}X_{s+pT+hT} - \hat{\eta}_N^2(s))|^2, \tag{4}$$

where $s = 1,2,\dots,T$.

The above defined function (4) plays an important role in formulating the limiting distribution of the estimator (3).

We are ready now to formulate the central limit theorem for the estimator (3).

3.1.1 Central Limit Theorem

We need to define the following sequences of random variables:

$$A_N(s) = N^{1-\zeta}(\hat{\eta}_N(s) - \eta(s)),$$

$$B_N(s) = N^{-2\zeta}\sum_{p=0}^{N-1}(X_{s+pT} - \hat{\eta}_N(s))^2,$$

$$C_N(s) = N^{-2\zeta+1}PCF_N(s).$$

Recall, that in the above definitions $s \in \{1,\dots,T\}$ where T is the length of the period.

Theorem 3. *Assume* **A1** *through* **A4** *from the definition of the model (2).*
Then the following joint weak convergence holds:

$$\big(A_N(s),B_N(s),C_N(s)\big) \stackrel{d}{\Rightarrow} \begin{cases} (S(s),U(s),0), & if\ 1/\alpha > (\beta+1)/2 \\ (V(s),0,\mu^4 C(s)^2), & if\ 1/\alpha < (\beta+1)/2 \\ (S(s)+V(s),U(s),(E(\sigma))^4 C(s)^2), & if\ 1/\alpha = (\beta+1)/2. \end{cases} \tag{5}$$

The symbol $\stackrel{d}{\Rightarrow}$ denotes the convergence in distribution.

To have a more convenient way of expressing the central limit theorem we will introduce the sequence $P_N(s)$ of random variables defined as:

$$P_N(s) = \frac{A_N(s)}{\sqrt{B_N(s)+C_N(s)}}.$$

Define also a random variable Q by

$$Q = \begin{cases} S(s)/\sqrt{U(s)}, & if\ 1/\alpha > (\beta+1)/2 \\ V(s)/\mu^2 C(s), & if\ 1/\alpha < (\beta+1)/2 \\ (S(s)+V(s))/\sqrt{U(s)+(E(\sigma))^4 C(s)^2}, & if\ 1/\alpha = (\beta+1)/2. \end{cases}$$

Recall, that the function $C(s)$ was defined in Definition 5. We have the following

Corollary 1. *Assume* **A1** *through* **A4** *from the definition of the model (2).*
Then

$$P_N(s) \stackrel{d}{\Rightarrow} Q.$$

In the above convergence of P_N we have that $S(s)$ is a $S\alpha S$ variable with zero location parameter, and scale $|f(s)|\sqrt{\gamma_G(0)/2}$, $s = 1,\ldots,T$. Moreover, $V(s)$ is a mean zero Gaussian variable with variance $\tilde{C}(s)\mu^2/(\beta+1)$, where $\tilde{C}(s) = |f(s)|^2 \cdot (C(s) - \gamma_G(0) I_{\{\beta=0\}})$. Additionally, we have that $S(s)$ and $V(s)$ are independent.

The random variables $U(s)$, $s = 1,\ldots,T$ are $\alpha/2$ stable with zero location parameter, skewness, and scale proportional to $f_s^2\gamma_G(0)$. The random variables $V(s)$ are independent of $U(s)$, but $S(s)$ and $U(s)$ are dependent.

Proofs of the Theorems 3 and 1 are in Appendix.

Let us summarize the results of the section above. The Eq. (2) in the beginning of this section is introducing the equation of our model. It is the Gegenbauer process scaled by heavy tails volatility time series. We have considered the cases of heavy tails behavior: the case of α-stable distributions. The fundamental assumptions necessary for further work are **A1** through **A4** listed under the model equation (2). The natural estimator of the unknown periodic mean function (2) is introduced in (Sect. 3.1.1). For this estimator we have a central limit theorem in which we clearly see the impact of the long memory assumption and heavy tails. It is quite straightforward to see that it is quite difficult to get the asymptotic confidence intervals for $\eta(s)$ (see (2)) due to a very complicated form of the asymptotic distribution. Therefore, the problem of finding confidence bands for our unknown mean function $\eta(s)$ can be solved using the concept of subsampling. In the next section we will introduce that concept and we will prove that it provides quantiles that are asymptotically valid which means that subsampling is consistent.

3.2 Subsampling

Subsampling is one of the resampling methods available for time series. From the research of Hall and Lahiri [33,38] we know that for the long-range dependent time series subsampling provides consistent procedure. In the presence of the long-range dependence, however, many other known resampling procedures (e.g. Moving Block Bootstrap, see [37]) fail to provide reliable results.

Let us assume, that we have sample $(X_1, X_2, X_3, ..., X_n)$ from a series $\{X_n\}$ with properties such as (2). Recall that our sample size n is equal $n = NT$. We will use the following form of subsampling:

- **Step 1**
  For each $s = 1, \ldots, T$ the estimator $\hat{\eta}_N(s)$ is recomputed from the $(X_s, \ldots, X_{s+(N-1)T})$ over "short" overlapping blocks of length b_s . The block length b_s depends on the length of the sample as well as on $s \in \{1, \ldots, T\}$.
- **Step 2**
  From Step 1 $N - b_s + 1$ statistics are obtained for each s. In our context those will be $a_{b_s}(\hat{\eta}_{N,b_s,i}(s) - \hat{\eta}_N(s))$ where $\hat{\eta}_{N,b_s,i}(s)$ is subsamplig version of the estimator $\hat{\eta}_N(s)$ and a_{b_s} is the normalizing sequence.
- **Step 3**
  We calculate the empirical distributions:

$$L_{N,b_s}(s)(x) = \frac{1}{N - b_s + 1} \sum_{i=1}^{N-b_s+1} I_{\{a_{b_s}(\hat{\eta}_{N,b_s,i}(s) - \hat{\eta}_N(s)) \le x\}}.$$

  We use $L_{N,b_s}(s)(x)$ to approximate the asymptotic distribution $L(s)(x)$ of the estimator $a_N(\hat{\eta}_N(s) - \eta(s))$.

The fundamental problem with subsampling procedure is its consistency. In particular, we need to prove that finite sample quantiles generated by the subsampling procedure converge asymptotically ($N \to \infty$) to the quantiles of the asymptotic distribution.

The following subsection presents our central result - the consistency of subsampling.

3.3 Consistency of the Subsampling Method

To prove the consistency of the subsampling procedure we usually start by studying if there exists a non-degenerated asymptotic distribution $L(s)(x)$ of the statistic $L_N(s)(x)$. It is not necessary to know the form of the asymptotic distribution. In our case this is already done since we have already established the central limit theorem so $L_N(s)(x)$ is the distribution of the normalized estimator $A_N(s)$ defined just before the statement of the central limit theorem. Therefore, the conclusion from the central limit theorems is that the empirical distribution functions converge weakly to the cumulative distribution functions of the limit random variables

$$L_N(s)(x) \to L(s)(x) \qquad if \qquad N \to \infty, \qquad s = 1, \ldots, T.$$

Denote the density of the limit distribution by $L'(s)$. We know that for stable distributions we have $\| L'(s) \|_\infty < \infty$ (for details see [40]).

We are now ready to formulate the subsampling consistency theorem.

Theorem 4. *Assume* **A1** *through* **A4**. *The consistency of the subsampling method holds, which means in particular that:*

1. *If x is the point of the continuity of $L(s)$, then $L_{N,b_s}(s)(x) \xrightarrow{p} L(s)(x)$.*
2. *If L is continuous then $\sup_x |L_{N,b_s}(s)(x) - L(s)(x)| \xrightarrow{p} 0$.*
3. *If $L(s)$ is continuous in $c(1-q)$ (where $(c(1-q)$ is a $q-$quantile) then if $N \to \infty$*

$$P\{N^{1-\zeta} \sum_{p=0}^{N-1} (\hat{\eta}_N(s) - \eta(s)) < c_{N,b_s}(1-q)\} \to 1-q.$$

where $c_{N,b_s}(1-q)$ is the $(1-q)$-quantile of the $L_{N,b_s}(x,s)$.
Proof can be found in the Appendix.

4 Simulation Study

4.1 Choosing b_s

One need to be careful in choosing b_s. It cannot be too small of course, but also it can't be too big otherwise the subsamlpling will not work (see [4]).

We remind that

$$L_{N,b_s}(x) = \frac{1}{N-b_s+1} \sum_{i=1}^{N-b_s+1} I_{\{a_{b_s} P_{b_s} \le x\}},$$

$$L_N(x) = \frac{1}{N} \sum_{i=1}^{N} I_{\{a_N(s) P_N(s) \le x\}}$$

and

$$L(x) = P(a_N P_N \le x)).$$

Let

$$\bar{L}_{b_s}(x) = \frac{1}{N-b_s+1} \sum_{i=1}^{N-b_s+1} I_{\{P_{b_s} \le x\}}$$

$$\tilde{L}_{b_s}(x) = P(P_{b_s} \le x)$$

Lemma 1. *If the sequence X_t is stationary and weak dependent, then $\bar{L}_b(x) = \tilde{L}_b(x) + o_P(1), N \to \infty$.*

The above Lemma is a simple corollary from the consistency of the subsampling method theorem.

Lemma 2. *Let $k_0 = sup\{x : L(x) = 0\}$ and $k_1 = inf\{x : L(x) = 1\}$ and assume that $L(x)$ is continuous and strictly increasing on (k_0, k_1) as a function of x. If the consistency of the subsampling method theorem is fulfilled then*

$$a_b \bar{L}_b^{-1}(x) = L^{-1}(x) + o_P(1), \tag{6}$$

for any $x \in (0,1)$ and $N \to \infty$.

Proof can be found in Appendix.

In order to choose the block length we will use a simple empirical tool proposed by Bertail [4].

Assuming that $a_N = N^{-\zeta}$ and taking logarithm in the Eq. (6) we get:

$$log(|\bar{L}_{b_s}^{-1}(x)|) = log(L^{-1}(x)) + \zeta log(b_s) + o_P(1).$$

If we take any $p_i \neq p_j \in (0,1)$ and draw log of some quantile range of subsampling distribution

$$|\bar{L}_{b_s}^{-1}(p_i) - \bar{L}_{b_s}^{-1}(p_j)| = \zeta log(b_s) + |L^{-1}(p_i) - L^{-1}(p_j)| + o_P(1)$$

we will get the diagram based on which we can make the best choice of b. The best choice of b (see [4]) is the largest one before the "unstable" behavior begins. We provide an example on Fig. 1. The R program was used to obtain the simulations.

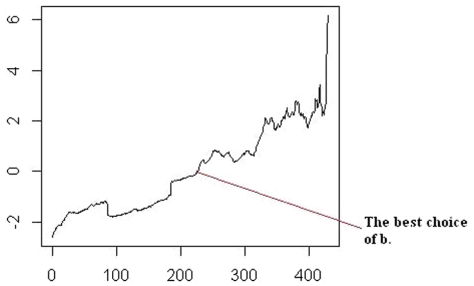


Fig. 1. The example of choosing b for the process X_{s+p+T} with parameters $\alpha = 1.5$, $\beta = 0.3$ and $s = 5$, $T = 24$.

In the sequel, on the Figs. 2, 3, 4 and 5, the simulation study for the mean is introduced. We assume that the mean function η is periodic function $\eta(t) = 2cos(\pi \cdot t/24)$, for $t \in Z$.

For each s = $1, \ldots, 24$, we found subsample size b_s by the method described by Bertail [4] and then draw the equal-tailed and symmetric 95% confidence intervals. The R program was used to obtain the simulations.

For the simulation study we chose the Gaussian Gegenbauer process with k = 1, innovations with mean zero and variance 1, $\nu = 1$ In this case the autocorrelation function is equal as follow [49]:

$$\gamma_{GG}(h) = \frac{\Gamma(1-\beta)}{\Gamma(\beta/2)\Gamma(1-\beta/2)} h^{\beta-1}$$

$$\gamma_{GG}(0) = \frac{\Gamma(1-\beta)}{\Gamma^2(1-\beta/2)}.$$

The constant C in the definition of long memory for each $s = 1, \ldots, T$ is:

$$C(s) = \mu^2 f_s^2 \frac{\Gamma(1-\beta)}{\beta/2\Gamma(\beta/2)\Gamma(1-\beta/2)}.$$

For the ε_t we chose $\alpha/2-$stable iid random variables with the skewness parameter 1, the location parameter 0 and the scale $(cos(\pi\alpha/4))^{2/\alpha}$.

The number of observations is $NT = 10320$, period $T = 24$.

The two-sided $1-q$ equal-tailed confidence interval for $\eta(t)$ were constructed as follows:

$$\left[\hat{\eta}_N(s) - \frac{\hat{\sigma}_N(s)}{\sqrt{N}} c_{N,b_s}(1-q/2), \hat{\eta}_N(s) + \frac{\hat{\sigma}_N(s)}{\sqrt{N}} c_{N,b_s}(q/2)\right], \tag{7}$$

where $c_{N,b_s}(1-q)$ is the $(1-q)$-quantile of the $L_{N,b_s}(s)(x)$ and

$$\hat{\sigma}_N^2(s) = \frac{1}{N}\sum_{p=0}^{N-1}(X_{s+pT} - \hat{\eta}_N(s))^2 + PCF_N(s).$$

The two-sided symmetric subsampling interval for $\eta(t)$ was constructed as follows:

$$\left[\hat{\eta}_N(s) - \frac{\hat{\sigma}_N(s)}{\sqrt{N}} c_{N,b_s,|\cdot|}(1-q/2), \hat{\eta}_N(s) + \frac{\hat{\sigma}_N(s)}{\sqrt{N}} c_{N,b_s,|\cdot|}(1-q/2)\right], \tag{8}$$

where

$$c_{N,b_s,|\cdot|}(1-q/2) = inf\{x : \frac{1}{N-b_s+1}\sum_{i=1}^{N-b_s+1} I_{(|P_{N,b_s}(s)| \le x)} \ge 1-q\}.$$

In the first case we took $\beta = 0.3$ and $\alpha = 1.5$. This is the "tail" case.

For each s = 1, …, 24, we found subsample size by the method proposed by Bertail [4]. We chose the minimum of found b and then draw the equal-tailed and symmetric 95% confidence intervals (Figs. 2, 3).

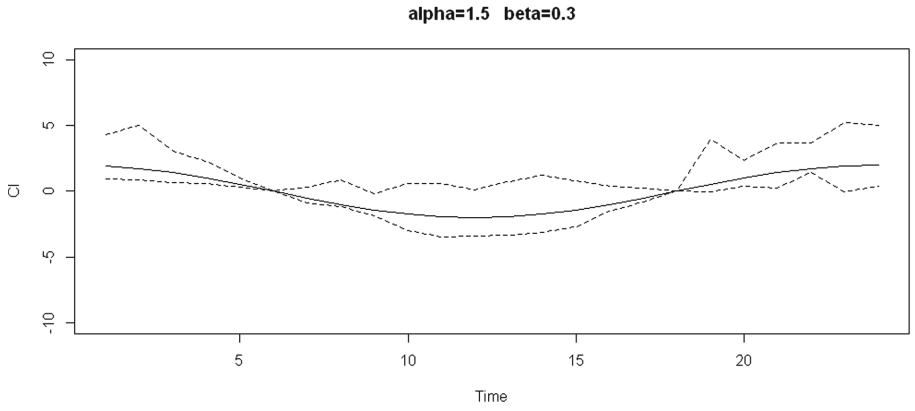


Fig. 2. Example of equal tailed confidence interval for the mean parameter of the process X_{s+p+T} with parameters $\alpha = 1.5$, $\beta = 0.3$ and $T = 24$.

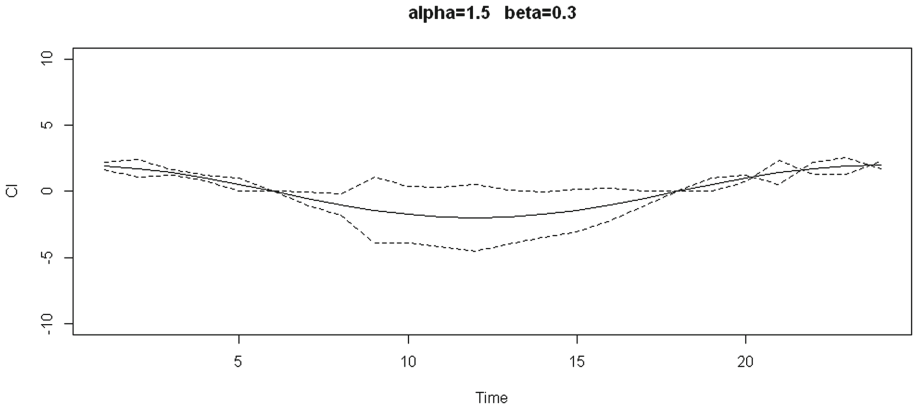


Fig. 3. Example of symmetric confidence intervals for the mean parameter of the process X_{s+p+T} with parameters $\alpha = 1.5$, $\beta = 0.3$ and $T = 24$.

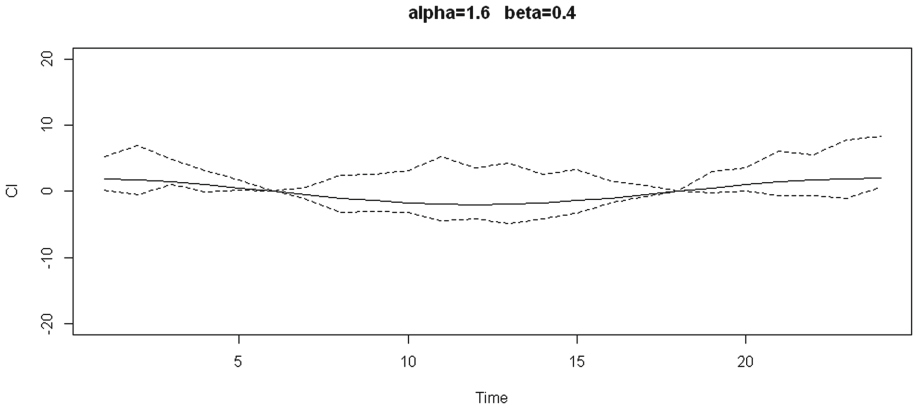


Fig. 4. Example of equal-tail confidence interval for the mean parameter of the process X_{s+p+T} with parameters $\alpha = 1.6$, $\beta = 0.4$ and $T = 24$.

In the second case we took $\beta = 0.4$ and $\alpha = 1.6$. This is the "memory" case. For each $s = 1, \ldots, 24$, we have done the same as in previous case (Figs. 4, 5).

4.2 Real Data Application

We consider the prices from the Nord Pool electricity data facility from January 1, 2000 to January 30, 2014 and obtain 5144 daily observations [44]. The Fig. 6 shows the data. The daily prices are computed as the average of 24 hourly prices.

We apply presented in the article subsampling to construct confidence interval for periodic mean of the Nord Pool electricity time series [44]. Data consists 5144 daily observations from January 1, 2000 to January 30, 2014. Daily prices were obtained as the average of 24 hourly prices. The series considered in the article was also considered as a series with a long memory and heavy tailed by Proietti, Haldrup and Knapik in [47].

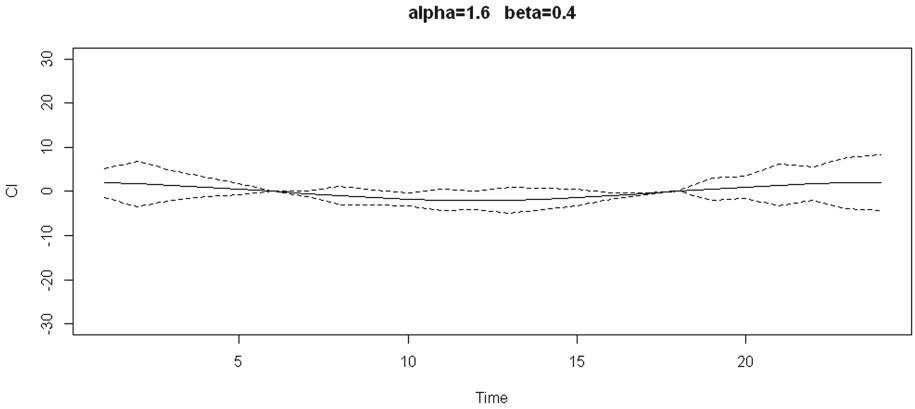


Fig. 5. Example of symmetric confidence intervals for the mean parameter of the process X_{s+p+T} with parameters $\alpha = 1.6$, $\beta = 0.4$ and $T = 24$.

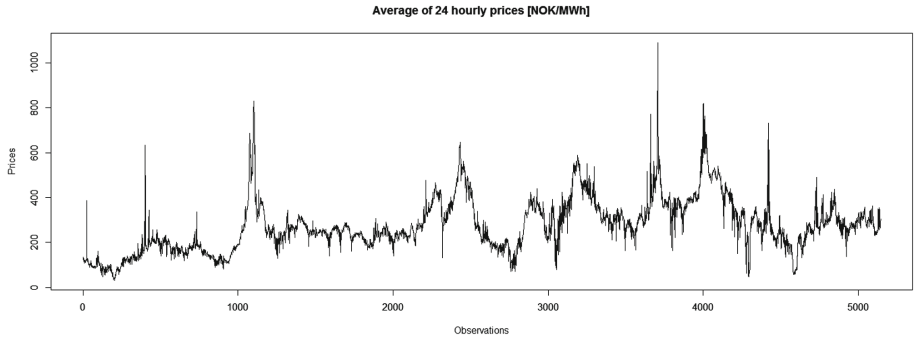


Fig. 6. Daily price for Nord Pool: Prices in Norwegian kroner (NOK) per MWh. The period of observation: January 1, 2000–January 30, 2014; number of observations is 5144

Our time series were split into seven blocks according to 7 days period of length 734 each. (It is possible to use a period such as monthly or quarterly, but seven days period seems to be the most visible.)

95% confidence interval for the means was calculated in each block.

As already mentioned the period in the time series is equal to $T = 7$. For each s = $1, \ldots, 7$, we calculated $P_N(s)$. The size of b was calculated by the method proposed by Bertail [4] and is 200.

Construction of confidence intervals is as in Eqs. (7) and (8). Note that the construction of $P_N(s)$ does not require knowledge of parameters of memory - β and tails - α. It is possible to estimate parameters α and β, [27]. $\alpha = 1.66$, $\beta = 0.96$. Then we drew confidence intervals for $\eta_N(s)$.

We can see that the mean changes over time in the 7 day period (Fig. 7).

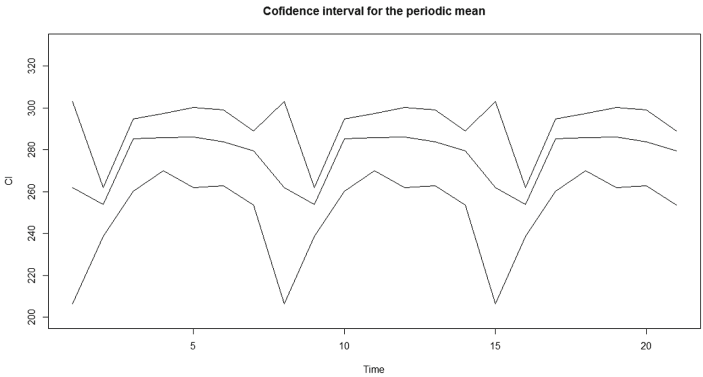


Fig. 7. The confidence interval for daily price for Nord Pool; The period of observation: January 1, 2000– January 30, 2014. The scale on vertical axle is 100.

5 Conclusions

Our results provide a relatively simple and practical algorithm of estimating the mean together with the confidence intervals in the case of heavy tailed, weakly dependent, long memory and periodically correlated time series. The method of subsampling provides a versatile tool to create confidence intervals without the hassle of exact calculation of parameters of the asymptotic distributions of the estimator of the periodic mean function. In the forthcoming work, we plan to apply our model to data where other heavy tailed distributions can be introduced, for example GED distributions. In such case, we will have an additional advantage of having all finite moments. For the estimator defined by the Eq. 3.1.1 and when the volatility process σ_t is from the GED family one can formulate the analogous limit theorem as 4. In such case, subsampling consistency theorem could provide another versatile tool to create confidence intervals.

Acknowledgment. Both Authors would like to express their enormous gratitude to Professor Paul Doukhan for his invaluable help while preparing this manuscript. The Author Leśkow is supported by the grant no 2013/10/M/ST1/00096 from Polish National Center for Science NCN.

Appendix

In the appendix the proofs of some results are presented.
Proof of Fact 1

The $\{GG_{s+pT}\}_{s\in Z}$, for each $s = 1,\ldots,T$ is a second ordered stationary process.
Assume that the process $\{GG_{s+pT}\}$ satisfies

- $\alpha-$mixing condition
- completely regular condition and
- completely linear regular condition

For Gaussian processes the relationship between above coefficient is as follows [36]: $\rho(k) = r(k)$, and $\alpha(k) \le r(k) \le 2\pi\alpha(k)$. From the theorem below the Gaussian - Gegenbauer process can't be strong mixing. [19] Since $\{\varepsilon_t\}_{t\in Z}$ in Eq. (1) is a Gaussian process, the long memory stationary k-factor Gegenbauer process (1) is not completely regular and hence is not strong mixing. From [15] we know that the Gaussian - Gegenbauer process has weak dependence properties. □

Proof of Theorem 3

The proof of the Theorem 3 is in [22], here only main points will be repeated.

Let: $N(\hat{\eta}_N(s) - \eta(s)) = \sum_{p=0}^{N-1} Y_{s+pT}$, $s = 1,2,...,T$ where $Y_t = X_t - \eta(t) = \sigma_t G G_t$.

First we need to show that $N^{-\zeta}\sum_{p=0}^{N-1} Y_{s+pT}$ converge weakly to some random variable.

Let $\mathscr{E}$ be the $\sigma-$field: $\mathscr{E} = \sigma(\varepsilon) = \sigma(\varepsilon_t, t \in Z)$. Let $\mathscr{G}$ be the $\sigma-$field: $\mathscr{G} = \sigma(GG) = \sigma(GG_t, t \in Z)$. From the assumption A in the definition of the model (2) the $\sigma-$fields $\mathscr{E}$ and $\mathscr{G}$ are independent with respect to the probability measure P. The properties of the characteristic function of normal variable sum are used in the proof.

$$E exp\{i\nu N^{-1/\alpha}\sum_{p=0}^{N-1} Y_{s+pT}\} = E[E[exp\{i\nu N^{-1/\alpha}\sum_{p=0}^{N-1}\sigma_{s+pT}GG_{s+pT}\}|\mathscr{E}]]$$

where ν is any real number and $s = 1,2,...,T$.

The inner conditional characteristic function, from the properties of Gaussian characteristic function is

$$E[exp\{i\nu N^{-1/\alpha}\sum_{p=0}^{N-1}\sigma_{s+pT}GG_{s+pT}\}|\mathscr{E}]$$
$$= exp\{-\frac{(\nu N^{-1/\alpha})^2}{2}\sum_{p,q=0}^{N-1}\sigma_{s+pT}\sigma_{s+qT}f_{s+pT}f_{s+qT}\gamma_N(T(p-q))\},\ s = 1,...,T.$$

This double sum is divided into the diagonal and off-diagonal terms:

$$N^{-\frac{2}{\alpha}}\Big(\sum_{p=0}^{N-1}\sigma_{s+pT}^2\gamma_G(0) + \sum_{p\neq q}\sigma_{s+pT}\sigma_{s+qT}\gamma_G((p-q)T)\Big) \tag{9}$$

In the case $1/\alpha > (\beta+1)/2$ the off-diagonal part of (9) is $O_P(N^{1-2/\alpha}N^{\beta})$ which tends to zero as $N \to \infty$. The characteristic function of the diagonal part of (9) is the characteristic function of a $S\alpha S$ variable with scale $\sqrt{\gamma_G(0)/2} = |f_s|\sqrt{\gamma_N(0)/2}$, for $s = 1,2,...,T$. (see [42]).

In the case $1/\alpha < (\beta+1)/2$ the formula (9) becomes

$$N^{-(\beta+1)}\Big(\sum_{p=0}^{N-1}\sigma_{s+pT}^2\gamma_G(0) + \sum_{q\neq p}\sigma_{s+pT}\sigma_{s+qT}\gamma_G((p-q)T)\Big). \tag{10}$$

The first term is $O_P(N^{2/\alpha-(\beta+1)})$ and tends to zero as $N \to \infty$.

The limiting characteristic function of the second part is characteristic function of a mean zero Gaussian with variance $\tilde{C}(s) = f_t(C(s) - \gamma_G(0)I_{\{\beta=0\}})$.

The case $1/\alpha = (\beta+1)/2$ is the combination of the two above cases. From the Slutsky's Theorem we get weak convergence the sum of two independent random variables.

The next step is to show joint convergence of the first and second moment of the model (2), for each $s = 1,..,N$ and $p = 0,...,N-1$.

In the proof the joint Fourier/Laplace Transform of the first and second sample moments [21] is considered.

For any real θ and $\phi > 0$,

$$Eexp\{i\theta N^{-\zeta}\sum_{p=0}^{N-1} Y_{s+pT} - \phi N^{-2\zeta}\sum_{p=0}^{N-1} Y_{s+pT}^2\}$$
$$= E[exp\{-\frac{1}{2}N^{-\zeta}\sum_{p.q}^{N-1}\sigma_{s+pT}\sigma_{s+qT}\gamma_G((p-q)T)(\theta+\sqrt{2\phi}W_{s+pT})(\theta+\sqrt{2\phi}W_{s+qT})\}].$$

The sequence of random variables W_s is i.i.d. standard normal, and is independent of the Y_s series. The information about W_s is denoted by $\mathscr{W}$. The double sum in the Fourier/Laplace Transform is broken into diagonal and off-diagonal terms.

The off-diagonal term is

$$N^{-2\zeta}\sum_{|h|>0}^{N-1}\sum_{p=0}^{N-|h|}\sigma_{s+pT}\sigma_{s+pT+hT}(\theta+\sqrt{2\phi}W_{s+pT})(\theta+\sqrt{2\phi}W_{s+pT+hT})\gamma_G(hT) \quad (11)$$

In the case $2/\alpha > \beta+1$, by the Markov inequality the off-diagonal term tends to zero in probability as $N \to \infty$.

In the case $2/\alpha < \beta+1$ (11) tends to constant (see the proof of Theorem 1 [42]).

In the case $2/\alpha = \beta+1$, the off-diagonal part, for fixed $s = 1,2,...,T$, tends to a constant.

The diagonal term is examined separately (by Dominated Convergence Theorem and above fact). Let $V_{s+pT} = \theta + \sqrt{2\phi}W_{s+pT}$

$$E[exp\{-\frac{1}{2}\gamma_G(0)N^{-2\zeta}\sum_{p=0}^{N-1}\sigma_{s+pT}^2 V_{s+pT}^2\}]$$
$$= E[exp\{-(\gamma_G(0)/2)^{\alpha/2}N^{-\alpha\zeta}\sum_{p=0}^{N-1}|V_{s+pT}|^\alpha\}|.$$

While $2/\alpha < \beta+1$ the sum $N^{-\alpha\zeta}\sum_{p=0}^{N-1}|V_{s+pT}|^\alpha \xrightarrow{p} 0$.

While $2/\alpha \geq \beta+1$ (by the Law of Large Numbers) $N^{-\alpha\zeta}\sum_{p=0}^{N-1}|V_{s+pT}|^\alpha \xrightarrow{p} E|V|^\alpha$. By the Dominated Convergence Theorem, the limit as $N \longrightarrow \infty$ can be taken through the expectation, so that $E[exp\{-(\gamma_G(0)/2)^{\alpha/2}N^{-\alpha\zeta}\sum_{p=0}^{N-1}|V_{s+pT}|^\alpha\}] \to exp\{-(\gamma_G(0)/2)^{\alpha/2}E|\theta+\sqrt{2\phi}N|^\alpha 1_{2/\alpha\geq\beta+1}\}$. Using the Fourier/Laplace Transform and argumentation as in [42] we get the proof of this part of the Theorem 3.

Using argumentation as in [42] and in [35] we obtain for $s = 1, ..., T$

$$N^{-\beta\rho}(\widehat{LM}(\rho, s))^{\rho} = o_P(1)$$
$$+\left|N^{-\beta\rho}\sum_{|h|>0}^{[N^{\rho}]}\frac{1}{N-|h|}\sum_{p=0}^{N-|h|}Y_{s+pT}Y_{s+pT+hT}\right| \xrightarrow{P} \mu^2 C(s).$$

At last from the Slutsky Theorem we get the convergence 5.

The second convergence in (5) follows from the continuous mapping theorem (denominators are different from zero). □

Proof of Theorem 4

Let us consider a sequence of statistics $P_N(s)$, for fixed $s = 1, 2, ..., T$ and $N = 1, 2, ...$ $L_N(s)(x) = P(Z_N(s) \leq x)$ is cumulative distribution function of $P_N(s)$.

There exist

$$r_N(s) = sup_{x\in R}|L_N(s)(x) - L(s)(x)| \longrightarrow 0, \quad N \to \infty$$

For overlapping samples the number of subsamples:

$Y_{b_s,p}(s) = (X_{N,s+pT}, X_{s+(p+1)T}..., X_{N,s+(p+b_s-1)T})$, $p = 0, 1, ..., N - b_s$ and the number of subsampling statistics:

$P_{N,b_s,p}(s) = \sqrt{b_s}(\hat{\eta}_{N,b_s,p}(s) - \hat{\eta}_N(s))/\hat{\sigma}_{N,b_s,p}(s)$ is $N - b_s + 1$.

Above statistics are used to approximate the distributions $L_N(s)(x)$ by empirical distribution functions: $L_{N,b_s,p}(s)(x) = \frac{1}{N-b+1}\sum_{p=0}^{N-b_s} I_{\{P_{N,b_s,p}(s)\leq x\}}$.

For $A_N(s)$ let define subsampled statistics:

$$U_{N,b_s,p}(s)(x) = \frac{1}{N-b+1}\sum_{p=0}^{N-b}\varphi(\frac{A_N(s)-x}{\varepsilon_n}).$$

The sequence ε_n is decreasing to zero and φ is the non-increasing continuous function such that $\varphi = 1$ or 0 according to $x \leq 0$ or $x \geq 1$ and which is affine between 0 and 1.

It is enough to investigate the variance of $U_{N,b_s,p}(s)$, $s = 1, ..., T$ (Theorem 11.3.1 [46]).

$$VarU_{N,b_s,p}(s)(x) = (N - b_s + 1)^{-2}(\sum_{|h|<N-b_s+1}(N - b_s + 1 - |h|)\gamma(h))$$

here $\gamma(h) = Cov(\varphi(A_N(s)), \varphi(A_N(s)))$.

From the assumption that we have $\lambda-$weak dependency

$$Cov(\varphi(A_{N,p}(s)), \varphi(A_{N,p+h}(s))) \leq \sqrt{b_s}\lambda_{h-b+1}/\varepsilon_N$$

For $\varepsilon_N = (N^{2(1-\beta)b})^{1/4}$ above covariance converge to zero. From Cesaro mean argument $VarU_{N,b_s,p}(s)(x)$ also goes to zero.

$$V_{N,b_s,p}(s)(x) = \frac{1}{N-b+1}\sum_{p=0}^{N-b}\varphi(\frac{A_N(s)P_{N,b_s,p}(s)-x}{\varepsilon_n}).$$

Under condition **A1** through **A4** and the Theorem 3.1, [50] we have:

$$lim_{N\to\infty}|E[V_{N,b_s,p}(s)(x) - E[V_{N,b_s,p}(s)(x)]]^2| = 0.$$

It implies that $Var(V_{N,b_s,p}(s)(x))$ tends to zero, it proves point 1. of the Theorem 4.

To prove the point 2. of the Theorem 4 we also use the Theorem 2 in [14].

$$lim_{N\to\infty}sup_{x\in R}|U_{N,b_s,p}(s)(x) - L(s)(x)| = 0,$$

in probability. The proof of point 3. if the 1. holds and under assumption of the model (2) is the same as the proof of 3. in the Theorem 11.3.1 [46]. □

Proof of the Lemma 2

The proof of the Lemma 2 strictly follows from Lemma 2, [5] and Theorem 2, [14]. □

References

1. Arcones M (1994) Limit theorems for nonlinear functionals of a stationary sequence of vectors. Ann Prob 22:2242–2274
2. Bardet J-M, Doukhan P, Lang G, Ragache N (2008) Dependent Lindeberg central limit theorem and some applications. ESAIM: Probab Stat 12:154–172
3. Beran J, Feng T, Ghosh S, Kulik R (2013) Long-memory processes. Probabilistic Properties and Statistical Methods. Springer, Heidelberg
4. Bertail P (2011) Comments on "Subsampling weakly dependent time series and application to extremes". TEST 20(3):487–490
5. Bertail P, Haefke C, Politis DN, White W (2004) Subsampling the distribution of diverging statistics with applications to finance. J Econ 120:295–326
6. Bickel P, Bühlmann P.: A new mixing notion and functional central limit theorems for a sieve bootstrap in time series, Bernoulli, 5-3, pp. 413–446 (1999)
7. Bradley R (2005) Basic properties of strong mixing conditions. Probab Surv 2:107–144
8. Brockwell P, Davis R (1991) Time series: theory and methods. Springer, New York
9. Cioch W, Knapik O, Leśkow J (2013) Finding a frequency signature for a cyclostationary signal with applications to wheel bearing diagnostics. Mech Syst Signal Process 38:55–64
10. Dedecker J, Doukhan P, Lang G, Léon JR, Louhichi S, Prieur C (2008) Weak dependence: with examples and applications. Lecture Notes in Statistics. Springer, New York
11. Dehay D, Hurd HL (1994) Representation and estimation for periodically and almost periodically correlated random processes. In: IEEE Cyclostationarity in communications and signal processing, New York, pp 295–326
12. Dehay D, Dudek A, Leśkow J (2014) Subsampling for continuous-time almost periodically correlated processes. J Stat Plan Inference 150:142–158
13. Dudek A, Leśkow J, Paparoditis E, Politis D (2013) A generalized block bootstrap for seasonal time series. J Time Ser Anal 35:89–114
14. Doukhan P, Prohl S, Robert CY (2011) Subsampling weakly dependent times series and application to extremes. TEST 20(3):487–490
15. Doukhan P, Lang G (2002) Rates in the empirical central limit theorem for stationary weakly dependent random fields. Statis Inf Stoch Porc 5:199–228 MR1917292
16. Doukhan P (1994) Mixing: properties and examples, vol 85. Lecture Notes in Statistics. Springer, New York

17. Doukhan P, Louhichi S (1999) A new weak dependence condition and applications to moment inequalities. Stoch Proc Appl 84:313–342
18. Evans M, Hastings N, Peacock B (2000) Statistical distributions, 3rd edn. Wiley, New York, pp 74–76
19. Guégan D, Ladoucette S (2001) Non-mixing properties of long memory sequences, C.R. Acad. Sci. Paris, 333, pp 373–376
20. Ferrara L, Guégan D (2001) Forecasting with k-factor Gegenbauer processes: theory and applications. J Forecast 20(8):581–601
21. Fitzsimmons P, McElroy T (2006) On joint Fourier-Laplace transforms, Mimeo. http://www.math.ucsd.edu/politis/PAPER/FL.pdf
22. Gajecka-Mirek E (2014) Cyclostationarity: Theory and Methods. Subsampling for weakly dependent and periodically correlated sequences. LNME. Springer, Cham. https://doi.org/10.1007/978-3-319-04187-2_4
23. Gardner WA, Napolitano A, Paura L (2006) Cyclostationarity: half a century of research. Signal Process. 86:639–697
24. Granger CWJ, Joyeux J (1980) An introduction to long memory time series models and fractional differencing. J. Time Series Anal. 1:15–30
25. Gray HL, Zhang N-F, Woodward WA (1989) On generalized fractional processes. J Time Ser Anal 10:233–257
26. Gray HL, Zhang N-F, Woodward WA (1994) On generalized fractional processes - a correction. J Time Ser Anal 15(5):561–562
27. Harmantzis F, Hatzinakos D (2001) Network traffic modeling and simulation using stable FARIMA processes. Technical report. Presented at INFORMS Annual Meeting, Miami
28. Hosking JRM (1981) Fractional differencing. Biometrika 68:165–176
29. Hosking JRM (1982) Some models of persistence in time series. In: Anderson OD (ed) Time series analysis: theory and practice, vol 1. North-Holland, Amsterdam
30. Hui YV, Li WK (1995) On Fractionally differenced periodic processes. Indian J Stat Ser B 57(1):19–31
31. Hurd HL, Miamee AG (2007) Periodically correlated random sequences: spectral theory and practice. Wiley, Hoboken
32. Hurd H, Makagon A, Miamee AG (2002) On AR(1) models with periodic and almost periodic coefficients. Stoch Process Appl 100:167–185
33. Hall P, Jing B, Lahiri S (1998) On the sampling window method for long-range dependent data. Stat Sin 8:1189–1204
34. Ibragimov IA, Rozanov YA (1978) Gaussian random processes. Springer, New York
35. Jach A, McElroy T, Politis DN (2012) Subsampling inference for the mean of heavy-tailed long memory time series. J Time Ser Anal 33(1):96–111
36. Kolmogorov AN, Rozanov YA (1960) On strong mixing conditions for stationary Gaussian processes. Theory Prob Appl 5:204–208
37. Künsch H (1989) The jackknife and the bootstrap for general stationary observations. Ann Statist 17:1217–1241
38. Lahiri S (1993) On the moving block bootstrap under long-range dependence. Stat Probab Lett 18:405–413
39. Lawrance AJ (1977) Kottegoda N T (1977) Stochastic modeling of river flow time series. J. Royal Stast Soc Ser B 140:1–47
40. Logan BF, Mallows CL, Rice SO, Shepp LA (1973) Limit distributions of self-normalized sums. Ann Probab 1:788–809
41. Marinucci D (2005) The empirical process for bivariate sequences with long memory. Statist Inf Stoch Proc 8(2):205–223
42. McElroy T, Politis DN (2007) Self-normalization for heavy-tailed time series with long memory. Stat Sin 17(1):199–220

43. McLeod AI, Hipel KW (1978) Preservation of the rescaled adjusted range, I a reassessment of the Hurst phenomenon. Water Resour Res 14:491–508
44. NordPool (2016) Nord pool spot. http://www.nordpoolspot.com/. Accessed 15 July 2017
45. Philippe A, Surgalis D, Viano MC (2006) Almost periodically correlated processes with long-memory, vol 187. Lecture Notes in Statistics. Springer, New York
46. Politis DN, Romano JP, Wolf M (1999) Subsampling. Springer, New York
47. Proietti T, Haldrup N, Knapik O (2017) Spikes and memory in (Nord Pool) electricity price spot prices, CREATES Research Paper 2017-39
48. Rosenblatt M (1956) A central limit theorem and a strong mixing condition. Proc Nat Acad Sci USA 42:43–47
49. Taqqu MS, Samorodntsky G (1994) Stable non-Gaussian random processes. Stochastic models with infinite variance. Chapman and Hall, New York
50. van der Vaart AW (1998) Asymptotic Statistics. Cambridge University Press, Cambridge

Bootstrapping the Autocovariance of PC Time Series - A Simulation Study

Anna E. Dudek[1,2(✉)] and Paweł Potorski[1]

[1] Department of Applied Mathematics, AGH University of Science and Technology, al. Mickiewicza 30, 30-059 Krakow, Poland
aedudek@agh.edu.pl

[2] Institut de Recherche Mathématique de Rennes, Université Rennes 2, Rennes, France

Abstract. In this paper a simulation comparison of the bootstrap confidence intervals for the coefficients of the autocovariance function of a periodically correlated time series is provided. Two bootstrap methods are used: the circular version of the Extension of Moving Block Bootstrap and the circular version of the Generalized Seasonal Block Bootstrap. The bootstrap pointwise and simultaneous confidence intervals for the real and the imaginary parts of the Fourier coefficients of the autocovariance function are constructed. The actual coverage probabilities, the average lengths and the average upper and lower quantiles values are calculated. On the basis of the performed simulation, the choice of the block length that is an integer multiple of the period length is advised when the Moving Block Bootstrap is used. Moreover, two methods for the block length choice designed for stationary data: the Minimum Volatility Method and the approach based on the logarithm of quantile are verified not to be valid for periodic nonstationary case. Finally, a heuristic method of the block length choice is proposed.

1 Introduction

Studies of periodically correlated (PC) processes were started in 1961 by Gladyshev [14]. Time series $\{X_t, t \in \mathbb{Z}\}$ with finite second moments is called PC with the period d, if it has periodic mean and covariance function, i.e.,

$$\mathrm{E}\left(X_{t+d}\right) = \mathrm{E}\left(X_t\right), \mathrm{Cov}\left(X_t, X_s\right) = \mathrm{Cov}\left(X_{t+d}, X_{s+d}\right).$$

For more details concerning PC processes we refer the reader to [15].

Over the last 60 years the theory of PC processes has developed fast and found applications in many branches of vibroacoustics, mechanics, signal analysis, hydrology, climatology and econometrics (e.g., energy markets). Many motivating examples can be found in [1,13,15,20,23]. The wide range of possible applications resulted in thousands of papers. Unfortunately, the analysis of PC processes is very difficult. One of the major problems is the estimation of the asymptotic covariance matrix for parameters of interest. In practice it is almost impossible and hence to construct confidence intervals resampling methods need to be used.

F. Chaari et al. (Eds.): CSTA 2017, ACM 16, pp. 41–55, 2020.
https://doi.org/10.1007/978-3-030-22529-2_3

The idea of resampling techniques is to approximate the distribution of the statistics of interest. Nowadays, bootstrap is the most popular resampling method. It was introduced in 1979 by Efron [11]. At the beginning it was designed for i.i.d. data. In such cases the bootstrap sample is created by random sampling with the replacement of observations from the original sample. However, this approach cannot be applied for dependent data. In 1986 Carlstein [3] proposed to select randomly blocks of observations. This allows us to preserve the dependence structure of the data inside the blocks. If the data are weakly dependent, i.e., observations that are far from each other are almost independent, this approach provides consistent estimators for many characteristics of stationary and also nonstationary time series. The number of block bootstrap techniques is constantly growing. New methods or modifications of existing techniques appear to provide estimators with lower bias, better rates of convergence or that mimic some specific structure of the data such as periodicity. In this paper we consider two block methods: the Extension of Moving Block Bootstrap (EMBB) and the Generalized Seasonal Block Bootstrap (GSBB). The EMBB is a modification of Carlstein's approach based on the Moving Block Bootstrap (MBB) method. It allows us to select any block of observations from the sample, while Carlstein's idea was to choose only non-overlapping blocks. On the other hand, the GSBB is designed for periodic data. To preserve periodicity during the block selection process, the set of blocks is restricted and varies in each step of the algorithm. The detailed description of the EMBB and the GSBB is presented in Sect. 2.

In this paper we focus on the second-order analysis of a periodically correlated time series. Proper detection of so-called second-order frequencies is crucial in many applications like e.g., diagnostics of rotating machines [6,7] and the analysis of the human walk [10]. These problems are of great importance and mistakes can be very costly, for example a large machine in a factory might be needlessly stopped for repair. Thus, we decided to perform a study that may help practitioners to choose the optimal bootstrap approach. Recently there appeared a few theoretical results for constructing bootstrap consistent pointwise and simultaneous confidence intervals for the coefficients of the autocovariance function of a PC time series. Unfortunately, the provided tools are not complete. There is no method for choosing the block length. Moreover, there is no study comparing properties of different approaches.

For our work we decided to use computational resources available in the Polish Grid Infrastructure PL-Grid. That allowed us to consider a wide range of block length values, from small to very large sample sizes, and different period lengths. In the sequel we try to provide answers for often posed questions like:

- does using the EMBB have any advantage over the GSBB;
- do some existing methods of block length choice for stationary time series also apply for our nonstationary case;
- does preserving periodic structure by the GSBB improve, compared with the EMBB, the coverage of the confidence intervals;
- how are the coverage probabilities affected by non-optimal block length choices;
- can the same block length be used to construct all pointwise confidence intervals for the considered frequencies;
- are simultaneous confidence intervals less sensitive to non-optimal block length choice than pointwise ones.

This paper is organized as follows. In Sect. 2 the necessary notation and definitions are introduced. Moreover, block bootstrap algorithms and bootstrap consistency results are recalled. Section 3 is dedicated to the results of the simulation study. The performance of the two considered block bootstrap methods is compared. Different sample sizes and block lengths are used. Pointwise and simultaneous confidence intervals for the coefficients of the autocovariance function are constructed. Additionally, the problem of the block length choice is discussed and a new heuristic approach is proposed. In Sect. 4 a short summary of the results is provided.

2 Problem Formulation

Let $\{X_t, t \in \mathbb{Z}\}$ be a PC time series with period d. From now on we assume that $\mathrm{E}(X_t) \equiv 0$. Moreover, let $B(t,\tau) = \mathrm{Cov}(X_t, X_{t+\tau})$ be its autocovariance function. The variables t and τ represent time and shift, respectively. Note that $B(t,\tau)$ is periodic in t. Its analysis in the frequency domain is performed using the following Fourier decomposition

$$B(t,\tau) = \sum_{\lambda \in \Lambda_\tau} a(\lambda,\tau)\exp(i\lambda t),$$

where $\Lambda_\tau = \{\lambda : a(\lambda,\tau) \neq 0\}$. The set Λ_τ is finite and is a subset of the set $\overline{\Lambda} = \{2k\pi/d, k = 0,\ldots,d-1\}$. This means that there is a finite number of the second-order significant frequencies. To detect them it is enough to point out the nonzero Fourier coefficients of $B(t,\tau)$. For that purpose one needs to construct confidence intervals for $a(\lambda,\tau)$. Below we recall results from literature to show how difficult this task is.

Let $X_1,\ldots,X_n$ be a sample from the considered time series. Without loss of generality we assume that $\tau \geq 0$. An estimator of $a(\lambda,\tau)$ is of the form

$$\widehat{a}_n(\lambda,\tau) = \frac{1}{n}\sum_{t=1}^{n-\tau} X_t X_{t+\tau}\exp(-i\lambda t).$$

Moreover, the asymptotic results that we present below require the mixing assumptions. To be precise X_t is assumed to be α-mixing i.e., $\alpha_X(k) \to 0$ as $k \to \infty$, where

$$\alpha_X(k) = \sup_t \sup_{\substack{A \in \mathscr{F}_X(-\infty,t) \\ B \in \mathscr{F}_X(t+k,\infty)}} |P(A \cap B) - P(A)P(B)|$$

and $\mathscr{F}_X(-\infty,t) = \sigma(\{X_s : s \leq t\})$, $\mathscr{F}_X(t+k,\infty) = \sigma(\{X_s : s \geq t+k\})$. If $\alpha_X(k) = 0$ it means that the observations that are k time units apart are independent. In general mixing assumptions are introduced to ensure that observations that are far from each other are 'almost' independent. An easy example of α-mixing process is a m-dependent time series.

Let us introduce some additional notation. By λ and τ we denote r-dimensional vectors of frequencies and shifts of the form $\lambda = (\lambda_1,\ldots,\lambda_r)'$, $\tau = (\tau_1,\ldots,\tau_r)'$. Additionally, $a(\lambda,\tau) = \big(\Re(a(\lambda_1,\tau_1)), \Im(a(\lambda_1,\tau_1)),\ldots,\Re(a(\lambda_r,\tau_r)), \Im(a(\lambda_r,\tau_r))\big)'$ and $\widehat{a}_n(\lambda,\tau)$ is its estimator. By $\Re(z)$ and $\Im(z)$ we denote the real and the imaginary part of the complex number z.

Theorem 1 below states asymptotic normality of $\widehat{a}_n(\lambda,\tau)$. For the proof we refer the reader to [26] (Theorem 2.6) and [18] (Theorem 1).

Theorem 1. *Assume that* $\{X_t, t \in \mathbb{Z}\}$ *is a PC* α*-mixing time series with* $\mathrm{E}(X_t) \equiv 0$ *and* $\mathrm{WP}(4)$ *such that:*

(i) $\sup_{t\in\mathbb{Z}} \mathrm{E}|X_t|^{4+2\delta} < \infty$ *for some* $\delta > 0$,
(ii) $\sum_{\tau=1}^{\infty} \tau \alpha_X^{\delta/(2+\delta)}(\tau) < \infty$.

Then

$$\sqrt{n}\left(\widehat{a}_n(\lambda,\tau) - a(\lambda,\tau)\right) \xrightarrow{d} N_{2r}\left(0, \Sigma(\lambda,\tau)\right),$$

where $\Sigma(\lambda,\tau) = \left[\sigma_{ef}\right]_{e,f=1,\ldots,2r}$ *and*

$$\sigma_{ef} = \begin{cases} \frac{1}{d}\sum_{s=1}^{d}\sum_{k=-\infty}^{\infty} C(s,k)\cos(\lambda_e s)\cos(\lambda_f(s+k)) & \text{for} \quad e = 2g_1 - 1, f = 2g_2 - 1, \\ \frac{1}{d}\sum_{s=1}^{d}\sum_{k=-\infty}^{\infty} C(s,k)\sin(\lambda_e s)\sin(\lambda_f(s+k)) & \text{for} \quad e = 2g_1, f = 2g_2, \\ -\frac{1}{d}\sum_{s=1}^{d}\sum_{k=-\infty}^{\infty} C(s,k)\sin(\lambda_e s)\cos(\lambda_f(s+k)) & \text{for} \quad e = 2g_1, f = 2g_2 - 1, \\ -\frac{1}{d}\sum_{s=1}^{d}\sum_{k=-\infty}^{\infty} C(s,k)\cos(\lambda_e s)\sin(\lambda_f(s+k)) & \text{for} \quad e = 2g_1 - 1, f = 2g_2, \end{cases}$$

with $C(s,k) = \mathrm{Cov}\left(X_s X_{s+\tau_e}, X_{s+k} X_{s+k+\tau_f}\right)$, $g_1, g_2 = 1, \ldots, r$ *and* $\det(\Sigma(\lambda,\tau)) \neq 0$.

$\mathrm{WP}(k)$ denotes weakly periodic times series of order k with period d. In other words time series is $\mathrm{WP}(k)$ when its k-th moments are periodic.

Theorem 1 enables construction of the asymptotic pointwise confidence intervals for $a(\lambda,\tau)$. Unfortunately, the asymptotic covariance matrix $\Sigma(\lambda,\tau)$ is very difficult to estimate. Thus, in practice resampling methods are often used to approximate the quantiles of the asymptotic distribution. Additionally, so-called percentile bootstrap confidence intervals do not require estimation of the covariance matrix (see e.g., [12]). Since PC time series are an example of the nonstationary processes, block bootstrap methods need to be used for them to keep the dependence structure contained in the data. So far consistency of two bootstrap techniques was shown for $a(\lambda,\tau)$. These are the EMBB and the GSBB. The EMBB was proposed in [7] and [8]. It is the modification of the MBB. The MBB was introduced independently in [16] and [19] and was designed for stationary time series. However, it turned out that it can be successfully applied in some nonstationary cases ([4,5,25]). The main disadvantage of the MBB and its modifications like the EMBB in the context of PC time series, is the fact that these methods do not preserve the periodic structure of the original data. On the other hand, the GSBB was designed for periodic processes. The method was proposed by Dudek et al. in [9]. Applied for a PC time series it perfectly retains the periodic structure contained in the data.

In this paper we will use a modification of the EMBB and the GSBB, whose idea is to consider data as wrapped on the circle. It was introduced in [21] to reduce edge effects caused by the MBB and it was called the Circular Block Bootstrap (CBB). The CBB guarantees that each observation is present in the same number of blocks which is not the case of the MBB, the EMBB and the GSBB. Below we recall the algorithms of the circular versions of the EMBB (cEMBB) and the GSBB (cGSBB).

Let B_i, $i = 1, \ldots, n$ be the block of observations from our sample $X_1, \ldots, X_n$, that starts with observation X_i and has the length $b \in \mathbb{N}$, i.e., $B_i = (X_i, \ldots, X_{i+b-1})$. If $i+b-1 > n$ then the missing part of the block is taken from the beginning of the sample and we get $B_i = (X_i, \ldots, X_n, X_1, \ldots, X_{b-n+i-1})$ for $i = n-b+2, \ldots, n$. To apply the cEMBB and the cGSBB to a PC data, we need to assume that the sample size n is an integer multiple of the period length d ($n = wd$). If $wd < n \leq (w+1)d$ then the observations $X_i, i = wd+1, \ldots, n$ need to be removed from the considered dataset.

cEMBB Algorithm

Since the cEMBB approach is based on the CBB algorithm, as first we recall the CBB method.

1. Choose a block size $b < n$. Then our sample can be divided into l blocks of length b and the remaining part is of length r, i.e., $n = lb + r$, $r = 0, \ldots, b-1$.
2. For $t = 1, b+1, 2b+1, \ldots, lb+1$, let $B_t^{*CBB} = \left(X_t^{*CBB}, \ldots, X_{t+b-1}^{*CBB}\right) = B_{k_t}$, where k_t are i.i.d. random variables drawn from a discrete uniform distribution $P(k_t = s) = 1/n$ for $s = 1, \ldots, n$.
3. Join the selected $l+1$ blocks $\left(B_1^{*CBB}, \ldots, B_{l+1}^{*CBB}\right)$ and take the first n observations to get the bootstrap sample $\left(X_1^{*CBB}, \ldots, X_n^{*CBB}\right)$ of the same length as the original one.

The cEMBB algorithm is a simple modification of the CBB one.

1. Define a bivariate series $Y_i = (X_i, i)$ and then do the CBB on the sample $(Y_1, \ldots, Y_n)$ to obtain $(Y_1^*, \ldots, Y_n^*)$.

Note that in the second coordinate of the series $Y_1^*, \ldots, Y_n^*$ we preserve the information on the original time indices of chosen observations. This simple modification of the CBB approach allows to construct the consistent estimators for the Fourier coefficients of the autocovariance function of a PC time series.

cGSBB Algorithm

The cGSBB algorithm differs from the CBB only in the second step, i.e.,

2'. For $t = 1, b+1, 2b+1, \ldots, lb+1$, let

$$B_t^{*cGSBB} = \left(X_t^{*cGSBB}, \ldots, X_{t+b-1}^{*cGSBB}\right) = B_{k_t} \tag{1}$$

where k_t is i.i.d. random variables drawn from a discrete uniform distribution $P(k_t = t + vd) = 1/w$ for $v = 0, 1, \ldots, w-1$.

Constructing the CBB or the cEMBB sample we are choosing among n blocks. The cGSBB case is much more subtle. The idea is to perfectly mimic the periodic structure of the data. Thus, during the block selection process the additional criteria need to be fulfilled.

Remark. It is well known that a PC time series with period d can be equivalently expressed as d-variate stationary time series (see e.g., [15]). Moreover, bootstrap methods and their properties are well investigated in the stationary case (see [17]) and hence it may seem useless to consider bootstrap for univariate periodic time series (especially when one is interested in estimation of the mean). However, as pointed out in [8] (Sect. 2.4) bootstrap approach via one dimensional time series is preferable. There are several reasons for that. The most important is that the EMBB and the GSBB allow to use blocks that do not contain an integer number of periods. Moreover, these blocks can start with any observation. None of those two properties hold when one wants to use the multivariate stationary representation of PC time series. As a result there are no bootstrap approaches designed for d-variate stationary time series that are equivalent to the EMBB and the GSBB for univariate PC time series.

Bootstrap Confidence Intervals

The consistency of the cGSBB and the cEMBB for $a(\lambda,\tau)$ was shown in [10] and [7], respectively. Since both results were obtained under the same assumptions, in order to recall them, to simplify notation, we use $\widehat{a}_n^*(\lambda,\tau)$ instead of $\widehat{a}_n^{*cEMBB}(\lambda,\tau)$ and $\widehat{a}_n^{*cGSBB}(\lambda,\tau)$. This means that whenever $\widehat{a}_n^*(\lambda,\tau)$ is used all statements are valid for both bootstrap approaches.

Theorem 2. *Assume that* $\{X_t, t\in\mathbb{Z}\}$ *is an* α*-mixing PC time series with* $\mathrm{E}(X_t)\equiv 0$ *and* $\mathrm{WP}(4)$ *such that:*

(i) $\sup_{t\in\mathbb{Z}}\mathrm{E}|X_t|^{8+2\delta}<\infty$ *for some* $\delta>0$,
(ii) $\sum_{\tau=1}^{\infty}\tau\alpha_X^{\delta/(4+\delta)}(\tau)<\infty$.

If $b\to\infty$ *as* $n\to\infty$ *such that* $b=o(n)$, *then*

$$\rho\Big(\mathscr{L}\left(\sqrt{n}\left(\widehat{a}_n(\lambda,\tau)-a(\lambda,\tau)\right)\right),\mathscr{L}^*\left(\sqrt{n}\left(\widehat{a}_n^*(\lambda,\tau)-\mathrm{E}^*\widehat{a}_n^*(\lambda,\tau)\right)\right)\Big)\xrightarrow{p}0,$$

where ρ *is any distance metricizing weak convergence of probability measures on* $\mathbb{R}^{2r}$.

$\mathscr{L}(\cdot)$ denotes probability low and $\mathscr{L}^*(\cdot)$ is its bootstrap counterpart. Moreover, E^* is the conditional expectation given the sample $(X_1,\ldots,X_n)$.

Using Theorem 2 one may construct bootstrap pointwise confidence intervals for the real and the imaginary part of $a(\lambda,\tau)$. For the sake of simplicity below we describe the basic ideas only for $\Re a(\lambda,\tau)$. As we mentioned before the asymptotic variance is very difficult to estimate. Since the construction of the bootstrap percentile confidence intervals does not require variance estimation, this kind of interval is the most often used for different characteristics of PC time series. To get it for a fixed frequency λ and shift τ we define the following statistic

$$K_{CBB}(x)=P^*\left(\sqrt{n}\Re\left(\widehat{a}_n^{*cEMBB}(\lambda,\tau)-\mathrm{E}^*\left(\widehat{a}_n^{*cEMBB}(\lambda,\tau)\right)\right)\le x\right),$$

where P^* is the conditional probability given the sample $(X_1,\ldots,X_n)$. Then, the equal-tailed 95% bootstrap confidence interval for $\Re a(\lambda,\tau)$ obtained using the cEMBB is of the form

$$\left(\Re\widehat{a}_n(\lambda,\tau)-\frac{u_{cEMBB}(0.975)}{\sqrt{n}},\Re\widehat{a}_n(\lambda,\tau)-\frac{u_{cEMBB}(0.025)}{\sqrt{n}}\right),$$

where $u_{cEMBB}(0.975)$ and $u_{cEMBB}(0.025)$ are 97.5% and 2.5% quantiles of $K_{cEMBB}(x)$, respectively. The confidence interval for the cGSBB is defined correspondingly.

In real data applications usually many different values of λ and τ are considered. In such situation the simultaneous confidence intervals are used. To obtain them the consistency of bootstrap for the smooth functions of $a(\lambda,\tau)$ is required. Such results for the cEMBB and the cGSBB can be found in [10] and [7]. The quantiles of the 95% equal-tailed bootstrap simultaneous intervals can be calculated using the maximum and

the minimum statistics. Let us assume that one is interested in frequencies $\lambda_1, \ldots, \lambda_r$ and shift τ. Again we describe the construction only for the cEMBB. We define:

$$K_{cEMBB,max}(x) = P^*\left(\sqrt{n}\max_i \Re\left(\widehat{a}_n^{*cEMBB}(\lambda_i,\tau) - \mathrm{E}^*\left(\widehat{a}_n^{*cEMBB}(\lambda_i,\tau)\right)\right) \le x\right),$$
$$K_{cEMBB,min}(x) = P^*\left(\sqrt{n}\min_i \Re\left(\widehat{a}_n^{*cEMBB}(\lambda_i,\tau) - \mathrm{E}^*\left(\widehat{a}_n^{*cEMBB}(\lambda_i,\tau)\right)\right) \le x\right)$$

and we get the confidence region of the form

$$\left(\widehat{a}_n(\lambda_i,\tau) - \frac{u_{cEMBB,max}(0.975)}{\sqrt{n}}, \widehat{a}_n(\lambda_i,\tau) - \frac{u_{cEMBB,min}(0.025)}{\sqrt{n}}\right), \tag{2}$$

where $i = 1, \ldots, r$ and $u_{cEMBB,max}(0.975)$ and $u_{cEMBB,min}(0.025)$ are 97.5% and 2.5% quantiles of $K_{cEMBB,max}(x)$ and $K_{cEMBB,min}(x)$, respectively.

In the next section we compare bootstrap pointwise and simultaneous confidence intervals obtained with the cEMBB and the cGSBB. We try to provide answers for questions posed in Sect. 1.

3 Simulation Study

The aim of our study is to compare the performance of the cEMBB and the cGSBB applied to construct the pointwise and the simultaneous confidence intervals for the real and the imaginary parts of the coefficients of the autocovariance function. For that purpose the actual coverage probabilities (ACPs), the average lengths and the average upper and lower quantiles values were calculated for all constructed intervals. Finally, confidence intervals were used to identify the significant frequencies. For our consideration we chose a few examples of PC time series that are listed below.

M1 $X_t = \cos(2\pi t/4)\varepsilon_t^1 + \cos(2\pi t/5)\varepsilon_t^2 + Z_t$,
M2 $X_t = \cos(2\pi t/4)\varepsilon_t^1 + \cos(2\pi t/6)\varepsilon_t^2 + Z_t$,
M3 $X_t = \cos(2\pi t/4)\varepsilon_t^1 + \cos(2\pi t/8)\varepsilon_t^2 + Z_t$,
M4 $X_t = 4\cos(2\pi t/4)\varepsilon_t^1 + \cos(2\pi t/5)\varepsilon_t^2 + Z_t$,
M5 $X_t = 8\cos(2\pi t/4)\varepsilon_t^1 + \cos(2\pi t/5)\varepsilon_t^2 + Z_t$,

where Z_t is zero-mean moving-average time series of the form

$$Z_t = 0.5\varepsilon_{t-3} + 0.3\varepsilon_{t-2} + 0.2\varepsilon_{t-1} + \varepsilon_t.$$

$\{\varepsilon_t\}_{t\ge1}, \{\varepsilon_t^1\}_{t\ge1}, \{\varepsilon_t^2\}_{t\ge1}$ are i.i.d. from the standard normal distribution. Moreover, the initial observations in each model were generated as standard normal random variables.

Each considered model **M1**–**M5** has exactly 4 true significant frequencies. Models **M1**–**M3** differ in period length and distance between the consecutive significant frequencies. In fact we were changing period lengths of the cosine function to see what will happen if two consecutive frequencies will be close to (**M3**) or far from each other (**M2**). Better separation may improve the detection. Moreover, if two frequencies λ_1

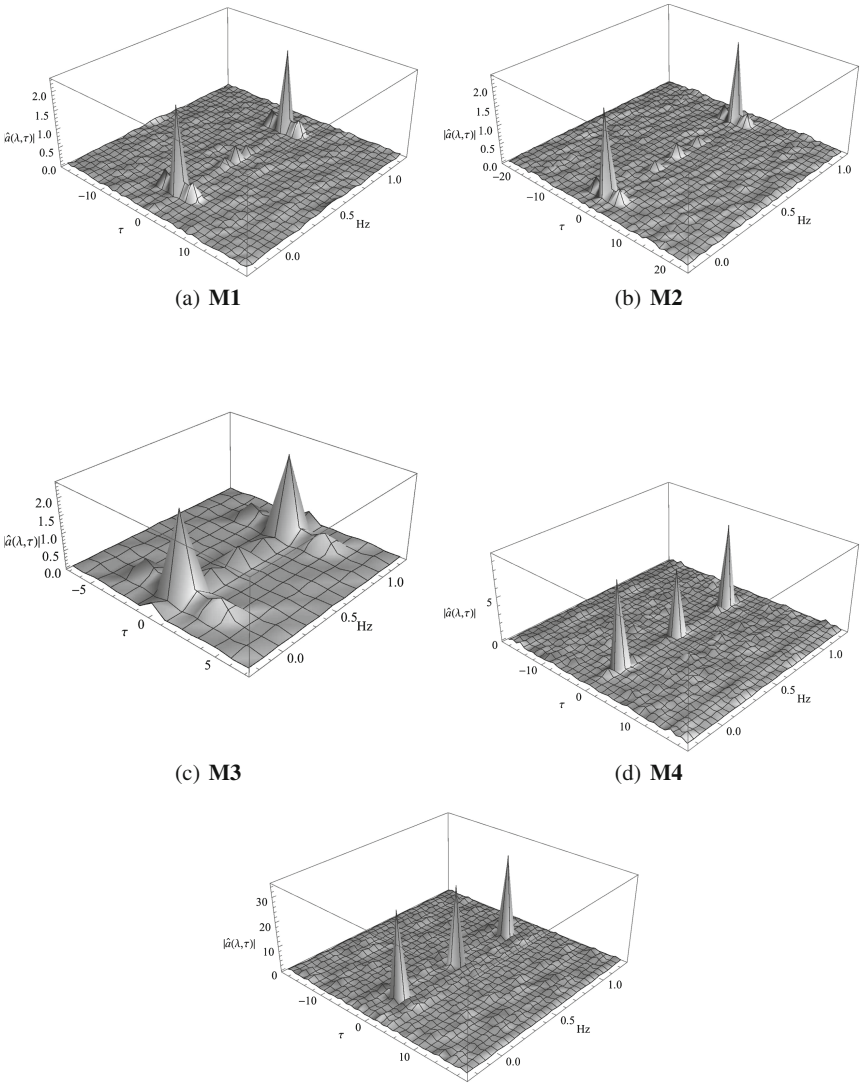
and λ_2 are close to each other and $a(\lambda_1,\tau)$ is much bigger than $a(\lambda_2,\tau)$, the detection of λ_2 can be more difficult than in the situation when $a(\lambda_1,\tau)$ and $a(\lambda_2,\tau)$ are comparable. The frequency $\lambda = 0$ is always the strongest (value of the corresponding coefficient is always the largest among all significant frequencies). Thus, to investigate its influence on our results the distance between the three remaining frequencies and $\lambda = 0$ was set to be short (**M3**), medium (**M2**) and large (**M1**). Finally, we added to our considerations models **M4**–**M5** to check how strengthening of some frequencies will affect the detections. Note that **M4** and **M5** were obtained from **M1** by multiplying the first cosine function by an appropriate constant. That increased the value of $a(\lambda,0)$ from 2.38 (**M1**) to 9.88 (**M4**) and 33.88 (**M5**) for $\lambda = 0$ Hz and from 0.5 (**M1**) to 8 (**M4**) and 32 (**M5**) for $\lambda = 0.5$ Hz. Values of $a(\lambda,0)$ for $\lambda = 0.4$ Hz and $\lambda = 0.6$ Hz remained unchanged and were equal to 0.25. In Fig. 1 we present the estimated values of $|a(\lambda,\tau)|$ for **M1**–**M5** for sample size $n = 1920$.

For our study we took three sample sizes, namely $n \in \{120, 480, 1920\}$. We set $\tau = 0$ and we considered all frequencies $\lambda \in \{2k\pi/d : k = 0,\ldots,d-1\}$. Using the cEMBB and the cGSBB we constructed the 95% equal-tailed pointwise and simultaneous confidence intervals (see Sect. 2) for the real and the imaginary part of $a(\lambda,\tau)$. The number of bootstrap resamples was $B = 500$ and the number of algorithm iteration was 1000. The block lengths b were chosen from the set $\{1,2,\ldots,100\}$. To compare the cEMBB and the cGSBB we calculated the ACPs, the average lengths and the average quantiles for all constructed confidence intervals. Since we considered 100 values of block length, the computation needed to be supported by the supercomputer available via the PL-Grid Infrastructure. For the sake of clarity we split the summary of our results into three parts.

Actual Coverage Probabilities

For the sake of simplicity, from now on we denote by cEMBB-ACPs and cGSBB-ACPs the ACPs obtained with the cEMBB and the cGSBB, respectively. The main conclusions are the same for all models. Generally, the cEMBB provides higher ACPs than the cGSBB. This means that confidence intervals constructed using the cEMBB are wider than corresponding ones obtained with the cGSBB. Moreover, in most of the cases taking $b < 20$ we get confidence intervals too wide (using the cEMBB) or too narrow (using the cGSBB).

Some of the cEMBB-APC curves contain periodic structure (see Fig. 2(a)). In fact for **M1**–**M5** this phenomena can be observed for frequencies that are close to $\lambda = 0$ Hz, i.e., $\lambda = 1/d$ Hz and $\lambda = 2/d$ Hz. Let us recall that $a(0,0)$ is always the largest among $a(\lambda,0)$ values for $\lambda \in \Lambda_0$. For **M1** $a(0,0)$ is almost 5 times higher than the corresponding value for $\lambda = 0.5$ Hz and 10 times than for $\lambda = 0.4$ Hz and $\lambda = 0.6$ Hz. Interestingly, for models **M4** and **M5** periodicity appears also for $\lambda = 0.4$ Hz and $\lambda = 0.45$ Hz. These two frequencies are in neighbourhood of another very strong frequency $\lambda = 0.5$ Hz. Sometimes periodicity is difficult to detect in APCs graphs. However, it can be much easier captured in figures presenting the average lengths of confidence intervals. Moreover, it is worth noticing that periodic structure appears not only in the context of the pointwise confidence intervals. We deal with it also in cEMBB-ACP graphs for the simultaneous confidence intervals. On the other hand, the cGSBB provides ACP curves without any periodicity (see Fig. 2(b)). We think that the periodicity phenomena may be



(a) **M1** (b) **M2**

(c) **M3** (d) **M4**

(e) **M5**

Fig. 1. Estimated values of $|a(\lambda, \tau)|$ for **M1**–**M5** with sample size $n = 1920$.

caused by the bias of the cEMBB estimator. Both bootstrap estimators are only asymptotically unbiased but bias of the cEMBB estimator may be strongly dependent on the chosen block length. Let us recall that the cGSBB preserves the periodic structure contained in the original data, while the cEMBB destroys it completely.

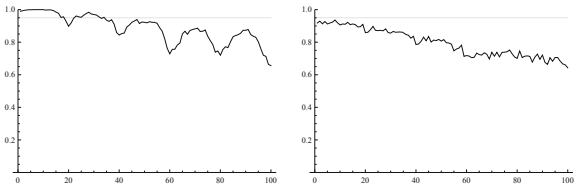


Fig. 2. Model **M1** with $n = 120$: ACPs (black lines) of pointwise confidence intervals for $\Re a(0.05\,\text{Hz},0)$ together with nominal coverage level (grey lines) for cEMBB (first column) and cGSBB (second column). On the horizontal axis block length $b \in \{1,2,\ldots,100\}$.

Block Length Choice - Validity of Some Existing Approaches

To popularize bootstrap techniques in the analysis of the PC time series, it is important to provide a method of block length choice. For now there is no such result. It concerns not only coefficients of the autocovariance function. Even in the simplest case of the overall mean estimation there are no indications for practitioners how to choose b. For stationary time series this issue is quite well investigated. Thus, we decided to check if some of the existing heuristic approaches can be applied for our non-stationary case.

In the literature it is often advised for periodic data to use the MBB with the block length equal to an integer multiple of the period length (see e.g., [24]). Our study confirms the validity of this approach. For $b = kd$, $k \in \mathbb{N}$ the cEMBB-ACP curves behave very similar to the cGSBB-ACP ones. In general median differences for the pointwise confidence intervals range from 0% to 3%, but for $n > 120$ from 0% to 0.5%. For the simultaneous confidence intervals the situation is very similar. For $n = 120$ the largest absolute difference is around 4% and for $n > 120$ all absolute differences belong to the interval $(0, 0.019)$. If periodic structure is present, a small change of the b value may result in big change of the actual coverage probability. Additionally, for large samples one may always find the block length being an integer multiple of the period length that provides cEMBB-ACPs close to nominal level.

The Minimum Volatility Method (see [22] p. 197) assumes that the block length b should be chosen from the region in which confidence interval is a stable function of b. For that purpose one may use a plot of the average confidence interval lengths or the average quantile values. We look for a region in which the curve is quite flat and there is not much variation. This method is not suitable for the cEMBB because of the periodic structure that appears sometimes. However, we tried to use it for the cGSBB. Unfortunately, we did not manage to find rules how to select b values for the different frequencies to obtain cGSBB-ACPs close to 95%. Curves of the average lengths of pointwise and simultaneous confidence intervals obtained with the cGSBB are quite

smooth and having low volatility. Some parts of the functions are quite flat, but often they do not contain block lengths corresponding to ACPs close to 95%.

The approach based on the logarithm of quantile for the subsampling method was recently discussed in [2] (see also references therein). The author proposed to find the largest b before the function becomes more erratic. Unfortunately, this methods does not work in the considered problem. When periodic structure is not present, the log of quantiles are smooth. We cannot localize regions with different behaviour.

Block Length Choice - New Heuristic Method for the cEMBB

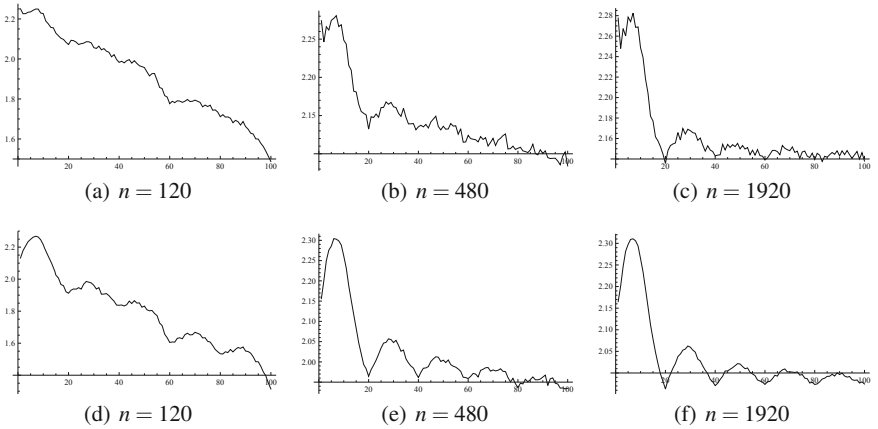
Inspired by the aforementioned methods, to choose the value b we investigated plots of average lengths and logarithm of upper quantiles. In the latter case we managed to detect the property that allows us to choose the block length for the cEMBB. However, we would like to indicate that for now we do not have any theoretical confirmation for the proposed heuristic approach. Thus, further research needs to be done to confirm its validity. Our study shows that it seems to work quite well for the simultaneous confidence intervals.

Block Length Choice Algorithm

1. If the log of quantile contains strong periodic structure, we look at local minima of the considered function. Their values create a string $m_1, m_2, \ldots$, which is decreasing starting from m_o, $o \in \mathbb{N}$. Finally, we choose the block length b for which the log of the quantile is equal to m_o.
2. If the log of quantile does not contain any periodic structure, we do not consider very small or very large block lengths. From the remaining "reasonable" block lengths we choose the largest b before the first break point, i.e., point after which the function starts to decrease sharply. In fact for very small b a sharp decrease can also be observed. But we are interested in the largest b from the first region where the function is quite flat. We would like to indicate that this choice is not always easy. Sometimes the region in which the function is not decreasing is very small.

In Fig. 3 we present the log of quantiles for the cEMBB simultaneous confidence intervals for $\Re a(\lambda,0)$ and $\Im a(\lambda,0)$ obtained for **M1** model. Using the proposed algorithm for **M1**–**M5** we chose the block lengths for the cEMBB simultaneous confidence intervals. Table 1 contains obtained values together with the ACPs generated by them for each sample size. In general, independently of the considered model and sample size, the chosen b differs for the real and imaginary part of $a(\lambda,0)$. Those for $\Im a(\lambda,0)$ are higher than the corresponding ones for $\Re a(\lambda,0)$. The ACPs obtained for $\Re a(\lambda,0)$ are usually below the nominal coverage level. The lowest value equal to 89% was got for **M3** with $n = 120$. The highest equal to 95.3% was observed for **M4** with $n = 1920$. For other cases the ACPs belong to interval $(91\%, 95\%)$. On the other hand the simultaneous confidence intervals for $\Im a(\lambda,0)$ are too wide independently on the sample size. The ACPs change from 95.7% to 97.9%.

For $n = 1920$ in the case of the real and the imaginary part of $a(\lambda,\tau)$ sometimes we were forced to choose some $b \leq 100$ while the minima of log of quantiles in the corresponding figures never started to decrease. In such a situation we were taking the largest



(a) $n = 120$ (b) $n = 480$ (c) $n = 1920$

(d) $n = 120$ (e) $n = 480$ (f) $n = 1920$

Fig. 3. Model **M1**: logarithms of the upper quantiles of simultaneous confidence intervals for $\Re a(\lambda, 0)$ (first row) and for $\Im a(\lambda, 0)$ (second row), $\lambda \in \overline{\Lambda}$ for cEMBB. On the horizontal axis block length $b \in \{1, 2, \ldots, 100\}$.

possible integer multiple of the period length. But in fact we should consider $b > 100$. For **M1**, **M4** and **M5** we performed an additional study taking $b \in \{120, 140, \ldots, 600\}$. The new choices of b values are in Table 2. One may observe that new cEMBB-ACPs are closer to 95%. Especially in the case of $\Im a(\lambda, \tau)$ the improvement is significant.

Finally, we compared results for **M1** with those for **M4** and **M5**. Let us recall that these models differ in values of $\Re a(0\,\text{Hz}, 0)$ and $\Re a(0.5\,\text{Hz}, 0)$. **M4** and **M5** contain two strongly significant frequencies ($\lambda = 0\,\text{Hz}$ and $\lambda = 0.5\,\text{Hz}$) in contrast to **M1** for which frequency $\lambda = 0\,\text{Hz}$ is strongly dominating all other frequencies. For **M4**–**M5** we obtained with our algorithm b values that resulted in cEMBB-ACPs that are closer to 95% than the corresponding ones for **M1**. Strengthening of some frequencies improved their detection.

The proposed algorithm works quite well for the simultaneous confidence intervals obtained with the cEMBB. Unfortunately, for the cGSBB we did not manage to find any properties of the log of quantiles curves that would allow us to choose the block length. For the cEMBB we take advantage of the fact that for $b = 1$ it provides too wide confidence intervals and hence the first stabilization area usually contains b close to optimal one. Using the cGSBB with small b results sometimes in too low and sometimes too high ACP and hence, for example, local maximum can be very misleading.

Simplification of Algorithm for the Pointwise Confidence Intervals Case

Choice of the block length for the pointwise confidence intervals is more problematic. In practice one would prefer to use one value of b for all considered frequencies. Moreover, when period length is long, the amount of calculations to get all plots is too big. Thus, we decided to check what would happen if we choose b only on the basis of curves that contain periodic structure. And to make procedure even simpler for each model we restricted our consideration only to two consecutive frequencies from $\overline{\Lambda}$ after $\lambda = 0\,\text{Hz}$, i.e., $\lambda = 1/d$ Hz and $\lambda = 2/d$ Hz. For the $n = 480$ and $n = 1920$ the absolute deviation

Table 1. For **M1**–**M5** and each sample size n and block length choice $b \leq 100$ for the cEMBB used to construct simultaneous confidence intervals for $\Re a(\lambda,0)$ (column 3) and $\Im a(\lambda,0)$ (column 5), where $\lambda \in \overline{\Lambda}$. In columns 4 and 6 are the obtained cEMBB-ACPs.

Model	n	Real part		Imaginary part	
		b	cEMBB-ACP	b	cEMBB-ACP
M1	120	10	0.935	20	0.968
	480	40	0.912	60	0.957
	1920	100	0.934	100	0.976
M2	120	5	0.918	24	0.962
	480	5	0.947	48	0.970
	1920	48	0.935	84	0.975
M3	120	8	0.890	16	0.966
	480	16	0.931	24	0.974
	1920	60	0.949	96	0.964
M4	120	8	0.944	20	0.965
	480	30	0.938	60	0.975
	1920	60	0.953	100	0.979
M5	120	5	0.947	20	0.957
	480	30	0.938	60	0.966
	1920	100	0.949	100	0.973

from 95% of majority of APC values is maximally 2.3% and 2.1% in the $\Re a(\lambda,\tau)$ and $\Im a(\lambda,\tau)$ case, respectively. It seems that our method of block length works quite well also for the pointwise confidence intervals.

Table 2. For **M1**, **M4**, **M5** and sample size $n = 1920$ and block length choice $b > 100$ for the cEMBB used to construct simultaneous confidence intervals for $\Re a(\lambda,0)$ (column 3) and $\Im a(\lambda,0)$ (column 5), where $\lambda \in \overline{\Lambda}$. In columns 4 and 6 are the obtained cEMBB-ACPs.

Model	n	Real part		Imaginary part	
		b	cEMBB-ACP	b	cEMBB-ACP
M1	1920	200	0.943	320	0.972
M4	1920	60	0.953	340	0.952
M5	1920	180	0.946	280	0.955

4 Conclusions

To obtain bootstrap confidence intervals for the coefficients of the autocovariance function of PC time series we can use two bootstrap approaches. These are the EMBB and

the GSBB or their circular versions. We performed an extensive study to compare their behaviour and we are unable to state, which of them is better. The ACPs obtained with the cEMBB are usually higher than the corresponding ones for the cGSBB. It means that the cEMBB confidence intervals are often wider than the cGSBB ones. For very large samples, the performance of both bootstrap methods is very comparable.

The cEMBB-ACPs curves sometimes contain periodic structure. It seems this happens for frequencies that are close to the true strong frequencies. This phenomena may be considered as a disadvantage of the cEMBB, because a slight change in the chosen block length has a significant affect on the confidence interval. However, it can also be used in the future to construct a test for the significant frequency detection. Moreover, we used this feature to propose a heuristic method of block length choice that seems to work well in the case of pointwise and the simultaneous confidence intervals. In fact for the pointwise confidence intervals we simplified it to make the choice only on the basis of 1–2 figures, which substantially reduces the amount of computations and provides satisfactory results. We also checked a few heuristic approaches for the block length choice that were designed for the stationary data, but none of them is working for PC processes. We did not succeed in finding any indication how the block length should be chosen when the cGSBB is used. Finally, we managed to confirm the suggestion appearing in the literature that for the cEMBB block lengths that are integer multiples of the period length should be considered. If we do not have any other method of block length choice, taking $b = kd$ allows us to avoid the periodicity effect and can provide the ACP close to the nominal one.

Acknowledgements. This research was partially supported by the PL-Grid Infrastructure. Anna Dudek has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 655394. Anna Dudek was partially supported by the Faculty of Applied Mathematics AGH UST statutory tasks within subsidy of Ministry of Science and Higher Education. The authors express their sincere gratitude to Prof. Patrice Bertail for his comments and ideas.

References

1. Antoni J (2009) Cyclostationarity by examples. Mech Syst Signal Process 23(4):987–1036
2. Bertail P (2011) Comments on: subsampling weakly dependent time series and application to extremes. TEST 20(3):487–490
3. Carlstein E (1986) The use of subseries methods for estimating the variance of a general statistic from a stationary time series. Ann Stat 14:1171–1179
4. Dehay D, Dudek AE (2015) Bootstrap method for Poisson sampling almost periodic process. J Time Ser Anal 36(3):327–351

5. Dehay D, Dudek AE (2017) Bootstrap for the second-order analysis of Poisson-sampled almost periodic processes. Electron J Stat 11(1):99–147
6. Dehay D, Dudek AE, Elbadaoui M (2018) Bootstrap for almost cyclostationary processes with jitter effect. Digit Signal Process 73:93–105
7. Dudek AE (2015) Circular block bootstrap for coefficients of autocovariance function of almost periodically correlated time series. Metrika 78(3):313–335
8. Dudek AE (2018) Block bootstrap for periodic characteristics of periodically correlated time series. J Nonparametric Stat 30(1):87–124
9. Dudek AE, Leśkow J, Paparoditis E, Politis DN (2014a) A generalized block bootstrap for seasonal time series. J Time Ser Anal 35:89–114
10. Dudek AE, Maiz S, Elbadaoui M (2014b) Generalized seasonal block bootstrap in frequency analysis of cyclostationary signals. Signal Process 104C:358–368
11. Efron B (1979) Bootstrap methods: another look at the jacknife. Ann Stat 7:1–26
12. Efron B, Tibshirani RJ (1993) An introduction to the bootstrap. Chapman & Hall/CRC, Boca Raton
13. Gardner WA, Napolitano A, Paura L (2006) Cyclostationarity: half a century of research. Signal Process 86(4):639–697
14. Gladyshev EG (1961) Periodically correlated random sequences. Soviet Math 2:385–388
15. Hurd HL, Miamee AG (2007) Periodically correlated random sequences: spectral theory and practice. Wiley, Hoboken
16. Künsch H (1989) The jackknife and the bootstrap for general stationary observations. Ann Stat 17:1217–1241
17. Lahiri SN (2003) Resampling methods for dependent data. Springer, New York
18. Lenart Ł, Leśkow J, Synowiecki R (2008) Subsampling in testing autocovariance for periodically correlated time series. J Time Ser Anal 29:995–1018
19. Liu R, Singh K (1992) Moving block jackknife and bootstrap capture weak dependence. In: Le Page R, Billard L (eds.) Exploring the limits of bootstrap, pp 225–248. Wiley, New York
20. Napolitano A (2012) Generalizations of cyclostationary signal processing: spectral analysis and applications. Wiley-IEEE Press, London
21. Politis DN, Romano JP (1992) A circular block-resampling procedure for stationary data. Wiley series in probability and mathematical statistics. Probability and mathematical statistics). Wiley, New York, pp 263–270
22. Politis DN, Romano JP, Wolf M (1999) Subsampling. Springer, New York
23. Salas JD, Delleur JW, Yevjevich V, Lane WL (1980) Applied modeling of hydrologic time series. Water Resources Publications, Littleton
24. Srinivasa VV, Srinivasan K (2005) Hybrid moving block bootstrap for stochastic simulation of multi-site multi-season streamflows. J Hydrol 302:307–330
25. Synowiecki R (2007) Consistency and application of moving block bootstrap for nonstationary time series with periodic and almost periodic structure. Bernoulli 13(4):1151–1178
26. Synowiecki R (2008) Metody resamplingowe w dziedzinie czasu dla niestacjonarnych szeregów czasowych o strukturze okresowej i prawie okresowej. PhD Thesis at the Departement of Applied Mathematics, AGH University of Science and Technology, Krakow, Poland. http://winntbg.bg.agh.edu.pl/rozprawy2/10012/full10012.pdf. Accessed 7 Aug 2017

The Coherent and Component Estimation of Covariance Invariants for Vectorial Periodically Correlated Random Processes and Its Application

I. Javors'kyj[1,2], I. Matsko[1(✉)], R. Yuzefovych[1,3], G. Trokhym[1], and P. Semenov[4]

[1] Karpenko Physico-Mechanical Institute of National Academy of Sciences of Ukraine, Naukova Str. 5, Lviv 79060, Ukraine
ihor.yavorskyj@gmail.com, matsko.ivan@gmail.com, roman.yuzefovych@gmail.com, george.trokhym@gmail.com
[2] Institute of Telecommunication, UTP University of Sciences and Technology, Al. Prof. S. Kaliskiego 7, 85796 Bydgoszcz, Poland
[3] Lviv Polytechnic National University, Bandera Str. 12, Lviv 79000, Ukraine
[4] Limited Liability Company "PORTTECHEXPERT", Pros. Peace, 25-B, Odessa Region, Izmail 68604, Ukraine

Abstract. The coherent and the component methods for the estimation of the linear covariance invariants of vectorial periodically correlated random processes (PCRP) are considered. The coherent estimators are calculated by averaging of the samples taken through the non-stationary period. The component estimators are built in the form of trigonometric polynomials, Fourier coefficients of which are calculated by weighted averaging of PCRP realization. The properties of the continuous and the discrete estimators are investigated, the asymptotical unbiasedness and mean square consistency are proved. The formulae for their biases and variances described dependency of these quantities on realization length, time sampling and PCRP covariance components are obtained. The conditions of the absence of the aliasing effects of the first and the second kinds are given. The comparison of the coherent and component estimators is carried out for the case of the amplitude modulated signals. It is shown that the advantage of the component method over coherent grows as a rate of PCRP correlations clumping increases. The example of the using of vectorial statistical processing for diagnosis of rolling bearing are given. The investigation results show that using of the covariance function invariants allow to improve the efficiency of the fault detection and to establish of the defect spatial properties.

Keywords: Vectorial periodically correlated random processes · Covariance invariant · Coherent and component methods · Bias and variance analysis · Rolling bearing

F. Chaari et al. (Eds.): CSTA 2017, ACM 16, pp. 56–91, 2020.
https://doi.org/10.1007/978-3-030-22529-2_4

1 Introduction

Vectorial random processes are the natural mathematical model for the vibration description, hence physical quantities, which characterize them, i.e. displacement, velocity, acceleration are vectors. The feature of the vectorial approach is putting under consideration of invariant quantities, which characterize observed object state independently on the coordinate system, in which the measurement is provided (Dragan et al. 1987; Javorśkyj 2013; Javorśkyj et al. 2017a, b, c).

For the analysis of rotary machinery vibration and detection of their faults the mathematical models in the form of PCRP are widely used (Javorskyj 2013; McCormick and Nandi 1998; Capdessus et al. 2000; Antoni et al. 2004; Antoni 2009; Javorskyj et al. 2017a, b, c). The consistency of such approach follows from the properties of vibration signals of rotary machineries, which are stochastically modulated by their nature with clearly shaped features of cyclic recurrence. Based on such position we will describe the components of vibration vector by PCRP, and their interdependency – by jointly PCRP. The properties of the stochasticity and recurrence are adequately represented in the mean functions of the vector $\vec{\xi}(t) = \vec{i}\,\xi_1(t) + \vec{j}\,\xi_2(t)$, i.e. in the functions $m_{\xi_p}(t) = E\xi_p(t), p = \overline{1,2}$, E – averaging operator, and also in auto- $b_{\xi_p}(t,\tau) = E\overset{\circ}{\xi}_p(t)\overset{\circ}{\xi}_p(t+\tau)$, $\overset{\circ}{\xi}_p(t) = \xi_p(t) - m_{\xi_p}(t)$ and cross-covariance functions $b_{\xi_p\xi_q}(t,\tau) = E\overset{\circ}{\xi}_p(t)\overset{\circ}{\xi}_q(t+\tau)$. These quantities periodically vary in time and can be represented in the Fourier series form:

$$m_{\xi_p}(t) = \sum_{k\in Z} m_k^{(\xi_p)} e^{ik\omega_0 t},\ b_{\xi_p}(t,\tau) = \sum_{k\in Z} b_k^{(\xi_p)}(\tau)e^{ik\omega_0 t}, \tag{1}$$

$$b_{\xi_p\xi_q}(t,\tau) = \sum_{k\in Z} b_k^{(\xi_p\xi_q)}(\tau)e^{ik\omega_0 t}, \tag{2}$$

where $\omega_0 = 2\pi/P$, P – nonstationarity period. We should not that issue of a cross covariance and cross spectral analysis is covered in (Gardner 1987; Sadler and Dandawate 1998). Covariance tensor-function

$$b_{\underset{\xi}{\rightarrow}}(t,\tau) = E\vec{\xi}(t)\otimes\vec{\xi}(t+\tau) = \begin{bmatrix} b_{\xi_1}(t,\tau) & b_{\xi_1\xi_2}(t,\tau) \\ b_{\xi_2\xi_1}(t,\tau) & b_{\xi_2}(t,\tau) \end{bmatrix}. \tag{3}$$

is a mutual covariance characteristic. Here $\otimes$ – sign of tensor product. On the basis of matrix (3) elements the quantities, which do not depend on coordinate system, where measurements of vector $\vec{\xi}(t)$ components were provided, can be formed. The linear

invariants $I_1(t,\tau)$ and $D(t,\tau)$, which are average values of scalar and skew products of vectors $\vec{\xi}(t)$ and $\vec{\xi}(t+\tau)$, respectively:

$$I_1(t,\tau) = \left|\overset{\circ}{\vec{\xi}}(t)\right|\left|\overset{\circ}{\vec{\xi}}(t+\tau)\right|\cos\left(\angle\overset{\circ}{\vec{\xi}}(t)\,\overset{\circ}{\vec{\xi}}(t+\tau)\right) = b_{\xi_1}(t,\tau)+b_{\xi_2}(t,\tau),$$

$$D(t,\tau) = \left|\overset{\circ}{\vec{\xi}}(t)\right|\left|\overset{\circ}{\vec{\xi}}(t+\tau)\right|\sin\left(\angle\overset{\circ}{\vec{\xi}}(t)\,\overset{\circ}{\vec{\xi}}(t+\tau)\right) = b_{\xi_1\xi_2}(t,\tau)-b_{\xi_2\xi_1}(t,\tau)$$

are the simplest. Invariant $I_1(t,\tau)$ is a sum of elements of the matrix (3) main diagonal and it can be considered as a measure of the collinearity of vectors $\overset{\circ}{\vec{\xi}}(t)$ and $\overset{\circ}{\vec{\xi}}(t+\tau)$. For $\tau = 0$ the linear invariant $I_1(t,0)$ is equal to sum of the variances of the vector components and characterizes the changes of vector $\overset{\circ}{\vec{\xi}}(t)$ intensity: $I_1(t,0) = b_{\xi_1}(t,0)+b_{\xi_2}(t,0) = E\left|\overset{\circ}{\vec{\xi}}(t)\right|^2$. Invariant $D(t,\tau)$ is a subtract of the elements of the matrix (3) second diagonal. If $D(t,\tau)\neq 0$ then the angle between vectors $\overset{\circ}{\vec{\xi}}(t)$ and $\overset{\circ}{\vec{\xi}}(t+\tau)$ in mean is not equal to zero. The value of $D(t,\tau)$ is maximum if this angle is equal to $\pi/2$. If direction of $\overset{\circ}{\vec{\xi}}(t)$ and $\overset{\circ}{\vec{\xi}}(t+\tau)$ in mean coincide then $D(t,\tau) = 0$. Thus in variant $D(t,\tau)$ can be considered as a rotation indicator.

Quadratic invariant $I_2(t,\tau)$ is a determinant of a symmetric part of the matrix (3). It is a main characteristic for the classification of the matrix (3) quadratic forms. Eigenvalues of invariant $I_2(t,\tau)$, namely $\lambda_1(t,\tau)$ and $\lambda_2(t,\tau)$, define extreme values of covariance function with respect to orthogonal directions. The second order curves, parameters of which are effective indicators for defining of defect geometric properties, can be built on their basis (Javorskyj et al. 2014a, b, c, 2017a, b, c).

The aim of the present paper is investigation of estimator properties of invariants and their Fourier coefficients, obtained using coherent (Javorskyj 2013; Javorskyj et al. 2007, 2014a, b, c) and component (Javorskyj 2013; Javorskyj et al. 2014a, b, c, 2010) methods. Estimators of invariants are formed using relations, defining them:

$$\hat{I}_1(t,\tau) = \hat{b}_{\xi_1}(t,\tau)+\hat{b}_{\xi_2}(t,\tau), \tag{4}$$

$$\hat{D}(t,\tau) = \hat{b}_{\xi_1\xi_2}(t,\tau)-\hat{b}_{\xi_2\xi_1}(t,\tau), \tag{5}$$

$$\hat{I}_2(t,\tau) = \hat{b}_{\xi_1}(t,\tau)\hat{b}_{\xi_2}(t,\tau)-\frac{1}{4}\left[\hat{b}_{\xi_1\xi_2}(t,\tau)+\hat{b}_{\xi_2\xi_1}(t,\tau)\right]^2,$$

$$\hat{\lambda}_{1,2}(t,\tau) = \frac{1}{2}\left[\hat{I}_1(t,\tau)\pm\left[\hat{I}_1^2(t,\tau)-4\hat{I}_2^2(t,\tau)\right]^{\frac{1}{2}}\right].$$

Estimator of quadratic invariant and estimators of eigen values $\lambda_1(t,\tau)$ and $\lambda_2(t,\tau)$ are non-linear functions of auto- and cross-covariance functions of vector components.

Because the derivation of the formulae for statistical characteristics of these estimators is complicated, a problem is simplified for analysis of statistics (4) and (5), characterizing properties of auto- and cross-covariance functions of vector random components in invariant form. Here necessary calculations are succeeded up to obtain of practically important formulas. On a basis of the last ones we can calculate systematic and mean-square estimation errors dependently on realization length θ, sampling interval h and also the signal parameters. Choosing also processing parameters, providing small errors of linear invariant estimation, we can hope to obtain admissible errors of invariant estimators $\hat{I}_2(t,\tau)$ and $\hat{\lambda}_{1,2}(t,\tau)$.

Synopsis. The paper consists of an introduction, five sections and conclusions. The second and third sections are divided into two subsections. Coherent estimators of the linear invariants are considered in Subsect. 2.1. In Subsect. 2.2 their Fourier coefficients estimators are analyzed. The properties of the linear invariant estimators in the form of a finite trigonometric polynomials (so called component estimators) are investigated in Sect. 3. The formulae for biases and their analysis are given in Subsect. 3.1. In Subsect. 3.2 for Gaussian processes the mean square convergence of the invariant components estimators is proved. The discrete estimators of Fourier coefficients are analyzed in Sect. 4. The interpolation formulae for invariants with finite number of Fourier coefficients are obtained. In Sect. 5 the obtained results are specified for amplitude-modulated vector components. The numeric results from analysis allow one to compare efficiency of coherent and component estimators. The results of vector covariance analysis for vibration of decanter Flottweg 24E are given in Sect. 6.

2 Coherent Covariance Analysis

2.1 Estimations of Covariance Invariants

Let us consider statistics (4) and (5). We assume at first that the estimators of auto- and cross-covariance functions of vectorial components $\xi_1(t)$ and $\xi_2(t)$ are calculated by averaging of the samples, taken via the non-stationary period P:

$$\hat{b}_{\xi_p}(t,\tau) = \frac{1}{N}\sum_{n=0}^{N-1}\left[\xi_p(t+\tau+nP) - \hat{m}_{\xi_p}(t+\tau+nP)\right] \times \left[\xi_p(t+nP) - \hat{m}_{\xi_p}(t+nP)\right], p = \overline{1,2}, \tag{6}$$

$$\hat{b}_{\xi_p\xi_q}(t,\tau) = \frac{1}{N}\sum_{n=0}^{N-1}\left[\xi_p(t+nP) - \hat{m}_{\xi_p}(t+nP)\right] \times \left[\xi_q(t+\tau+nP) - \hat{m}_{\xi_q}(t+\tau+nP)\right], \quad p,q = \overline{1,2}, \tag{7}$$

here with

$$\hat{m}_{\xi_{p,}\xi_q}(t) = \frac{1}{N}\sum_{n=0}^{N-1} \xi_{p,q}(t+nP),$$

and N is natural number.

Proposition 2.1. *The invariant estimators* (4) *and* (5) *when auto- and cross-covariance functions are calculated using formulae* (6) *and* (7) *if conditions*

$$\lim_{|\tau|\to\infty} \hat{b}_{\xi_p,\xi_q}(t,\tau) = 0, \lim_{|\tau|\to\infty} \hat{b}_{\xi_p\xi_q}(t,\tau) = 0, \qquad p,q = \overline{1,2}, \tag{8}$$

are satisfied, are asymptotically unbiased and for a finite number N their biases $\varepsilon\left[\hat{I}_1(t,\tau)\right] = E\hat{I}_1(t,\tau) - I_1(t,\tau)$ *and* $\varepsilon\left[\hat{D}(t,\tau)\right] = E\hat{D}(t,\tau) - D(t,\tau)$ *are:*

$$\varepsilon\left[\hat{I}_1(t,\tau)\right] = \sum_{k\in\mathbb{Z}} \varepsilon_k^{(I_1)}(\tau) e^{ik\omega_0 t}, \tag{9}$$

$$\varepsilon\left[\hat{D}(t,\tau)\right] = \sum_{k\in\mathbb{Z}} \varepsilon_k^{(D)}(\tau) e^{ik\omega_0 t}, \tag{10}$$

where

$$\varepsilon_k^{(I_1)}(\tau) = -\frac{1}{N}\left[B_k^{(I_1)}(\tau) + \frac{1}{N}\sum_{n=-N+1}^{N-1} |n| B_k^{(I_1)}(\tau+nP)\right], \tag{11}$$

$$\varepsilon_k^{(D)}(\tau) = -\frac{1}{N}\left[B_k^{(D)}(\tau) + \frac{1}{N}\sum_{n=-N+1}^{N-1} |n| B_k^{(D)}(\tau+nP)\right], \tag{12}$$

and $B_k^{(I_1)}(\tau) = B_k^{(\xi_1)}(\tau) + B_k^{(\xi_2)}(\tau)$, $B_k^{(D)}(\tau) = B_k^{(\xi_1\xi_2)}(\tau) - B_k^{(\xi_2\xi_1)}(\tau)$.

Proof. For analysis simplification we rewrite statistic (6) in the form:

$$\begin{aligned}
\hat{b}_{\xi_p}(t,\tau) &= \frac{1}{N}\sum_{n=0}^{N-1}\left[\left[\xi_p(t+\tau+nP) - m_{\xi_p}(t+\tau)\right] - \left[\hat{m}_{\xi_p}(t+\tau+nP) - m_{\xi_p}(t+\tau)\right]\right] \\
&\times \left[\left[\xi_p(t+nP) - m_{\xi_p}(t)\right] - \left[\hat{m}_{\xi_p}(t+nP) - m_{\xi_p}(t)\right]\right] \\
&= \frac{1}{N}\sum_{n=0}^{N-1}\left[\overset{\circ}{\xi}_p(t+nP)\overset{\circ}{\xi}_p(t+\tau+nP) - \overset{\circ}{\hat{m}}_{\xi_p}(t+nP)\overset{\circ}{\xi}_p(t+\tau+nP)\right. \\
&\left. - \overset{\circ}{\hat{m}}_{\xi_p}(t+\tau+nP)\overset{\circ}{\xi}_p(t+nP) + \overset{\circ}{\hat{m}}_{\xi_p}(t+nP)\overset{\circ}{\hat{m}}_{\xi_p}(t+\tau+nP)\right]
\end{aligned} \tag{13}$$

where

$$\overset{\circ}{\xi}_{p,q}(t) = \xi_{p,q}(t) - \underset{\xi_p,\xi_q}{m}(t),\ \overset{\circ}{\hat{m}}_{\xi_p,\xi_q}(t) = \frac{1}{N}\sum_{n=0}^{N-1}\overset{\circ}{\xi}_{p,q}(t+nP).$$

Similarly we have

$$\begin{aligned}\hat{b}_{\xi_p\xi_q}(t,\tau) = \frac{1}{N}\sum_{n=0}^{N-1}\Big[&\overset{\circ}{\xi}_p(t+nP)\overset{\circ}{\xi}_q(t+\tau+nP) - \overset{\circ}{\hat{m}}_{\xi_p}(t+nP)\overset{\circ}{\xi}_q(t+\tau+nP)\\ &- \overset{\circ}{\xi}_p(t+nP)\overset{\circ}{\hat{m}}_{\xi_q}(t+\tau+nP) + \overset{\circ}{\hat{m}}_{\xi_p}(t+nP)\overset{\circ}{\hat{m}}_{\xi_q}(t+\tau+nP)\Big]\end{aligned} \tag{14}$$

After ensemble averaging for expressions (13) and (14) we obtain

$$E\hat{b}_{\xi_p}(t,\tau) = b_{\xi_p}(t,\tau) - \frac{1}{N}\left[b_{\xi_p}(t,\tau) + \frac{1}{N}\sum_{n=-N+1}^{N-1}|n|b_{\xi_p}(t+\tau+nP)\right],$$

$$E\hat{b}_{\xi_p\xi_q}(t,\tau) = b_{\xi_p\xi_q}(t,\tau) - \frac{1}{N}\left[b_{\xi_p\xi_q}(t,\tau) + \frac{1}{N}\sum_{n=-N+1}^{N-1}|n|b_{\xi_p\xi_q}(t,\tau+nP)\right]$$

Hence

$$\varepsilon\left[\hat{I}_1(t,\tau)\right] = -\frac{1}{N}\left[I_1(t,\tau) + \frac{1}{N}\sum_{n=-N+1}^{N-1}|n|I_1(t,\tau+nT)\right],$$

$$\varepsilon\left[\hat{D}(t,\tau)\right] = -\frac{1}{N}\left[D(t,\tau) + \frac{1}{N}\sum_{n=-N+1}^{N-1}|n|D(t,\tau+nT)\right].$$

Taking into consideration Fourier series (1) and (2), we come to the formulae (11) and (12). If conditions (8) are satisfied then $\sum_{n=-N+1}^{N-1}|n|B_k^{(I_1)}(\tau+nT) \sim \mathrm{O}(N^\alpha)$ and $\sum_{n=-N+1}^{N-1}|n|B_k^{(D)}(\tau+nT) \sim \mathrm{O}(N^\alpha)$, where $\alpha<2$. Thus $\varepsilon\left[\hat{I}_1(t,\tau)\right] \to 0$ and $\varepsilon\left[\hat{D}(t,\tau)\right] \to 0$ if $N \to \infty$ □.

If covariance vanish within a time interval less than the period P (i.e. if $B_k^{(I_1)}(nT) = B_k^{(D)}(nT) \approx 0$ for all $k \in \mathbb{N}$), then

$$\varepsilon\left[\hat{I}_1(t,0)\right] \approx -\frac{I_1(t,0)}{N},\ \varepsilon\left[\hat{D}(t,0)\right] \approx \frac{D(t,0)}{N}.$$

We assume that the mean functions are estimated only for $t \in [0, P]$ and $\hat{m}_{\xi_1,\xi_2}(t+nP) = \hat{m}_{\xi_1,\xi_2}(t)$ when $t \notin [0, P]$. Then statistics (6) and (7) can be rewritten in the form:

$$\hat{b}_{\xi_p}(t,\tau) = \frac{1}{N}\sum_{n=0}^{N-1} \xi_p(t+nP)\xi_p(t+\tau+nP) - \hat{m}_{\xi_p}(t)\hat{m}_{\xi_p}(t+\tau), \tag{15}$$

$$\hat{b}_{\xi_p\xi q}(t,\tau) = \frac{1}{N}\sum_{n=0}^{N-1} \xi_p(t+nP)\xi_q(t+\tau+nP) - \hat{m}_{\xi_p}(t)\hat{m}_{\xi_q}(t+\tau). \tag{16}$$

Proposition 2.2. *The invariant estimators* (4) *and* (5) *when auto- and cross-covariance functions are calculated using formulae* (15) *and* (16)*, are asymptotically unbiased if conditions* (8) *are satisfied and for a finite number N their biases are defined by the Fourier series* (9) *and* (10)*, where*

$$\varepsilon_k^{(I_1)}(\tau) = -\frac{1}{N}\sum_{n=-N+1}^{N-1}\left(1-\frac{|n|}{N}\right)B_k^{(I_1)}(\tau+nP), \tag{17}$$

$$\varepsilon_k^{(D)}(\tau) = -\frac{1}{N}\sum_{n=-N+1}^{N-1}\left(1-\frac{|n|}{N}\right)B_k^{(D)}(\tau+nP). \tag{18}$$

Proof. We represent the statistics in the form:

$$\hat{b}_{\xi_p}(t,\tau) = \frac{1}{N}\sum_{n=0}^{N-1}\left[\overset{\circ}{\xi}_p(t+nP)\overset{\circ}{\xi}_p(t+\tau+nP) - \frac{1}{N}\sum_{r=0}^{N-1}\overset{\circ}{\xi}_p(t+nP)\overset{\circ}{\xi}_p(t+\tau+rP)\right],$$

$$\hat{b}_{\xi_p\xi q}(t,\tau) = \frac{1}{N}\sum_{n=0}^{N-1}\left[\overset{\circ}{\xi}_p(t+nP)\overset{\circ}{\xi}_q(t+\tau+nP) - \frac{1}{N}\sum_{r=0}^{N-1}\overset{\circ}{\xi}_p(t+nP)\overset{\circ}{\xi}_q(t+\tau+rP)\right].$$

Averaging these expressions, for biases of invariant estimators (4) and (5) we find:

$$\varepsilon\left[\hat{I}_1(t,\tau)\right] = -\frac{1}{N}\sum_{n=-N+1}^{N-1}\left(1-\frac{|n|}{N}\right)I_1(t,\tau+nP),$$

$$\varepsilon\left[\hat{D}(t,\tau)\right] = -\frac{1}{N}\sum_{n=-N+1}^{N-1}\left(1-\frac{|n|}{N}\right)D(t,\tau+nP).$$

Substituting representations (1)–(2) to these equalities, we come to Fourier series with coefficients (17) and (18) □.

We should note that for estimation of auto- and cross-covariance functions the following statistics can be used:

$$\hat{b}_{\xi_p}(t,\tau) = \frac{1}{N}\sum_{n=0}^{N-1}\left[\xi_p(t+nP)\xi_p(t+\tau+nP) - \hat{m}_{\xi_p}(t+nP)\hat{m}_{\xi_p}(t+\tau+nP)\right], \quad (19)$$

$$\hat{b}_{\xi_p\xi q}(t,\tau) = \frac{1}{N}\sum_{n=0}^{N-1}\left[\xi_p(t+nP)\xi_q(t+\tau+nP) - \hat{m}_{\xi_p}(t+nP)\hat{m}_{\xi_q}(t+\tau+nP)\right]. \quad (20)$$

It is easy to show that the biases of these estimators coincide with biases for estimators (15), (16).

For the variances of invariants estimators when auto- and cross-covariance functions are calculated using statistics (6), (7) or (15), (16) in the first approximation we find:

$$\mathrm{V}ar\left[\hat{I}_1(t,\tau)\right] = \frac{1}{N^2}\sum_{n,m=0}^{N-1}\left[G_{nnmm}^{(\xi_1\xi_1\xi_1\xi_1)}(t,\tau) + 2G_{nnmm}^{(\xi_1\xi_1\xi_2\xi_2)}(t,\tau) + G_{nnmm}^{(\xi_2\xi_2\xi_2\xi_2)}(t,\tau)\right],$$

$$\mathrm{V}ar\left[\hat{D}(t,\tau)\right] = \frac{1}{N^2}\sum_{n,m=0}^{N-1}\left[G_{nnmm}^{(\xi_1\xi_2\xi_1\xi_2)}(t,\tau) + 2G_{nnmm}^{(\xi_1\xi_2\xi_2\xi_1)}(t,\tau) + G_{nnmm}^{(\xi_2\xi_1\xi_2\xi_1)}(t,\tau)\right],$$

where for Gaussian processes

$$\begin{aligned} G_{nmpq}^{(\xi_1\xi_2\xi_3\xi_4)}(t,\tau) \;=\; & \left[b_{\xi_1\xi_2}(t,(p-n)P)b_{\xi_3\xi_4}(t+\tau,(q-m)P)\right. \\ & \left. +\; b_{\xi_1\xi_4}(t,\tau+(q-n)P)b_{\xi_2\xi_3}(t+\tau,(p-m)P-\tau)\right]. \end{aligned}$$

We introduce the functions

$$\tilde{b}_{\xi_p}(t,s-t,\tau) = b_{\xi_p}(t,s-t)b_{\xi_p}(t+\tau,s-t) + b_{\xi_p}(t,s-t+\tau)b_{\xi_p}(t+\tau,s-t-\tau),$$

$$\begin{aligned} \tilde{b}_{\xi_p\xi_q}(t,s-t,\tau) \;=\; & b_{\xi_p\xi_q}(t,s-t)b_{\xi_p\xi_q}(t+\tau,s-t) \\ & +\; b_{\xi_p\xi_q}(t,s-t+\tau)b_{\xi_p\xi_q}(t+\tau,s-t-\tau)\,. \end{aligned}$$

Using expressions (1) and (2) we represent these functions in the form of Fourier series:

$$\tilde{b}_{\xi_p}(t,s-t,\tau) = \sum_{k\in\mathbb{Z}}\tilde{B}_k^{(\xi_p)}(s-t,\tau)e^{ik\omega_0 t}, \quad (21)$$

$$\tilde{b}_{\xi_p\xi_q}(t,s-t,\tau) = \sum_{k\in\mathbb{Z}}\tilde{B}^{(\xi_p\xi_q)}(s-t,\tau)e^{ik\omega_0 t}, \quad (22)$$

where

$$\tilde{B}_k^{(\xi_p)}(s-t,\tau)=\sum_{r\in\mathbb{Z}}e^{-ir\omega_0\tau}\left[B_{k+r}^{(\xi_p)}(s-t)\bar{B}_r^{(\xi_p)}(s-t)+B_{k+r}^{(\xi_p)}(s-t-\tau)\bar{B}_r^{(\xi_p)}(s-t+\tau)\right],$$

$$\begin{aligned}\tilde{B}_k^{(\xi_p\xi_q)}(s-t,\tau)\;=\;\sum_{r\in\mathbb{Z}}e^{-ir\omega_0\tau}\Big[&B_{k+r}^{(\xi_p\xi_q)}(s-t)\bar{B}_r^{(\xi_p\xi_q)}(s-t)\\&+B_{k+r}^{(\xi_p\xi_q)}(s-t-\tau)\bar{B}_r^{(\xi_p\xi_q)}(s-t+\tau)\Big],\end{aligned}$$

here "¯" denotes complex conjugation. Then for variances we have:

$$\mathrm{V}ar\big[\hat{I}_1(t,\tau)\big]=\sum_{k\in\mathbb{Z}}\alpha_k^{(I_1)}(\tau)e^{ik\omega_0 t},\tag{23}$$

$$\mathrm{V}ar\big[\hat{D}(t,\tau)\big]=\sum_{k\in\mathbb{Z}}\alpha_k^{(D)}(\tau)e^{ik\omega_0 t}\tag{24}$$

where

$$\alpha_k^{(I_1)}(\tau)=\frac{1}{N}\left[\tilde{B}_k^{(I_1)}(0,\tau)+2\sum_{n=1}^{N-1}\left(1-\frac{n}{N}\right)\tilde{B}_k^{(I_1)}(nT,\tau)\right]\tag{25}$$

and

$$\tilde{B}_k^{(I_1)}(s-t,\tau)=\tilde{B}_k^{(\xi_1)}(s-t,\tau)+\tilde{B}_k^{(\xi_2)}(s-t,\tau)+\tilde{B}_k^{(\xi_1\xi_2)}(s-t,\tau)+\tilde{B}_k^{(\xi_2\xi_1)}(s-t,\tau).\tag{26}$$

The coefficient $\alpha_k^{(D)}(\tau)$ we find, using representations of functions

$$\tilde{b}_{\xi_p}(t,s-t,u)=b_{\xi_p}(t,s-t)b_{\xi_p}(t+u,s-t)+b_{\xi_p\xi_q}(t,s-t+\tau)b_{\xi_q\xi_p}(t+\tau,s-t-\tau),$$

$$\tilde{b}_{\xi_q\xi_p}(t,s-t,\tau)=b_{\xi_q\xi_p}(t,s-t)b_{\xi_q\xi_p}(t+\tau,s-t)+b_{\xi_p}(t,s-t+\tau)b_{\xi_q}(t+\tau,s-t-\tau)$$

in the form of Fourier series

$$\tilde{b}_{\xi_p}(t,s-t,\tau)=\sum_{k\in\mathbb{Z}}\tilde{B}_k^{(\xi_p)}(s-t,\tau)e^{ik\omega_0 t},\tag{27}$$

$$\tilde{b}_{\xi_p\xi_q}(t,s-t,\tau)=\sum_{k\in\mathbb{Z}}\tilde{B}_k^{(\xi_p\xi_q)}(s-t,\tau)e^{ik\omega_0 t}.\tag{28}$$

The formulae for Fourier coefficients we find, taking into account (1) and (2):

$$\tilde{B}_k^{(\zeta_p)}(s-t,\tau)=\sum_{r\in\mathbb{Z}} e^{ik\omega_0\tau}\left[B_{r+k}^{(\xi_p)}(s-t)\bar{B}_r^{(\xi_q)}(s-t)+B_{r+q}^{(\xi_p\xi_q)}(\tau+s-t)\bar{B}_r^{(\xi_q\xi_p)}(s-t-\tau)\right],$$

$$\tilde{B}_k^{(\zeta_p\zeta_q)}(\tau)=\sum_{r\in\mathbb{Z}} e^{-ik\omega_0\tau}\left[B_{r+k}^{(\xi_p\xi_q)}(s-t)\bar{B}_r^{(\xi_q\xi_p)}(s-t)+B_{r+q}^{(\xi_p)}(s-t+\tau)\bar{B}_r^{(\xi_q)}(s-t-\tau)\right].$$

Using these expressions, we have:

$$\alpha_k^{(D)}(\tau)=\frac{1}{N}\left[\tilde{B}_k^{(D)}(0,\tau)+2\sum_{n=1}^{N-1}\left(1-\frac{n}{N}\right)\tilde{B}_k^{(D)}(nT,\tau)\right], \tag{29}$$

where

$$\tilde{B}_k^{(D)}(s-t,\tau)=\tilde{B}_k^{(\xi_1)}(s-t,\tau)+\tilde{B}_k^{(\xi_2)}(s-t,\tau)+\tilde{B}_k^{(\xi_1\xi_2)}(s-t,\tau)-\tilde{B}_k^{(\xi_2\xi_1)}(s-t,\tau). \tag{30}$$

Taking into consideration the foregoing we now formulate the following proposition:

Proposition 2.3. *The invariant estimators* (4) *and* (5) *of Gaussian vectorial PCRP when auto- and cross-covariance functions are calculated using formulae* (6), (7) *or* (15), (16) *are consistent if conditions* (8) *are satisfied and for a finite number N their variances are defined by the Fourier series* (23) *and* (24), *which coefficients in the first approximation are given by* (25)–(30).

We note that the component of statistics (6), (7) and (15), (16) caused by previous computation of mean functions estimators were neglected above for finding (26) and (30). The variance of estimators (4), (5) and (19), (20) obtained in this approximation is dependent on mean functions of vector components. It is obvious that such dependence is undesirable and therefore the statistics (6), (7) or (15), (16) have advantage over (19), (20) when covariance invariants are calculated on the basis of experimental data.

2.2 The Estimation of Invariants Fourier Components

We analyze now the estimators of invariants Fourier components which are calculated by formulae:

$$\hat{B}_k^{(I_1)}(\tau)=\frac{1}{P}\int_0^P \hat{I}_1(t,\tau)e^{-ik\omega_0 t}dt, \tag{31}$$

$$\hat{B}_k^{(D)}(\tau)=\frac{1}{P}\int_0^P \hat{D}(t,\tau)e^{-ik\omega_0 t}dt, \tag{32}$$

where $\hat{I}_1(t,\tau)$ and $\hat{D}(t,\tau)$ are estimated using formulae (6) and (7). Taking expectation we have:

$$E\hat{B}_k^{(I_1)}(\tau) = \frac{1}{P}\int_0^P E\hat{I}_1(t,\tau)e^{-ik\omega_0 t}dt,$$

$$E\hat{B}_k^{(D)}(\tau) = \frac{1}{P}\int_0^P E\hat{D}(t,\tau)e^{-ik\omega_0 t}dt.$$

Taking into account the series (1) and (2), we conclude that the formulae for biases of estimators coincide with the expressions for Fourier coefficients of invariants estimators biases that are (11) and (12).

We assume now that the estimators of mean functions of vector components are periodical on time t. It is obvious that

$$\hat{B}_k^{(I_1)}(\tau) = \hat{B}_k^{(\xi_1)}(\tau) + \hat{B}_k^{(\xi_2)}(\tau), \hat{B}_k^{(D)}(\tau) = \hat{B}_k^{(\xi_1\xi_2)}(\tau) - \hat{B}_k^{(\xi_2\xi_1)}(\tau).$$

The estimators of covariance components for $\theta = NP$ can be rewritten in the form:

$$\hat{B}_k^{(\xi_p)}(\tau) = \frac{1}{\theta}\int_0^\theta \left[\overset{\circ}{\underset{p}{\xi}}(s)\overset{\circ}{\underset{p}{\xi}}(s+\tau) - \overset{\circ}{\underset{\xi_p}{m}}(s)\overset{\circ}{\underset{\xi_p}{m}}(s+\tau)\right]e^{-ik\omega_0 s}ds, \tag{33}$$

$$\hat{B}_k^{(\xi_p\xi_q)}(\tau) = \frac{1}{\theta}\int_0^\theta \left[\overset{\circ}{\underset{p}{\xi}}(s)\overset{\circ}{\underset{p}{\xi}}(s+\tau) - \overset{\circ}{\underset{\xi_p}{m}}(s)\overset{\circ}{\underset{\xi_p}{m}}(s+\tau)\right]e^{-ik\omega_0 s}ds, \tag{34}$$

where $\theta = NP$. Then

$$\mathrm{V}ar\left[\hat{B}_k^{(I_1)}(\tau)\right] = \frac{1}{\theta^2}\int_0^\theta\int_0^\theta \Big[\left[\tilde{b}_{\xi_1}(t,s-t,\tau) + \tilde{b}_{\xi_2}(t,s-t,\tau)\right]e^{-ik\omega_0(s-t)} + 2\tilde{b}_{\xi_1\xi_2}(t,s-t,\tau)\cos k\omega_0(s-t)\Big]dtds,$$

$$\mathrm{V}ar\left[\hat{B}_k^{(D)}(\tau)\right] = \frac{1}{\theta^2}\int_0^\theta\int_0^\theta \Big[\left[\tilde{b}_{\zeta_1}(t,s-t,\tau) + \tilde{b}_{\zeta_2}(t,s-t,\tau)\right]e^{ik\omega_0(s-t)} - 2b_{\zeta_1\zeta_2}(t,s-t,\tau)\cos k\omega_0(s-t)\Big]dtds.$$

After introducing a new variable $\tau_1 = s - t$ and changing the order of integration we find:

$$\mathrm{V}ar\left[\hat{B}_k^{(I_1)}(\tau)\right] = \frac{2}{\theta^2}\int_0^{\theta}\int_0^{\theta-\tau_1}\left[\tilde{b}_{\xi_1}(t,\tau_1,\tau) + \tilde{b}_{\xi_1}(t,\tau_1,\tau) + \tilde{b}_{\xi_1\xi_2}(t,\tau_1,\tau) + b_{\xi_2\xi_1}(t,\tau_1,\tau)\right]\cos k\omega_0\tau_1 dt d\tau_1,$$

$$\mathrm{V}ar\left[\hat{B}_k^{(D)}(\tau)\right] = \frac{1}{\theta^2}\int_0^{\theta}\int_0^{\theta-\tau_1}\left[\tilde{b}_{\zeta_1}(t,\tau_1,\tau) + \tilde{b}_{\zeta_2}(t,\tau_1,\tau) - \tilde{b}_{\zeta_1\zeta_2}(t,\tau_1,\tau) - \tilde{b}_{\zeta_2\zeta_1}(t,\tau_1,\tau)\right]\cos k\omega_0\tau_1 dt d\tau_1.$$

If to substitute the series (21), (22) and (27), (28) into these expressions, and integrate with respect to t, in the first approximation we obtain:

$$\mathrm{V}ar\left[\hat{B}_k^{(I_1)}(\tau)\right] = \frac{2}{\theta}\int_0^{\theta}\left(1-\frac{\tau_1}{\theta}\right)\tilde{B}_0^{(I_1)}(\tau_1,\tau)\cos k\omega_0\tau_1 d\tau_1, \tag{35}$$

$$\mathrm{V}ar\left[\hat{B}_k^{(D)}(\tau)\right] = \frac{2}{\theta}\int_0^{\theta}\left(1-\frac{\tau_1}{\theta}\right)\tilde{B}_0^{(D)}(\tau_1,\tau)\cos k\omega_0\tau_1 d\tau. \tag{36}$$

Thus, we can pose the following proposition:

Proposition 2.4. *The estimators* (31) *and* (32) *if conditions* (8) *are satisfied are asymptotically unbiased and consistent for Gaussian vectorial PCRP and for a finite* $\theta = NP$ *their biases and variances are defined in the first approximation by formulae* (11), (12) *and* (35), (36).

It follows from relationships (35) and (36) that variances of estimators for invariants Fourier components depend in the main on quantities, which are time average values of covariance function for products of random vector components. The frequency of cosine weight function depends only on the number of invariant Fourier component which is estimated.

3 Component Estimators of Covariance Invariants

3.1 The Biases of Estimators

Coherent estimation which are obtained on a basis of averaging of the samples, taken via the non-stationary period P, uses only single values of the processes $\xi_1(t)$ and $\xi_2(t)$ realizations at the period.

Let us now consider component statistics

$$\hat{I}_1(t,\tau) = \sum_{k=-N_2}^{N_2} \hat{B}_k^{(I_1)}(\tau) e^{ik\omega_0 t}, \tag{37}$$

$$\hat{D}(t,\tau) = \sum_{k=-N_2}^{N_2} \hat{B}_k^{(D)}(\tau) e^{ik\omega_0 t}, \tag{38}$$

where $\hat{B}_k^{(I_1)}(\tau)$ and $\hat{B}_k^{(D)}(\tau)$ are defined by formulae (31), (32) and the covariance component estimator are formed on a basis of integral transformations:

$$\hat{B}_k^{(\xi_p)}(\tau) = \frac{1}{\theta}\int_0^\theta \left[\xi_p(s) - \hat{m}_{\xi_p}(s)\right]\left[\xi_p(s+\tau) - \hat{m}_{\xi_p}(s+\tau)\right] e^{-ik\omega_0 s} ds, \tag{39}$$

$$\hat{B}_k^{(\xi_p\xi_q)}(\tau) = \frac{1}{\theta}\int_0^\theta \left[\xi_p(s) - \hat{m}_{\xi_p}(s)\right]\left[\xi_p(s+\tau) - \hat{m}_{\xi_q}(s+\tau)\right] e^{-ik\omega_0 s} ds, \tag{40}$$

$$\hat{m}_{\xi_p}(t) = \sum_{k=-N_1}^{N_1} e^{ik\omega_0 t}\left[\frac{1}{\theta}\int_0^\theta \xi_p(s) e^{-ik\omega_0 s} ds\right], \; p = \overline{1,2}. \tag{41}$$

Here $\theta = NP$ is a length of realization section, N_1 and N_2 are amounts of harmonic components in trigonometric polynomials (37), (38) and (41). Number N_1 and N_2 can be defined using results of coherent processing of experimental data.

Proposition 3.1. *The component estimators of covariance invariants which are defined by formulae* (37), (38) *and* (31), (32), (39)–(41) *if conditions* (8) *are satisfied, are asymptotically unbiased and for a finite* θ *their biases are:*

$$\varepsilon\left[\hat{I}_1(t,\tau)\right] = \sum_{k=-N_2}^{N_2} \varepsilon_k^{(I_1)}(\tau) e^{ik\omega_0 t},$$

$$\varepsilon\left[\hat{D}(t,\tau)\right] = \sum_{k=-N_2}^{N_2} \varepsilon_k^{(D)}(\tau) e^{ik\omega_0 t},$$

where

$$\varepsilon_k^{(I_1)}(\tau) = \frac{1}{\theta}\int_0^\theta \left(1 - \frac{\tau_1}{\theta}\right)\left[\left[B_k^{(I_1)}(\tau - \tau_1) e^{-ik\omega_0\tau_1} + B_k^{(I_1)}(\tau + \tau_1) e^{-ik\omega_0\tau_1}\right] h(N_1, \tau_1) + B_k^{(I_1)}(\tau_1) h_k^{(1)}(N_1, \tau, \tau_1)\right] d\tau_1, \tag{42}$$

$$\varepsilon_k^{(D)}(\tau) = \frac{1}{\theta}\int_0^\theta \left(1-\frac{\tau_1}{\theta}\right)\left[\left[B_k^{(D)}(\tau+\tau_1)e^{-ik\omega_0\tau_1} - B_k^{(D)}(\tau-\tau_1)e^{ik\omega_0\tau_1}\right]h(N_1,\tau_1) + B_k^{(D)}(\tau_1)h_k^{(2)}(N_1,\tau,\tau_1)\right]d\tau_1, \tag{43}$$

and also

$$h(N_1,u_1) = 1+2\sum_{l=1}^{N_1}\cos l\omega_0 u_1,$$

$$h_k^{(1)}(N_1,\tau,\tau_1) = \sum_{l\in\bar{R}} e^{il\omega_0\tau}\left(e^{-il\omega_0\tau_1}+e^{i(l-k)\omega_0\tau_1}\right),$$

$$h_k^{(2)}(N_1,\tau,\tau_1) = \sum_{l\in\bar{R}} e^{il\omega_0\tau}\left(e^{-il\omega_0\tau_1}-e^{i(l-k)\omega_0\tau_1}\right),$$

$$\bar{R} = \{-N_1,\ldots,N_1\} - R, R = \{-N_1,\ldots,N_1\}\cap\{k-N_1,\ldots,k+N_1\}.$$

Proof. The expressions (39) and (40) in case $p = q$ are identical, thus we consider only (40) as more general. Rewrite this expression in the form:

$$\hat{B}_k^{(\xi_p\xi_q)}(\tau) = \frac{1}{\theta}\int_0^\theta \left[\overset{\circ}{\xi}_p(t)\overset{\circ}{\xi}_q(t+\tau) - \overset{\circ}{\hat{m}}_{\xi_p}(t)\overset{\circ}{\xi}_q(t+\tau) - \overset{\circ}{\xi}_p(t)\overset{\circ}{\hat{m}}_{\xi_q}(t+\tau) + \overset{\circ}{\hat{m}}_{\xi_p}(t)\overset{\circ}{\hat{m}}_{\xi_q}(t+\tau)\right]e^{-ik\omega_0 t}dt, \tag{44}$$

where

$$\overset{\circ}{\hat{m}}_{\xi_p}(t) = \sum_{k=-N_1}^{N_1} e^{ik\omega_0 t}\left[\frac{1}{\theta}\int_0^\theta \overset{\circ}{\xi}_p(s)e^{-ik\omega_0 s}ds\right],$$

$$\overset{\circ}{\xi}_p(s) = \xi_p(s) - m_p(s).$$

Let us introduce the functions

$$g_k(N_1,s_1,s_2) = e^{-ik\omega_0 s_2}\sum_{l=-N_1}^{N_1} e^{il\omega_0(s_2-s_1)},$$

$$h_k(N_1,s_1,s_2,\tau) = e^{-ik\omega_0 s_1}\sum_{l=-N_1}^{N_1} e^{il\omega_0(s_1-s_2+\tau)}.$$

Then the second and third terms in (44) can be written in the form:

$$
\begin{aligned}
&-\frac{1}{\theta}\int_0^\theta \left[\overset{\circ}{\hat{m}}_{\xi_p}(t)\overset{\circ}{\xi}_q(t+\tau)+\overset{\circ}{\xi}_p(t)\,\underset{\xi_p}{\hat{m}}(t+\tau)\right]e^{-ik\omega_0 t}dt \\
&=-\frac{1}{\theta}\int_0^\theta\int_0^\theta \Big[\overset{\circ}{\xi}_p(s_1)\overset{\circ}{\xi}_q(s_2+\tau)g_k(N_1,s_1,s_2) \\
&\qquad\qquad +\overset{\circ}{\xi}_p(s_1)\overset{\circ}{\xi}_q(s_2)h_1(N_1,s_1,s_2,\tau)\Big]ds_1ds_2.
\end{aligned}
\tag{45}
$$

For the fourth term taking into account the expression $\frac{1}{\theta}\int_0^\theta e^{i(k-l+m)t}dt=\delta_{k,l-m}$, where $\delta_{k,l-m}$ is Kronecker delta, we have

$$
\frac{1}{\theta}\int_0^\theta \overset{\circ}{\hat{m}}(t)\,\overset{\circ}{\hat{m}}(t+u)e^{-ik\omega_0 t}dt=\frac{1}{\theta^2}\int_0^\theta\int_0^\theta \overset{\circ}{\xi}(s_1)\,\overset{\circ}{\xi}(s_2)p_k(N_1,s_1,s_2,\tau)ds_1ds_2, \tag{46}
$$

where

$$
p_k(N_1,s_1,s_2,\tau)=e^{-ik\omega_0 s}\sum_{l\in\bar{R}}e^{il\omega_0(s_1-s_2+\tau)}.
$$

Taking into consideration equalities (45) and (46) after ensemble averaging we obtain:

$$
\begin{aligned}
E\hat{B}_k^{\left(\xi_p\xi_q\right)}(\tau)=B_k^{\left(\xi_p\xi_q\right)}(\tau)\;-\;&\frac{1}{\theta^2}\int_0^\theta\int_0^\theta \Big[b_{\xi_p\xi_q}(s_1,s_2-s_1+\tau)g_k(N_1,s_1,s_2) \\
&+b_{\xi_p\xi_q}(s_2,s_2-s_1)h_k(N_1,s_1,s_2,\tau)\Big]ds_1ds_2,
\end{aligned}
\tag{47}
$$

where

$$
h_k(N_1,s_1,s_2,\tau)=e^{-ik\omega_0 s_2}\sum_{r\in\bar{R}}e^{ir\omega_0(s_2-s_1-\tau)}.
$$

After substituting Fourier series (2) into (47) and integrating on t we have

$$
\begin{aligned}
EB_k^{\left(\xi_p\xi_q\right)}(\tau)\;=\;B_k^{\left(\xi_p\xi_q\right)}(\tau)-\frac{1}{\theta}\sum_{r=-N_2}^{N_2}\Bigg[\int_0^\theta \Big[\Big[B_r^{\left(\xi_p\xi_q\right)}(\tau_1-\tau)e^{ir\omega_0\tau}+B_r^{\left(\xi_p\xi_q\right)}(\tau_1+\tau)e^{-ir\omega_0\tau_1}\Big] \\
\times h(N_1,\tau_1)+B_k^{\left(\xi_p\xi_q\right)}(\tau_1)h_k^{(1)}(N_1,\tau,\tau_1)\Big]f_{r,k}(0,\theta-\tau_1)d\tau_1\Bigg].
\end{aligned}
$$

Here with $f_{r,k}(0,\theta-u_1)=\theta^{-1}\int\limits_0^{\theta-u_1} e^{i(k-r)\omega_0 t}dt$. Using expressions $E\hat{B}_k^{(I_1)}(\tau)=E\hat{B}_k^{(\xi_1)}(\tau)+E\hat{B}_k^{(\xi_2)}(\tau)$ and $E\hat{B}_k^{(I_1)}(\tau)=E\hat{B}_k^{(\xi_1\xi_2)}(\tau)-E\hat{B}_k^{(\xi_2\xi_1)}(\tau)$ we come in the first approximation to formulae (42) and (43) □.

When $k=0$ $\bar{R}$ is an empty set and

$$\varepsilon\left[\hat{B}_0^{(I_1)}(\tau)\right]=-\int\limits_0^{\theta}\left(1-\frac{\tau_1}{\theta}\right)\left[B_0^{(I_1)}(\tau-\tau_1)+B_0^{(I_1)}(\tau_1+\tau)\right]h(N_1,\tau_1)d\tau_1, \tag{48}$$

$$\varepsilon\left[\hat{B}_0^{(D)}(\tau)\right]=-\int\limits_0^{\theta}\left(1-\frac{\tau_1}{\theta}\right)\left[B_0^{(D)}(\tau-\tau_1)+B_0^{(D)}(\tau_1+\tau)\right]h(N_1,\tau_1)d\tau_1. \tag{49}$$

These quantities define time averaged biased of component invariant estimators. As following from (42), (43) and (48), (49) the estimators biases in the first approximation depend only on the value of the same component being estimated.

3.2 The Variances for the Invariant Estimators

Let us analyze the variances of invariant estimators represented in the form of trigonometric polynomials (37) and (38) which are defined by expressions:

$$\mathrm{V}ar\left[\hat{I}_1(t,\tau)\right]=E\left[\hat{I}_1(t,\tau)-E\hat{I}_1(t,\tau)\right]^2=\sum_{k,l=-N_2}^{N_2}R_{lk}^{(I_1)}(\tau)e^{i(l-k)\omega_0 t}, \tag{50}$$

$$\mathrm{V}ar\left[\hat{D}(t,\tau)\right]=E\left[\hat{D}(t,\tau)-E\hat{D}(t,\tau)\right]^2=\sum_{k,l=-N_2}^{N_2}R_{lk}^{(D)}(\tau)e^{i(l-k)\omega_0 t}. \tag{51}$$

Functions $R_{lk}^{(I_1)}(\tau)$ and $R_{lk}^{(D)}(\tau)$ are defined by covariances of estimators for invariant components:

$$R_{lk}^{(I_1)}(\tau)=E\hat{B}_l^{(I_1)}(\tau)\overline{\hat{B}_k^{(I_1)}}(\tau)-E\hat{B}_l^{(I_1)}(\tau)\overline{\hat{B}_k^{(I_1)}}(\tau),$$

$$R_{lk}^{(D)}(\tau)=E\hat{B}_l^{(D)}(\tau)\overline{\hat{B}_k^{(D)}}(\tau)-E\hat{B}_l^{(D)}(\tau)\overline{\hat{B}_k}(\tau).$$

We can represent the series (50), (51) in the form:

$$\mathrm{V}ar\left[\hat{I}_1(t,\tau)\right]=\sum_{n=-2N_2}^{2N_2}\alpha_n^{(I_1)}(\tau)e^{ik\omega_0 t}, \tag{52}$$

$$\mathrm{Var}\left[\hat{D}(t,\tau)\right] = \sum_{n=-2N_2}^{2N_2} \alpha_n^{(D)}(\tau) e^{ik\omega_0 t}, \tag{53}$$

where

$$\alpha_n^{(I_1,D)}(\tau) = \sum_{k\in L} R_{n+k,k}^{(I_1,D)}(\tau)$$

and $L = \{-N_2 - n, N_2\}$ for $n \le 0$ and $L = \{-N_2, N_2 - n\}$ for $n > 0$. For Gaussian processes in the first approximation we have

$$\hat{R}_{lk}^{(I_1)}(\tau) = \frac{1}{\theta^2}\int_0^\theta\int_0^\theta \left[\tilde{b}_{\xi_1}(t,s-t,\tau) + \tilde{b}_{\xi_2}(t,s-t,\tau) + \tilde{b}_{\xi_1\xi_2}(t,s-t,\tau) + \tilde{b}_{\xi_2\xi_1}(t,s-t,\tau)\right] e^{i\omega_0(ks-lt)}dtds,$$

$$\hat{R}_{lk}^{(D)}(\tau) = \frac{1}{\theta^2}\int_0^\theta\int_0^\theta \left[\tilde{b}_{\zeta_1}(t,s-t,\tau) + \tilde{b}_{\zeta_2}(t,s-t,\tau) + \tilde{b}_{\zeta_1\zeta_2}(t,s-t,\tau) + \tilde{b}_{\zeta_2\zeta_1}(t,s-t,\tau)\right] e^{i\omega_0(ks-lt)}dtds.$$

After introducing a new variable $\tau_1 = s - t$ and changing the order of integration we obtain:

$$R_{k+n,k}^{(I_1)}(\tau) = \frac{1}{\theta^2}\int_0^\theta\int_0^{\theta-\tau_1} \left[\left[\tilde{b}_{\xi_1}(t,\tau_1,\tau) + \tilde{b}_{\xi_2}(t,\tau_1,\tau)\right]\left(e^{i\omega_0(k+n)\tau_1} + e^{i\omega_0 k\tau_1}\right) + \tilde{b}_{\xi_1\xi_2}(t,\tau_1,\tau)e^{ik\omega_0\tau_1} + \tilde{b}_{\xi_2\xi_1}(t,\tau_1,\tau)e^{-i(k+n)\omega_0\tau_1}\right] e^{-ik\omega_0 t}dtd\tau_1,$$

$$R_{k+n,k}^{(D)}(\tau) = \frac{1}{\theta^2}\int_0^\theta\int_0^{\theta-\tau_1} \left[\left[\tilde{b}_{\zeta_1}(t,\tau_1,\tau) + \tilde{b}_{\zeta_2}(t,\tau_1,\tau)\right]\left[e^{-i(k+n)\omega_0\tau_1} + e^{i\omega_0 k\tau_1}\right] - \tilde{b}_{\zeta_1\zeta_2}(t,\tau_1,\tau)e^{ik\omega_0\tau_1} - \tilde{b}_{\zeta_2\zeta_1}(t,\tau_1,\tau)e^{-i(k+n)\omega_0\tau_1}\right] e^{-ik\omega_0 t}dtd\tau_1.$$

If to substitute the Fourier series (21), (22) and (27), (28) to these expressions and integrate out, then in the first approximation:

$$\alpha_n^{(I_1)}(\tau) = \frac{1}{\theta}\int_0^\theta \left(1 - \frac{u_1}{\theta}\right)\tilde{B}_n^{(I_1)}(\tau_1,\tau)\sum_{k\in L}\left[e^{-i\omega_0(k+n)\tau_1} + e^{ik\omega_0\tau_1}\right]d\tau_1, \tag{54}$$

$$\alpha_n^{(D)}(\tau) = \frac{1}{\theta}\int_0^\theta \left(1 - \frac{u_1}{\theta}\right)\tilde{B}_n^{(D)}(\tau_1,\tau)\sum_{k\in L}\left[e^{-i\omega_0(k+n)\tau_1} + e^{ik\omega_0\tau_1}\right]d\tau_1. \tag{55}$$

Summarizing the statements written above we form proposition:

Proposition 3.2. *For Gaussian vectorial PCRP the component invariant estimators* (37), (38) *if condition* (8) *are satisfied are consistent and for a finite* θ *their variances are defined by formulae* (52)–(55).

Taking into account that

$$\lim_{N_2\to\infty}\sum_{k\in L}\left(e^{ik\omega_0\tau}+e^{-i(k+n)\omega_0\tau}\right)=P\left(1+e^{-in\omega_0\tau}\right)\sum_{r\in\mathbb{Z}}\delta(\tau+rP)$$

for $\alpha_n^{(I_1)}(\tau)$ and $\alpha_n^{(D)}(\tau)$ we obtain the formulae which coincide with expressions for the Fourier coefficients of the variances of coherent invariant estimators, that is (17) and (18).

As it has been noted above, the component invariant estimators are formed on the basis of initial information about the amount of harmonic components in Fourier series for estimated invariants. Using the coherent method, we allow implicitly the existence of a finite number of components. Just this feature distinguishes the coherent method from the component one, and we can anticipate a better efficiency for the component estimators against coherent ones in the case of fast damping of correlations between processes values.

4 Discrete Estimation

We form discrete estimators of invariant components replacing integral transformations (31) and (32) by integral sums:

$$\hat{B}_k^{(I_1)}(jh)=\frac{1}{M+1}\sum_{n=0}^{M}\hat{I}_1(nh,jh)e^{-ik\frac{2\pi}{M+1}n},\tag{56}$$

$$\hat{B}_k^{(D)}(jh)=\frac{1}{M+1}\sum_{n=0}^{M}\hat{D}(nh,jh)e^{-ik\frac{2\pi}{M+1}n}.\tag{57}$$

Here $\hat{I}_1(nh,jh)=\hat{b}_{\xi_1}(nh,jh)+\hat{b}_{\xi_2}(nh,jh)$, $\hat{D}(nh)=\hat{b}_{\xi_1\xi_2}(nh,jh)-\hat{b}_{\xi_2\xi_1}(nh,jh)$, $h=\frac{P}{M+1}$, M – natural number, and

$$\hat{b}_{\xi_p}(nh,jh)=\tfrac{1}{N}\sum_{k=0}^{N-1}\xi_p[(n+j)h+k(M+1)h]\xi_p[nh+k(M+1)h]\\ -\hat{m}_{\xi_p}[(n+j)h]\hat{m}_{\xi_p}[nh],$$

$$\hat{b}_{\xi_p\xi_q}(nh,jh)=\tfrac{1}{N}\sum_{k=0}^{N-1}\xi_q[(n+j)h+k(M+1)h]\xi_p[nh+k(M+1)h]\\ -\hat{m}_{\xi_p}[(n+j)h]\hat{m}_{\xi_p}[nh].$$

Taking into account Fourier series (1), (2) and expressions

$$E\hat{b}_{\xi_p}(nh,jh) = b_{\xi_p}(nh,jh) - \frac{1}{N}\sum_{p=-N+1}^{N-1}\left(1-\frac{|p|}{N}\right)b_{\xi_p}[nh,(j+(M+1))h],$$

$$E\hat{b}_{\xi_p\xi_q}(nh,jh) = b_{\xi_p\xi_q}(nh,jh) - \frac{1}{N}\sum_{p=-N+1}^{N-1}\left(1-\frac{|p|}{N}\right)b_{\xi_p\xi_q}[nh,(j+(M+1))h],$$

for biases of (56) and (57) we obtain

$$\varepsilon\left[\hat{B}_k^{(I_1,D)}(jh)\right] = \varepsilon_0\left[\hat{B}_k^{(I_1,D)}(jh)\right] + \varepsilon_N\left[\hat{B}_k^{(I_1,D)}(jh)\right],$$

where

$$\varepsilon_0\left[\hat{B}_k^{(I_1,D)}(jh)\right] = \sum_{\substack{q\in\mathbb{Z},\\ q\neq 0}} B_{q(M+1)}^{(I_1,D)}(jh),$$

$$\varepsilon_N\left[\hat{B}_k^{(I_1,D)}(jh)\right] = -\frac{1}{N}\sum_{p=-N+1}^{N-1}\left(1-\frac{|p|}{N}\right)B_{k+q(M+1)}[(i+p(M+1))h].$$

Thus we have

Fact 4.1. *Aliasing effects of the first and the second kinds* (Javorśkyj 1984) *significantly affect the values of biases of invariant component estimators in the cases when invariants $I_1(t,\tau)$ and $D(t,\tau)$ are represented by _ infinite Fourier series. These effects are avoided when the covariance invariants are represented by trigonometric polynomials*

$$I_1(t,\tau) = \sum_{k=-N_2}^{N_2} B_k^{(I_1)}(\tau)e^{ik\omega_0 t},\ D(t,\tau) = \sum_{k=-N_2}^{N_2} B_k^{(D)}(\tau)e^{ik\omega_0 t},$$

and the sampling interval satisfies inequality $h \le \frac{P}{2N_2+1}$.

Proposition 4.1. *Statistics* (56) *and* (57) *are asymptotically unbiased estimators for finite number components if condition* (8) *and inequality $h \le \frac{P}{2N_2+1}$ are satisfied and for a finite number N their biases are:*

$$\varepsilon\left[\hat{B}_k^{(I_1)}(jh)\right] = -\frac{1}{N}\sum_{p=-N+1}^{N-1}\left(1-\frac{|p|}{N}\right)B_k^{(I_1)}[(j+p(M+1))h],$$

$$\varepsilon\left[\hat{B}_k^{(D)}(jh)\right] = -\frac{1}{N}\sum_{p=-N+1}^{N-1}\left(1-\frac{|p|}{N}\right)B_k^{(D)}[(j+p(M+1))h].$$

We represent the estimators of covariance invariants for $\forall\, t \in [0,P]$ in the forms:

$$\hat{I}(t,jh) = \sum_{k=-N_2}^{N_2} e^{\mathrm{i}k\omega_0 t}\left[\frac{1}{M+1}\sum_{h=0}^{M}\hat{I}(nh,jh)e^{-\mathrm{i}k\frac{2\pi}{M+1}n}\right], \tag{58}$$

$$\hat{D}(t,jh) = \sum_{k=-N_2}^{N_2} e^{\mathrm{i}k\omega_0 t}\left[\frac{1}{M+1}\sum_{h=0}^{M}\hat{I}(nh,jh)e^{-\mathrm{i}k\frac{2\pi}{M+1}n}\right], \tag{59}$$

and prove Proposition.

Proposition 4.2. *Statistics* (58) *and* (59) *are asymptotically unbiased estimators for covariance invariants* $\forall t \in [0,P]$ *if conditions* (8) *are satisfied and* $h \le \frac{P}{2N_2+1}$ *and for* $M = 2N_2$ *they coincide with Shannon-Kotelnikov formulae, i.e.*

$$\hat{I}_1(t,jh) = \sum_{l=0}^{2N_2}\hat{I}(lh,jh)\varphi_l(t), \tag{60}$$

$$D(t,jh) = \sum_{l=0}^{2N_2}\hat{D}(lh,jh)\varphi_l(t), \tag{61}$$

where $\hat{I}(lh,jh)$ *and* $\hat{D}(lh,jh)$ *are coherent estimators for covariance invariants and*

$$\varphi_l(t) = \frac{\sin(2N_2+1)\frac{\pi}{P}(t-lh)}{(2N_2+1)\sin\frac{\pi}{P}(t-lh)}.$$

Proof. Taking into account Propositions 2.1 and 2.2 we have

$$\lim_{N\to\infty} E\hat{I}(t,jh) = I(t,jh),$$

$$\lim_{N\to\infty} E\hat{D}(t,jh) = D(t,jh).$$

So, estimators (60) and (61) are asymptotically unbiased. Using the formula for the sum of finite geometric sequence, we obtain

$$\sum_{k=-N_2}^{N_2} e^{\mathrm{i}k\frac{2\pi}{P}(t-nh)} = \frac{\sin(2N_2+1)\frac{\pi}{P}(t-nh)}{\sin\frac{\pi}{P}(t-nh)}.$$

After transformations we come to expressions (60) and (61) □.

We will analyze the variances of estimators (56), (57) neglecting the components caused by previous computation of mean functions of processes $\xi_1(t)$ and $\xi_2(t)$. These components lead to appearance of additional summands in the expressions of variances that are of higher order of smallness.

Therefore we write

$$\hat{B}_k^{\left(\overset{\circ}{\xi}_p\right)}(jh) = \frac{1}{K}\sum_{n=0}^{K-1} \overset{\circ}{\xi}_p(nh)\, \overset{\circ}{\xi}_p[(n+j)h] e^{-\mathrm{i}k\frac{2\pi}{M+1}n},$$

$$\hat{B}_k^{\left(\overset{\circ}{\xi}_p \overset{\circ}{\xi}_q\right)}(jh) = \frac{1}{K}\sum_{n=0}^{K-1} \overset{\circ}{\xi}_p(nh)\, \overset{\circ}{\xi}_q[(n+j)h] e^{-\mathrm{i}k\frac{2\pi}{M+1}n},$$

where $K = N(M+1)$. Then

$$\begin{aligned} Var\left[\hat{B}_k^{(I_1)}(jh)\right] &= \frac{1}{K^2}\sum_{m,n=0}^{K-1} \Big[\big[\tilde{b}_{\xi_1}(nh,(m-n)h,jh) + \tilde{b}_{\xi_2}(nh,(m-n)h,jh)\big] \\ &\quad \times e^{\mathrm{i}k\frac{2\pi}{M+1}(m-n)} + 2\tilde{b}_{\xi_1\xi_2}(nh,(m-n)h,jh)\cos k\frac{2\pi}{M+1}(m-n)\Big], \\ Var\left[\hat{B}_k^{(D)}(jh)\right] &= \frac{1}{K^2}\sum_{m,n=0}^{K-1} \Big[\big[\tilde{b}_{\zeta_1}(nh,(m-n)h,jh) + \tilde{b}_{\zeta_2}(nh,(m-n)h,jh)\big] \\ &\quad \times e^{\mathrm{i}k\frac{2\pi}{M+1}(m-n)} + 2\tilde{b}_{\zeta_1\zeta_2}(nh,(m-n)h,jh)\cos k\frac{2\pi}{M+1}(m-n)\Big]. \end{aligned}$$

After transformation we find

$$Var\left[\hat{B}_k^{(I_1)}(jh)\right] = \frac{1}{K}\left[\hat{B}^{(I_1)}_{q(M+1)}(0,jh) + 2\sum_{p=1}^{K-1}\left(1-\frac{p}{K}\right)\hat{B}^{(I_1)}_{q(M+1)}(ph,jh)\cos k\frac{2\pi}{M+1}p\right],$$

$$Var\left[\hat{B}_k^{(D)}(jh)\right] = \frac{1}{K}\left[\hat{B}^{(D)}_{q(M+1)}(0,jh) + 2\sum_{p=1}^{K-1}\left(1-\frac{p}{K}\right)\hat{B}^{(D)}_{q(M+1)}(ph,jh)\cos k\frac{2\pi}{M+1}p\right].$$

Thus we have

Proposition 4.3. *For Gaussian vectorial PCRP the discrete estimators of invariant components (56)–(57) if conditions (8) are satisfied and* $h \le \frac{P}{2N_2+1}$ *are consistent and for a finite number K their variances are defined by expressions:*

$$Var\left[\hat{B}_k^{(I_1)}(jh)\right] = \frac{1}{K}\left[\hat{B}^{(I_1)}_{0}(0,jh) + 2\sum_{p=1}^{K-1}\left(1-\frac{p}{K}\right)\hat{B}^{(I_1)}_{0}(ph,jh)\cos k\frac{2\pi}{M+1}p\right],$$

$$\mathrm{V}ar\left[\hat{B}_k^{(D)}(jh)\right] = \frac{1}{K}\left[\hat{B}_0^{(D)}(0,jh) + 2\sum_{p=1}^{K-1}\left(1-\frac{p}{K}\right)\hat{B}_0^{(D)}(ph,jh)\cos k\frac{2\pi}{M+1}p\right].$$

The variances of the covariance invariant estimators (60) and (61) are defined by expressions:

$$\mathrm{V}ar\left[\hat{I}(t,jh)\right] = \sum_{r=-2N_2}^{2N_2}\tilde{\alpha}_r^{(I_1)}(jh)e^{ir\omega_0 t}, \tag{62}$$

$$\mathrm{V}ar\left[\hat{D}(t,jh)\right] = \sum_{r=-2N_2}^{2N_2}\tilde{\alpha}_r^{(D)}(jh)e^{ir\omega_0 t}, \tag{63}$$

where

$$\begin{aligned}\tilde{\alpha}_r^{(I_1)}(jh) &= \frac{1}{K}\left[(2N_2+|r|+1)\tilde{B}_r^{(I_1)}(0,jh) + \sum_{p=1}^{K-1}\left(1-\frac{p}{K}\right)\tilde{B}_r^{(I_1)}(ph,jh)\right.\\ &\left.\times\sum_{k\in L}\left(e^{\mathrm{i}(k+r)\frac{2\pi}{M+1}p} + e^{\mathrm{i}k\frac{2\pi}{M+1}p}\right)\right],\end{aligned} \tag{64}$$

$$\begin{aligned}\tilde{\alpha}_r^{(D)}(jh) &= \frac{1}{K}\left[(2N_2+|r|+1)\tilde{B}_r^{(D)}(0,jh) + \sum_{p=1}^{K-1}\left(1-\frac{p}{K}\right)\tilde{B}_r^{(D)}(ph,jh)\right.\\ &\left.\times\sum_{k\in L}\left(e^{\mathrm{i}(k+r)\frac{2\pi}{M+1}p} + e^{\mathrm{i}k\frac{2\pi}{M+1}p}\right)\right],\end{aligned} \tag{65}$$

if $h \le \frac{P}{2N_2+1}$. From above expression we pose

Proposition 4.4. *For Gaussian vectorial PCRP estimators of covariance invariants which are defined by interpolation formulae* (58) *and* (59) *are consistent if conditions* (8) *are satisfied and* $h \le \frac{P}{2N_2+1}$ *and for a finite number K their variances are defined by expressions* (64) *and* (65).

Formulae obtained here can be used for comparison of the effectiveness of discrete and continuous estimators. Obviously that comparative analysis of estimators should be carried out for those sampling intervals at which the aliasing effect of both first and second kinds are absent.

5 The Covariance Invariant Estimators for Vectorial Amplitude-Modulated Signals

We consider one case of vectorial PCRP when its components are amplitude-modulated, exact $\xi_1(t) = \mu(t)\cos\omega_0 t$, $\xi_2(t) = v(t)\sin\omega_0 t$ at that

$$E\mu(t) = m_\mu,\ E\,\mathring{\mu}(t)\,\mathring{\mu}(t+\tau) = R_\mu(\tau),\ \mathring{\mu}(t) = \mu(t) - m_\mu,$$

$$Ev(t) = m_v,\ E\,\mathring{v}(t)\,\mathring{v}(t+\tau) = R_v(\tau),\ \mathring{v}(t) = v(t) - m_v.$$

Expression for auto- and cross-covariance function of the components $\xi_1(t)$ and $\xi_2(t)$ contains the zeroth and the second harmonic components only. Invariants $I_1(t,\tau)$ and $D(t,\tau)$ are defined with relations

$$I_1(t,\tau) = \sum_{k\in 0,\pm 2} B_k^{(I)}(\tau) e^{ik\omega_0 t},$$

$$D(t,\tau) = \sum_{k\in 0,\pm 2} B_k^{(D)}(\tau) e^{ik\omega_0 t},$$

where

$$B_0^{(I)}(\tau) = \frac{1}{2}\left[R_\mu(\tau) - R_v(\tau)\right]\cos\omega_0\tau,$$
$$B_2^{(I)}(\tau) = \frac{1}{4}\left[R_\mu(\tau) - R_v(\tau)\right]e^{i\omega_0\tau},\ B_{-2}^{(I)}(\tau) = \bar{B}_2^{(I)}(\tau),$$
$$B_0^{(D)}(\tau) = \frac{1}{2}\left[R_{\mu v}(\tau) - R_{v\mu}(\tau)\right]\sin\omega_0\tau,$$
$$B_2^{(D)}(\tau) = -\frac{i}{4}\left[R_{\mu v}(\tau) - R_{v\mu}(\tau)\right]e^{i\omega_0\tau},\ B_{-2}^{(D)}(\tau) = \bar{B}_2^{(D)}(\tau),$$

We suppose

$$R_\mu(\tau) = A_1 e^{-\alpha_1|\tau|},\ R_v(\tau) = A_2 e^{-\alpha_2|\tau|},\ R_{\mu v}(\tau) = Ae^{-\alpha|\tau|}$$

and introduce function

$$r_l(\alpha) = \frac{1}{T}\int_0^{NT}\left(1 - \frac{\tau}{NT}\right)e^{-\alpha\tau}\cos l\omega_0\tau \mathrm{d}\tau,$$

$$\tilde{r}_l(\alpha,\tau_1) = \frac{1}{T}\int_0^{NT}\left(1 - \frac{\tau_1}{NT}\right)e^{-\alpha(|\tau+\tau_1| + |\tau-\tau_1||)}\cos l\omega_0\tau_1 \mathrm{d}\tau_1,$$

$$r_l^*(\alpha_1,\alpha_2,\tau) = \frac{1}{T}\int_0^{NT}\left(1 - \frac{\tau_1}{NT}\right)e^{-(\alpha_1|\tau+\tau_1| + \alpha_2|\tau-\tau_1|)}\cos l\omega_0\tau_1 \mathrm{d}\tau_1.$$

Then zeroth components of expressions (23) and (24), defining time averaged values of invariant estimator variances are the following

$$\alpha_0^{(I)}(\tau) = \frac{1}{4N}\left[\sum_{k=1}^{2} A_k^2[2[r_0(2\alpha_k) + \tilde{r}_0(\alpha_k,\tau)] + 3[r_2(2\alpha_k) + \tilde{r}_2(\alpha_k,\tau)]\right.$$
$$+ [r_0(2\alpha_k) + \tilde{r}_0(\alpha_k,\tau) + 2[r_2(2\alpha_k) + \tilde{r}_2(\alpha_k,\tau)]]\cos 2\omega_0\tau + r_4(2\alpha_k) + \tilde{r}_4(2\alpha_k)]$$
$$\left. + 2A^2[r_2(2\alpha) + \tilde{r}_2(2\alpha) + r_0(2\alpha) + \tilde{r}_0(2\alpha) + 2[r_2(2\alpha) + \tilde{r}_2(\alpha,\tau)]]\cos 2\omega_0\tau\right],$$

$$\alpha_0^{(D)}(\tau) = \frac{1}{4N}\left[A_1A_2\left[2\left[\begin{array}{l}2r_0(\alpha_1+\alpha_0) + 3r_2(\alpha_1+\alpha_2)\\ -[r_0(\alpha_1+\alpha_2) + 2r_2(\alpha_1+\alpha_2)]\cos 2\omega_0\tau + r_4(\alpha_1+\alpha_2)\end{array}\right]\right.$$
$$+ r_2^*(\alpha_1,\alpha_2,\tau) + r_2^*(\alpha_2,\alpha_1,\tau) - \left[\begin{array}{l} r_0^*(\alpha_1,\alpha_2,\tau) + r_0^*(\alpha_2,\alpha_1,\tau)\\ + 2r_2^*(\alpha_1,\alpha_2,\tau) + 2r_2^*(\alpha_2,\alpha_1,\tau)\end{array}\right] \times \cos 2\omega_0\tau$$
$$\left. - r_4^*(\alpha_1,\alpha_2,\tau) - r_4^*(\alpha_2,\alpha_1,\tau)\right] + 2A^2[2\tilde{r}_0(\alpha,\tau) + 3\tilde{r}_2(\alpha,\tau) - [\tilde{r}_6(\alpha,\tau) + 2\tilde{r}_2]$$
$$\times \cos 2\omega_0\tau + \tilde{r}_4(\alpha,\tau) - r_2(2\alpha) - [r_0(2\alpha) + 2r_2(2\alpha)]\cos 2\omega_0\tau - r_4(2\alpha)]].$$

Expression of such quantities in case of coherent estimations are the following

$$\alpha_0^{(I)}(\tau) = \frac{1}{8N}\left[\sum_{k=1}^{2}\left[A_k^2\left[1 + e^{-2\alpha_k|\tau|} + 2[s(2\alpha_k) + \tilde{s}(2\alpha_k,\tau)]\right](2 + \cos 2\omega_0\tau)\right.\right.$$
$$\left.\left. + 2A^2\left[1 + e^{-2\alpha_k|\tau|} + 2[s(2\alpha) + \tilde{s}(\alpha,\tau)]\right]\cos 2\omega_0\tau\right],$$

$$\alpha_0^{(D)}(\tau) = \frac{1}{8N}\left[2\left(A_1A_2 + A^2e^{-\alpha|\tau|}\right)(2 - \cos 2\omega_0\tau) - 2\left(A^2 + A_1A_2e^{-(\alpha_1+\alpha_2)|\tau|}\right)\right.$$
$$\times \cos 2\omega_0\tau + 4\left[A_1A_2s(\alpha_1+\alpha_2) + A^2\tilde{s}(\alpha,\tau)\right](2 - \cos 2\omega_0\tau)$$
$$\left. - 2\left[2A^2s(2\alpha) + A_1A_2[s^*(\alpha_1,\alpha_2,\tau) + s^*(\alpha_2,\alpha_1,\tau)]\right]\cos 2\omega_0\tau\right],$$

where

$$s(\alpha) = \sum_{n=1}^{N-1}\left(1 - \frac{n}{N}\right)e^{-\alpha nT},$$

$$\tilde{s}(\alpha,\tau) = \sum_{n=1}^{N-1}\left(1 - \frac{n}{N}\right)e^{-\alpha(|\tau+nT| + |\tau-nT|)},$$

$$s^*(\alpha_1,\alpha_2,\tau) = \sum_{n=1}^{N-1}\left(1 - \frac{n}{N}\right)e^{-(\alpha_1|\tau+nT| + \alpha_2|\tau-nT|)}.$$

In Fig. 1 there are represented dependences of time averaged estimator variances $\alpha_0^{(I_1)}(\tau)$ and $\alpha_0^{(D)}(\tau)$ on lag τ, calculated for $D_1 = D_2 = 1.0$, $D = 0.8$, $N = 30$, $T = 10$, $\alpha_1 = \alpha_2 = 0.2$ in cases of coherent and component estimation. As it can be seen, these quantities are oscillating attenuating functions and for great lags they are periodic curves. Maximal values of attenuating oscillations and also amplitudes _ stated by coherent estimating are 1.5 times greater than analogous values obtained in case of

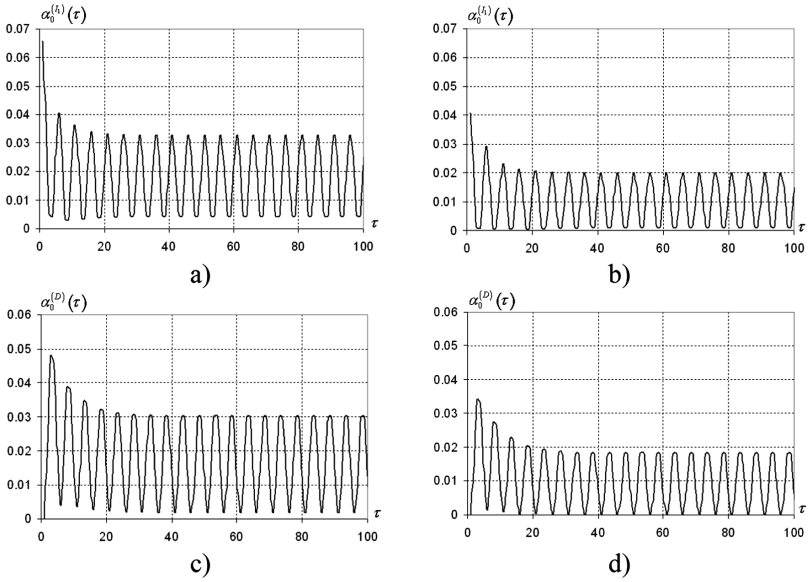


Fig. 1. Dependencies of quantities $\alpha_0^{(I_1)}(\tau)$ and $\alpha_0^{(D)}(\tau)$ on lag τ for coherent (a, b) and for component (c, d) estimators.

component estimation. A difference between variances of coherent and component estimations decreases with decrease of attenuation rate of covariance functions and it increases essentially in case of its increase (Table 1).

Table 1. Time averaged invariant estimator variances for coherent and component estimation

α_1			0.04	0.12	0.4	0.8	1.2
α_2			0.04	0.12	0.4	0.8	1.2
α_3			0.02	0.06	0.2	0.4	0.6
$\alpha_0^{(I_1)}(\tau)$	$\tau = 0$	Comp. estimation	0.17002	0.06206	0.02574	0.01767	0.01393
		Coher. estimation	0.17631	0.07903	0.06108	0.06067	0.06066
$\alpha_0^{(D)}(\tau)$	$\tau = 4$	Comp. estimation	0.06139	0.02137	0.00657	0.00357	0.00259
		Coher. estimation	0.06286	0.02530	0.01443	0.01281	0.01252

Non-attenuating oscillations of time averaged variances of invariants estimator in case of large lags are the results of periodic non-stationarity of the signals. In case of estimation of characteristics of stationary signals the estimators variances tend to definite constant. Such property of the estimators variances must be taken into account in case of statistical processing of experimental data, since it leads to increase of relative mean square estimation error that results in necessity of correlograms truncation (Javorśkyj 2013; Javorśkyj et al. 2007, 2010).

6 The Covariance Invariant Analysis of Real Data

Coherent and component methods given above were used for the analysis of vibrations of decanter Flottweg 24E bearing unit. This decanter is operating on the milk plant in Ukraine. The aim of the analysis was an evaluation of decanter technical state. The scheme of the decanter units is given in Fig. 2.

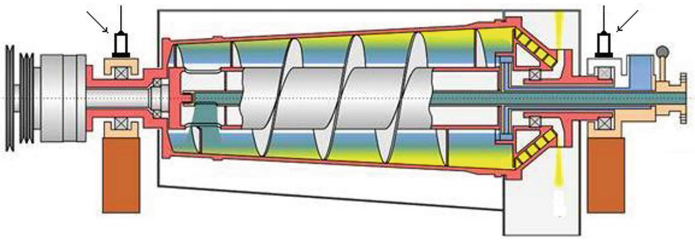


Fig. 2. Mechanical scheme of decanter (arrows show places of accelerometers fastening)

Acquisition and processing of vibro-acceleration signals were provided with the use of multichannel vibro-acoustic system "Vector". Treble frequency of the selective filter – 5 kHz, sampling frequency – $f_{sample} = 10$ kHz. Realization length – 20 s. The fragment of decanter bearing vertical vibration is shown in Fig. 3.

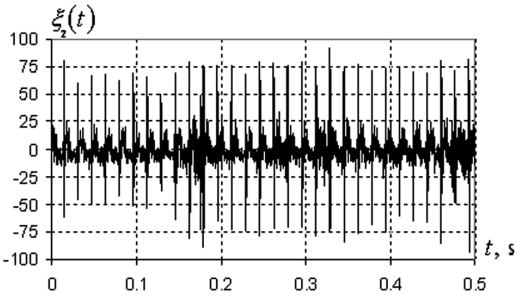


Fig. 3. Realization of decanter bearing vertical vibration

It can be clearly seen from the figure, that signal contains powerful impacts with frequency that corresponds to shaft rotation. Using methods of PCRP statistics the covariance structure of vibration was analyzed. Using such approach, the first stage of analysis is separation of deterministic and stochastic parts. In order to reach this purpose we used coherent method, for which estimator of mean function, that describes deterministic part of vertical or horizontal vibration, has a form (Javorskyj 1984; Javorskyj and Mykhajlyshyn 1996; Javorśkyj et al. 2014a, b, c, 2017a, b, c).

$$\hat{m}_{\xi_1,\xi_2}(t,T) = \frac{1}{2N+1}\sum_{n=-N}^{N}\left\{\begin{array}{c}\xi_1(t+nT)\\ \xi_2(t+nT)\end{array}\right\},$$

where T – test period. Since discrete values of vibration signals are recorded with interval $h = 10^{-4}$ s, so values of t and T may be changed with the same interval. In order to specify value of period we made resampling using Shannon-Kotelnikov formula

$$\xi_{1,2}(t) = \sum_{n=-M}^{M}\xi_{1,2}(nh)\frac{\sin\omega_{\max}(t-nh)}{\omega_{\max}(t-nh)},$$

where $\omega_{\max}$ – the highest frequency of signal. Time t was chosen close to maximum of estimator $\hat{m}_{\xi_2}(t,T)$. Dependence of $\hat{m}_{\xi_2}(t,T)$ on test period T is shown in Fig. 4. This quantity reaches maximum at $T = 1.6667 * 10^{-2}$ s. We consider this value as a true value of period estimator of mean function $\hat{m}_{\xi_2}(t,\hat{P})$ of vertical vibration.

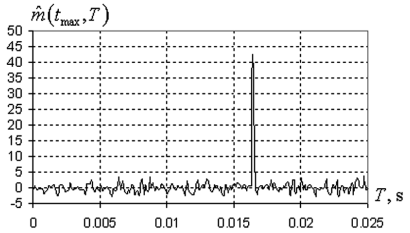


Fig. 4. Dependence of estimator $\hat{m}_{\xi_2}(t_{\max},T)$ on test period T

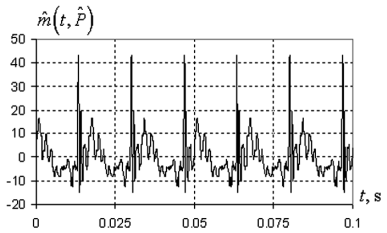


Fig. 5. Dependence of mean function estimator on time t

Frequency 59.998 Hz corresponds to this value of period. Dependence of mean function estimator $\hat{m}_{\xi_2}(t_{\max},\hat{P})$ on time, that describes vibration deterministic part, as it may be seen from the Fig. 5, is characterized by consecutive impacts, which repeat through period of shaft rotation. On the base of mean function values within interval $[0,\ \hat{P}]$ we can calculate amplitude spectrum of vertical vibration's deterministic part. Estimators of harmonics amplitudes $\left|\hat{m}_k^{(\xi_2)}\right|$ were calculated using formulas (Javorśkyj 2013; Javorśkyj et al. 2016)

$$\left\{\begin{array}{c}\hat{m}_k^c\\ \hat{m}_k^s\end{array}\right\} = \frac{2}{L+1}\sum_{n=0}^{L}\hat{m}_{\xi_2}(n\tilde{h},\hat{P})\left\{\begin{array}{c}\cos k\frac{2\pi}{L+1}n\\ \sin k\frac{2\pi}{L+1}n\end{array}\right\},\ \left|\hat{m}_k^{(\xi_2)}\right| = \sqrt{\left(\hat{m}_k^c\right)^2+\left(\hat{m}_k^s\right)^2},$$

here $\tilde{h} = \hat{P}/(L+1)$. Values of estimators of harmonic amplitudes are shown in Fig. 6 in the form of a diagram. As it can be seen from the figure, amplitude spectrum is quite wide, it contains almost 40 harmonics, that correspond to frequency of almost 2.5 kHz.

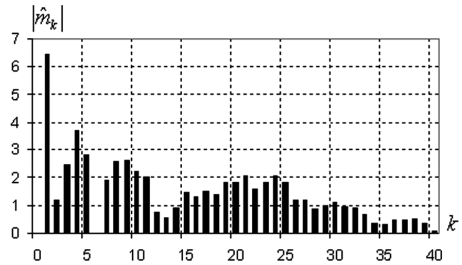


Fig. 6. Amplitude spectrum of deterministic part of vertical vibration

Similar results were obtained while processing of horizontal vibration. We should notice, that period estimator calculated in this case differs from the previous one only with sixth digit after comma.

We obtain the stochastic parts of both vibration components after centering of initial data on estimators of mean functions, namely

$$\overset{\circ}{\xi}_1(t) = \xi_1(t) - \hat{m}_{\xi_1}\left(t, \hat{P}\right), \ \overset{\circ}{\xi}_2(t) = \xi_2(t) - \hat{m}_{\xi_2}\left(t, \hat{P}\right).$$

Before invariant covariance processing of these components, we carry out searching for hidden periodicities in the time variety of covariance properties of each of them. In order to reach this purpose we use symmetric coherent averaging of covariance product (Javorskyj 1984; Javorskyj and Mykhajlyshyn 1996; Javorśkyj et al. 2014a, b, c, 2017a, b, c)

$$\left\{ \begin{matrix} \hat{b}_{\xi_1}(t,\tau,T) \\ \hat{b}_{\xi_2}(t,\tau,T) \end{matrix} \right\} = \frac{1}{2N+1} \sum_{n=-N}^{N} \left\{ \begin{matrix} \overset{\circ}{\xi}_1(t+nT)\, \overset{\circ}{\xi}_1(t+\tau+nT) \\ \overset{\circ}{\xi}_2(t+nT)\, \overset{\circ}{\xi}_2(t+\tau+nT) \end{matrix} \right\}. \tag{66}$$

These statistics at $\frac{\partial^2 \hat{b}_{\xi_p}(t,\tau,T)}{\partial t^2} \neq 0$ have extremes for these T, which are close to period of nonstationarity $\hat{P}$, if analyzed data are realizations of PCRP. The most clearly such extremes appear for these moments of time, which are close to covariance function $b_{\xi_p}(t,\tau)$ extremes. Graphs of dependencies of statistics (66) on test period T at $\tau = 0$ for moments of time t, which are close to estimators extremes, are shown in Fig. 7. As it can be seen, statistics (66) have pronounced peaks, which are significantly over the level of fluctuations.

Points of extremes are close to period of shaft rotation. More detail calculations show that they differ from period estimators for mean functions only with sixth digit. Thus, in this case we also accept that $\hat{P} = 1.6667 * 10^{-2}$ s.

Following the accepted value of period let us calculate the estimators of linear invariants, using the first stage coherent method. We will use formulas (4), (5) and (6), (7) substituting period P by its estimator $\hat{P}$. It is shown in (Javorskyj et al. 2017a, b, c) that mean and covariance functions coherent statistics using a period estimator, are asymptotically unbiased and consistent. Graphs of dependencies of estimators $\hat{I}_1(t,0)$

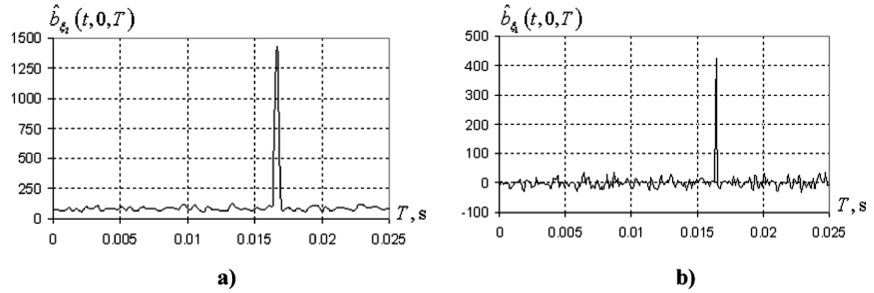


Fig. 7. Dependencies of statistics (66) on test period T at $\tau = 0$

and $\hat{D}(t, \tau_m)$ on time t are shown in Figs. 8a and 9a respectively. Since invariant $\hat{D}(t, \tau)$ at $\tau = 0$ equals _ zero, thus for illustration of its time dependency we chose another time lag, which corresponds to its first maximum. Time dependencies of both quantities contain significant peaks within the period of rotation, which amplitudes significantly over levels of other invariants time values.

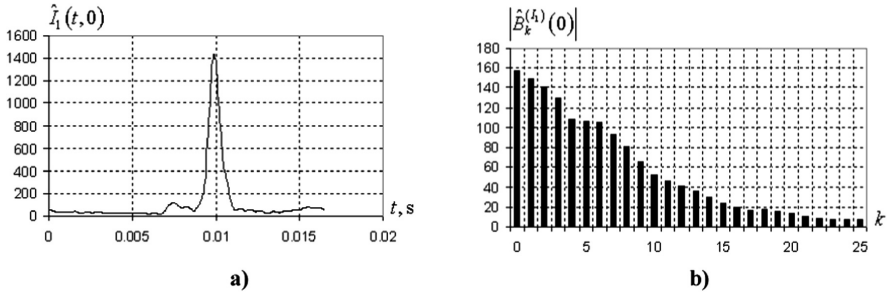


Fig. 8. Dependence of estimator $\hat{I}_1(t, 0)$ on time (a) and its Fourier coefficients (b)

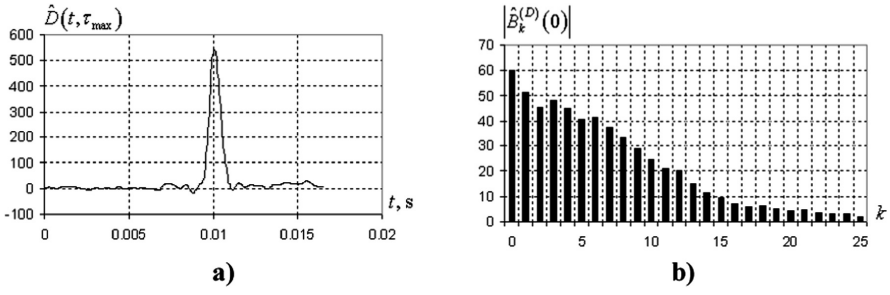


Fig. 9. Dependence of estimator $\hat{D}(t, \tau_m)$ on time (a) and its Fourier coefficients (b)

Estimators of invariants Fourier coefficients, calculated using formulas (56) and (57), in the form of diagrams are shown in Figs. 7b and 8b. Widths of invariants amplitude spectra, as it can be seen, are similar but narrower than amplitude spectra of mean function estimators. They contain almost 20 significant components. Harmonic amplitudes monotonically decrease as their number k increases.

Estimator $\hat{I}_1(t,0)$ is defined by a sum of estimators of variances of vector components, namely it characterizes power of stochastic changes of acceleration vector. The measure of nonstationarity of such vectorial random process is a quantity

$$I = \frac{\sum_{k=1}^{20} \hat{B}_k^{(I_1)}(0)}{\hat{B}_0^{(I_1)}(0)},$$

which can be chosen as an indicator, that defines degree of defect growth. In the present case _ its value equals _ $I = 7.1$, that is quite large and testifies about the presence of powerful stochastic modulation of signal harmonics. Taking into account this result and also significant widths of amplitude spectra of deterministic oscillations as well as periodic variety of power of vector stochastic parts we conclude that defect present in the rotary unit is local and quite developed.

Time variety invariant $\hat{D}(t,\tau_m)$ estimator is similar in shape to time variety of estimator $\hat{I}_1(t,0)$. Amplitude spectrum is similar too but amplitude of $\hat{D}(t,\tau_m)$ and amplitudes of its harmonics are significantly smaller. Such ratio between values of these invariants was expected following their definitions. Invariant $\hat{D}(t,\tau)$ is an indicator of rotation, so even the fact that its value does not equals to zero and its time variety has the clearly defined shape and is characterized by a wide amplitude spectrum testifies that present defect is developed in moving part of a rotary unit and has a percussive character. The same conclusion is testified by analysis of linear invariants with respect to time lag τ.

In Fig. 10 the dependencies of estimators $\hat{I}_1(t_{\max},\tau)$ and $\hat{D}(t_{\max},\tau)$ on time lag τ, calculated using formulas (4)–(7) for these moments of time t_m, which correspond to maxima of graphs in Figs. 7a and 8a, are shown. These dependencies are similar. They have a group structure.

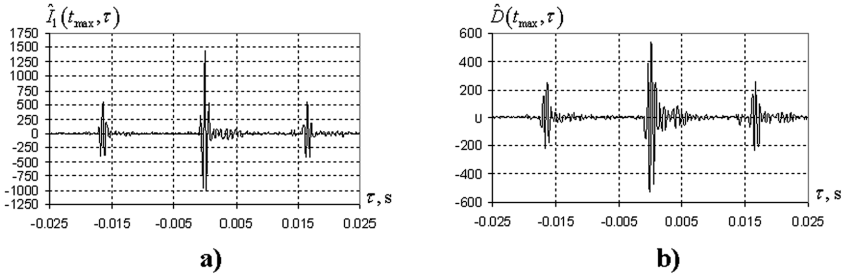


Fig. 10. Dependencies of linear invariants estimators on time lag

Time intervals between individual groups, the power of which decreases as time lag increases, are close to period of shaft rotation. Periods of oscillations of individual groups differ insignificantly and are close to the value, which is almost 25 times smaller than the period of shaft rotation. We can suppose that such oscillations are forced oscillations of a rotary unit, which appear under the influence of impacts caused by present defect.

Estimators of invariants covariance components have similar structure. Nonzero covariance components are complex ones: $B_k^{(I_1,D)}(\tau) = \frac{1}{2}\left[C_k^{I_1,D}(\tau) - iS_k^{I_1,D}(\tau)\right]$. Estimators of these quantities were represented in the same form: $\hat{B}_k^{(I_1,D)}(\tau) = \frac{1}{2}\left[\hat{C}_k^{I_1,D}(\tau) - i\hat{S}_k^{I_1,D}(\tau)\right]$ and sine and cosine components were calculated on the base of formulas:

$$\begin{Bmatrix} \hat{C}_k^{(I_1)}(\tau) \\ \hat{S}_k^{(I_1)}(\tau) \end{Bmatrix} = \frac{2}{M+1}\sum_{n=0}^{M} \hat{I}_1(nh,jh) \begin{Bmatrix} \cos k\frac{2\pi}{M+1}n \\ \sin k\frac{2\pi}{M+1}n \end{Bmatrix},$$

$$\begin{Bmatrix} \hat{C}_k^{(D)}(\tau) \\ \hat{S}_k^{(D)}(\tau) \end{Bmatrix} = \frac{2}{M+1}\sum_{n=0}^{M} \hat{D}(nh,jh) \begin{Bmatrix} \cos k\frac{2\pi}{M+1}n \\ \sin k\frac{2\pi}{M+1}n \end{Bmatrix},$$

Dependencies of estimators for the zeroth and the first covariance components of invariants on time lag are shown in Figs. 11 and 12. In order to specify internal

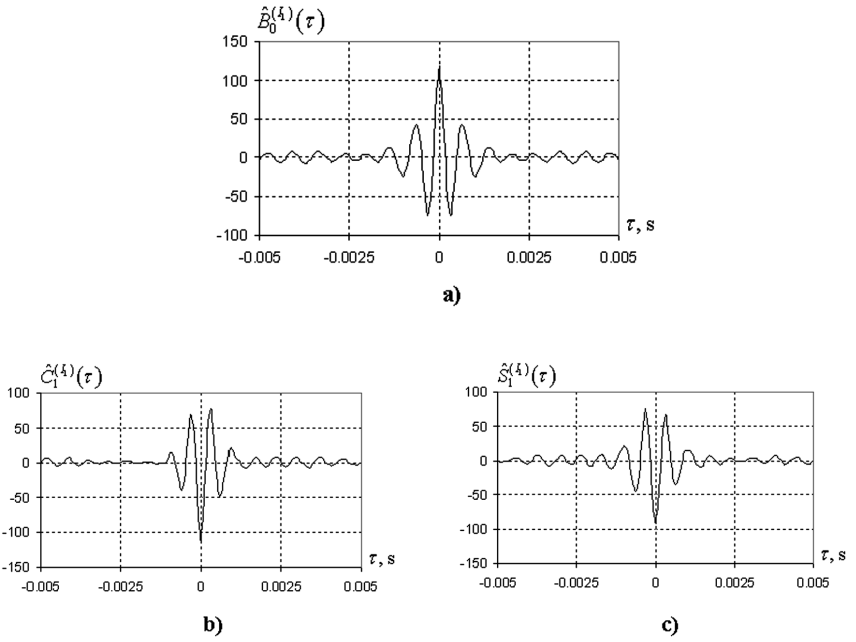


Fig. 11. Estimators of covariance components of invariant $\hat{I}_1(t_m, \tau)$

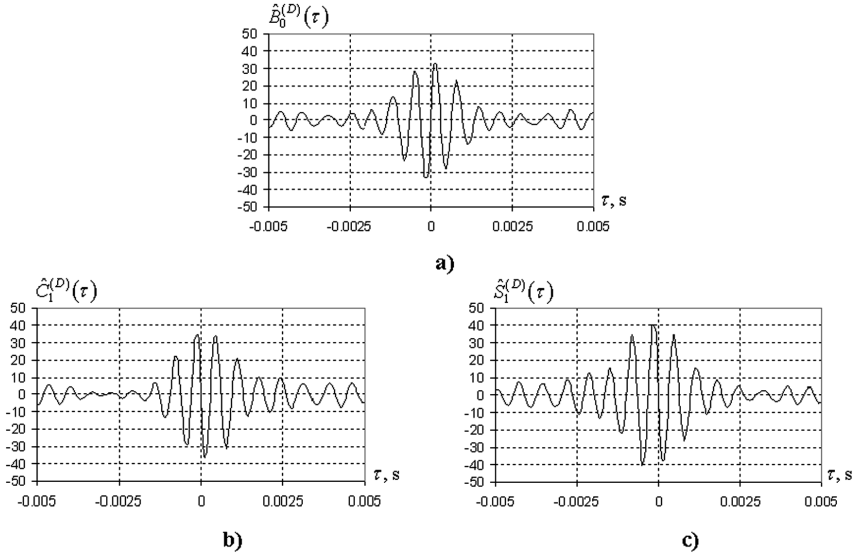


Fig. 12. Estimators of covariance components of invariant $D(t_m, \tau)$

structure of individual groups, which are similar, we illustrated only a group, which corresponds to a set of small time lags.

Comparing graphs of estimators of both invariants components, we should notice two differences between them. The first one is that oscillations, which describe the behavior of estimators of invariant $\hat{I}_1(t_m, \tau)$ covariance components, damp faster as time lag increases, and the second – estimators of covariance components $\hat{B}_k^{(I_1)}(\tau)$ mainly are characterized by pair dependence on time lag, but estimators of covariance components $\hat{B}_k^{(D)}(\tau)$ – by odd one. It can be easily proved by separating their pair and odd parts.

As an example, in the Fig. 13 the graphs of pair $\left[\hat{B}_0^{(I_1, D)}(\tau)\right]^+ = \frac{1}{2}\begin{bmatrix} \hat{B}_0^{(I_1, D)}(\tau) + \\ + \hat{B}_0^{(I_1, D)}(-\tau) \end{bmatrix}$ and odd $\left[\hat{B}_0^{(I_1, D)}(\tau)\right]^- = \frac{1}{2}\begin{bmatrix} \hat{B}_0^{(I_1, D)}(\tau) - \\ -\hat{B}_0^{(I_1, D)}(-\tau) \end{bmatrix}$ parts of zeroth covariance components are represented.

As mentioned above, linear invariants have almost 25 significant Fourier coefficients. Thus, for enhancement of statistical processing while the estimation of quadratic invariants the component method was used and the number of components N_2 was chosen 25 too. Computation shows, that periodic non-stationarity of the second order detected in linear invariants properties is stronger represented in the time variety of quadratic invariant $I_2(t, \tau)$ estimator, which is a determinant of matrix (3). Such dependence is similar in shape to the previous ones (Figs. 7a and 8a), but its amplitude is significantly larger. It also should be noticed, that amplitudes of the first harmonics decrease slower.

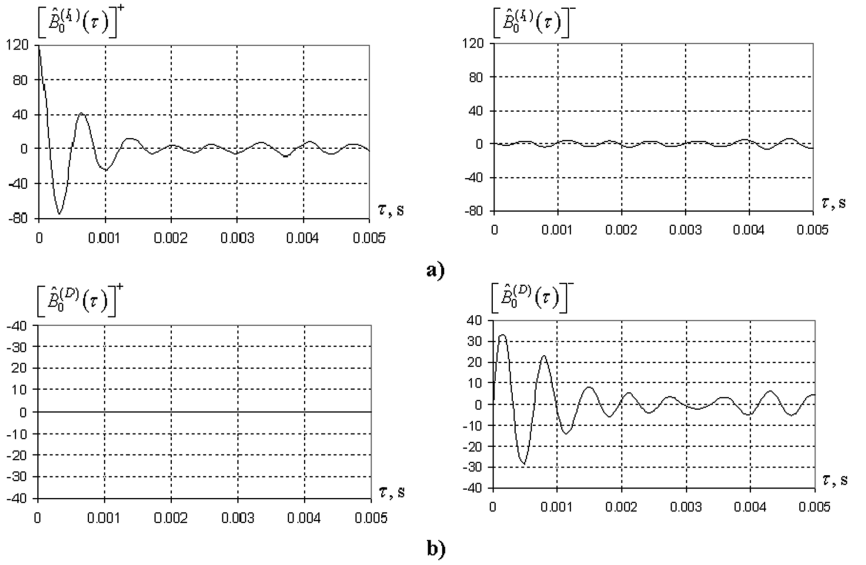


Fig. 13. Pair and odd parts of zeroth components of linear invariants $\hat{B}_0^{(I_1)}(\tau)$ (a) and $\hat{B}_0^{(D)}(\tau)$ (b)

Top-shaped form is also proper to time varieties of estimators $\lambda_{1,2}(t,\tau)$, which are the eigen values of matrix (3) symmetric part and define covariance functions of vector components in the own basis, and also

$$\lambda_1(t,\tau)+\lambda_2(t,\tau)=I_1(t,\tau),$$

$$\lambda_1(t,\tau)\lambda_2(t,\tau)=I_2(t,\tau).$$

It is clearly seen from the Figs. 14a and 15a, in which graphs of time varieties of $\lambda_1(t,0)$ and $\lambda_2(t,0)$ are represented, that estimators values within the period are positive, and value of $\lambda_1(t,0)$ significantly exceeds value of $\lambda_2(t,0)$. It means _ that the second order curves built on their basis are ellipses. Significant difference between time

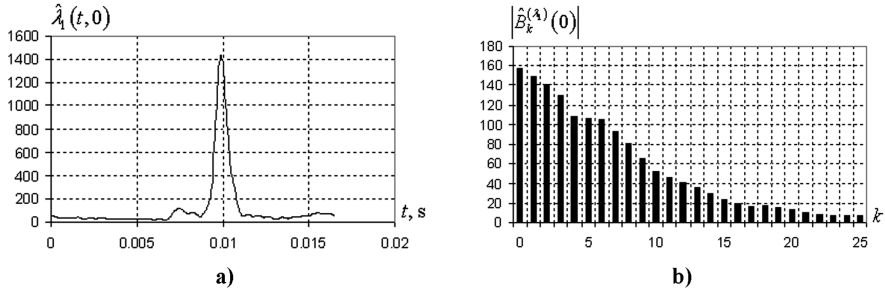


Fig. 14. Dependence of estimator $\hat{\lambda}_1(t,0)$ on time (a) and its Fourier coefficients $\left|\hat{B}_k^{(\lambda_1)}(0)\right|$ (b)

dependencies of $\lambda_1(t,0)$ and $\lambda_2(t,0)$ is a displacement of their maxima with respect to time. Such displacement, in practice, does not represent the relationship between amplitudes of harmonic of invariants estimators (Fig.s 14b and 15b) but only changes their phase relationships.

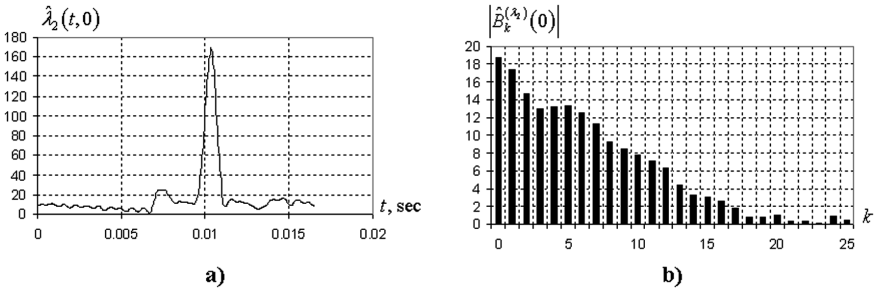


Fig. 15. Dependence of estimator $\hat{\lambda}_2(t,0)$ on time (a) and its Fourier coefficients $\left|\hat{B}_k^{(\lambda_2)}(0)\right|$ (b)

Quantities, which define these relationships were computed using formulas:

$$\hat{B}_0^{(\lambda_1,\lambda_2)}(\tau) = \frac{1}{M+1}\sum_{n=0}^{M}\hat{\lambda}_{1,2}(nh,\tau),$$

$$\hat{C}_k^{(\lambda_1,\lambda_2)}(\tau) = \frac{2}{M+1}\sum_{n=0}^{M}\hat{\lambda}_{1,2}(nh,\tau)\cos k\frac{2\pi}{M+1}n,$$

$$\hat{S}_k^{(\lambda_1,\lambda_2)}(\tau) = \frac{2}{M+1}\sum_{n=0}^{M}\hat{\lambda}_{1,2}(nh,\tau)\sin k\frac{2\pi}{M+1}n,$$

$$\left|\hat{B}_k^{(\lambda_1,\lambda_2)}(\tau)\right| = \sqrt{\left[\hat{C}_k^{(\lambda_1,\lambda_2)}(\tau)\right]^2 + \left[\hat{S}_k^{(\lambda_1,\lambda_2)}(\tau)\right]^2}.$$

As a result of invariants estimators time shifting the orientation of ellipses changes in time (Fig. 16) and _ some angle sectors, where power of vector stochastic changes is the greatest, can be allocated. Directions of the greatest power changes happen can be allocated on the base of graphical representation of quantity

$$b_\beta(t,\tau) = \lambda_1(t,\tau)\cos^2\beta + \lambda_2(t,\tau)\sin^2\beta,$$

which characterizes correlations with respect to direction. Here β is an angle between given direction and own basis. If $\tau = 0$ then this quantity defines power of stochastic oscillations. As it is seen from the Fig. 17, this direction is close to vertical one.

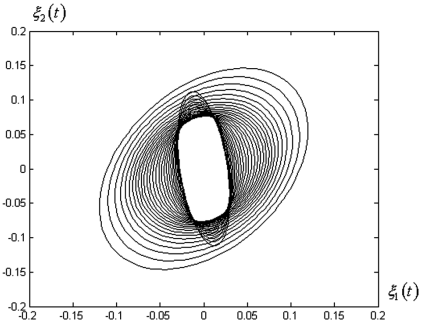


Fig. 16. Second order curves

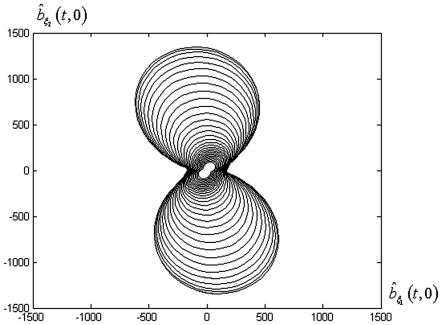


Fig. 17. Time variety of stochastic vibration power with respect to direction

7 Conclusions

Covariance invariants of vectorial PCRP are functions dependent on two variables – time t and time lag τ. Time variety of these quantities is periodic, thus for their analysis on the base of real data the coherent method can be used which is based on synchronous averaging of data through period of non-stationarity. On the base of values of invariants, calculated within period, their Fourier coefficients – covariance components can be calculated. Such calculations are made for each value of time lag from the previous chosen set. Covariance components values decrease as their number k increases. Component method, for which estimators have form of trigonometric functions, allows to neglect insignificant harmonic components while statistic processing of data. Sufficient condition of asymptotic unbiasedness of coherent and component estimators of invariants is damping of correlations as time lag increases. For Gaussian vectorial PCRP such condition also provides consistency of estimators. Formulas for variance and bias of estimators obtained in the paper allow to compute mean-square errors of statistic processing in dependence on realization length, sampling interval and parameters, which describe covariance structure of analyzed signals, and on this base to choose grounded parameters for processing.

The carried out theoretical investigations became a basis for algorithms grounding and creation of software for vectorial covariance processing and analysis of experimental data, which were verified using vibration stand (Javorśkyj et al. 2017a, b, c) and while diagnose machineries on the enterprise in Ukraine. One of the examples of vectorial covariance analysis usage for evaluation of industrial object state is given in the paper. Investigation results showed that using linear and quadratic invariants of covariance tensor allows defect detection on more early stages than it is possible while conducting covariance analysis of each of them separately. Calculation of eigen values of covariance tensor and basis of its orientation allows to define caused by defect appearance spatial properties of vibration vector and to localize it.

References

Dragan Y, Rozhkov V, Javorśkyj I (1987) The methods of probabilistic analysis of oceanological rhythms. Gidrometeoizdat, Leningrad. (in Russian)

Javorśkyj I (2013) Mathematical models and analysis of stochastic oscillations. Physico-Mechanical Institute of NAS of Ukraine, Lviv. (in Ukrainian)

Javorśkyj I, Matsko I, Yuzefovych R, Dzeryn O (2017a) Vectorial periodically correlated random processes and their covariance invariant analysis. In: Chaari F, Leskow J, Napolitano A, Zimroz R, Wylomanska A (eds) Cyclostationarity: theory and methods III. Springer, New York, pp 121–150

McCormick AC, Nandi AK (1998) Cyclostationarity in rotating machine vibrations. Mech Syst Signal Process 12(2):225–242

Capdessus C, Sidahmed M, Lacoume JL (2000) Cyclostationary processes: application in gear faults early diagnosis. Mech Syst Signal Process 14(3):371–385

Antoni J, Bonnardot F, Raad A, El Badaoui M (2004) Cyclostationary modeling of rotating machine vibration signals. Mech Syst Signal Process 18:1285–1314

Antoni J (2009) Cyclostationarity by examples. Mech Syst Signal Process 23:987–1036

Javorśkyj I, Kravets I, Matsko I, Yuzefovych R (2017b) Periodically correlated random processes: application in early diagnostics of mechanical systems. Mech Syst Signal Process 83:406–438

Gardner WA (1987) Statistical spectral analysis: a nonprobabilistic theory. Prentice Hall, Englewood Cliffs

Sadler BM, Dandawate AV (1998) Nonparametric estimation of the cyclic cross spectra. IEEE Trans Inf Theory 44(1):351–358

Javorśkyj I, Kravets I, Yuzefovych R et al (2014a) Vectorial diagnosis of rolling bearing with growing defect on the outer race. Vibratciji v technici I technologijah 2(76):101–110. (in Ukrainian)

Javorśkyj I, Isayev I, Zakrzewski Z, Brooks S (2007) Coherent covariance analysis of periodically correlated random processes. Signal Process 87(1):13–32

Javorśkyj I, Yuzefovych R, Kravets I, Matsko I (2014b) Methods of periodically correlated random processes and their generalizations. In: Chaari F, Leśkow J, Sanches-Ramirez A (eds) Cyclostationarity: theory and methods. Springer, New York, pp 73–93

Javorśkyj I, Isayev I, Yuzefovych R, Majewski J (2010) Component covariance analysis for periodically correlated random processes. Signal Process 90:1083–1102

Javorśkyj I (1984) Application of Buys-Bullot scheme in statistical analysis of rhythmic signals. Radioelectron Telecommun Syst 27(11):28–33

Javorśkyj I, Mykhajlyshyn V (1996) Probabilistic models and investigation of hidden periodicities. Appl Math Lett 9:21–23

Javorśkyj I, Dehay D, Kravets I (2014c) Component statistical analysis of second order hidden periodicities. Digit Signal Process 26:50–70

Javorśkyj I, Yuzefovych R, Matsko I, Zakrzewski Z, Majewski J (2017c) Coherent covariance analysis of periodically correlated random processes for unknown non-stationarity period. Digit Signal Process 65:27–51

Javorśkyj I, Yuzefovych R, Matsko I, Zakrzewski Z, Majewski J (2016) Discrete estimators for periodically correlated time series. Digit Signal Process 53:25–40

On Extreme Values in Stationary Weakly Dependent Random Fields

Paul Doukhan[1,2(✉)] and José G. Gómez[1,3(✉)]

[1] UMR AGM 8088 CNRS, Cergy-Pontoise University, 95000 Cergy-Pontoise, France
paul.doukhan@u-cergy.fr, jose-gregorio.gomez-garcia@unicaen.fr
[2] CIMFAV, Universidad de Valparaiso, Valparaíso, Chile
[3] UMR LMNO 6139 CNRS, University of Caen-Normandy, 14000 Caen, France

Abstract. The existing literature on extremal types theorems for stationary random processes and fields is, until now, developed under either mixing or "Coordinatewise (Cw)-mixing" conditions. However, these mixing conditions are very restrictives and difficult to verify in general for many models. Due to these limitations, we extend the existing theory, concerning the asymptotic behaviour of the maximum of stationary random fields, to a weaker and simplest to verify dependence condition, called *weak dependence*, introduced by Doukhan and Louhichi [Stochastic Processes and their Applications 84 (1999): 313–342]. This stationary weakly dependent random fields family includes models such as Bernoulli shifts, chaotic Volterra and associated random fields, under reasonable addition conditions. We mention and check the weak dependence properties of some specific examples from this list, such as: linear, Markovian and LARCH(∞) fields. We show that, under suitable weak-dependence conditions, the maximum may be regarded as the maximum of an approximately independent sequence of sub-maxima, although there may be high local dependence leading to clustering of high values. These results on asymptotic max-independence allow us to prove an extremal types theorem and discuss domain of attraction criteria in this framework. Finally, a numerical experiment using a non-mixing weakly dependent random field is performed.

Keywords: Weak dependence · Extreme values · Random field · Extremal types theorem · Domain of attraction

1 Introduction

This is a well-known fact that, given independent and identically distributed random variables $X_1, \ldots, X_n$, the asymptotic distribution of the normalised maxima $\mathbb{P}(a_n^{-1}(M_n - b_n) \leqslant x)$, for $M_n = \max\{X_1, \ldots, X_n\}$ and for some sequences some $a_n > 0$ and $b_n \in \mathbb{R}$, converges to a non-degenerate distribution G, which is of extremal type. The same result is obtained by Leadbetter *et al.* [15] for dependent stationary sequences $(X_i)_{i\in\mathbb{N}}$ under a weak mixing condition. Even more,

F. Chaari et al. (Eds.): CSTA 2017, ACM 16, pp. 92–110, 2020.
https://doi.org/10.1007/978-3-030-22529-2_5

in [15] extend this result to stationary processes $\{X_t,\ t \geqslant 0\}$ of continuous time, redefining the maximum as $M_T = \sup\{X_t : 0 \leqslant t \leqslant T\}$, again under the same mixing condition. In the case of stationary random fields, we must begin by quoting [1, § 6.9], where it has been shown that such result holds for $M_T = \sup\{X_{\mathbf{t}} : \mathbf{t} \in [0,T]^d\}$ $(d > 1)$, with $(X_{\mathbf{t}})_{\mathbf{t}}$ Gaussian with weak covariance conditions. Specifically,

$$\mathbb{P}\left(\frac{M_T - b_T}{a_T} \leqslant x\right) \underset{T\to\infty}{\longrightarrow} G(x), \tag{1}$$

where $G(x) = \exp(-\exp(-x))$ and whose normalisations are defined by

$$a_T = (2d\ \log T)^{-1/2}$$

$$b_T = (2d\ \log T)^{1/2} + \frac{\frac{1}{2}(d-1)(\log\log T) + \log\left((2\pi)^{-1}\ \sqrt{\det(\Lambda)(d/\pi)^{d-1}}\right)}{(2d\ \log T)^{1/2}},$$

with Λ denoting the usual matrix of second-order spectral moments. Finally, using all those ideas, Leadbetter and Rootzén (1998) [16] proved the result for stationary random fields subject to appropriate long-range dependence restrictions, showing that, if the limit (1) holds for some normalisation constants $a_T > 0$ and b_T, then G is of extreme value type. Precisely, there the result is shown considering a weaker version of dependence than the usual mixing condition of Rosenblatt (1956) [17], called "*Coordinatewise (Cw) mixing*", which is the usual strong mixing condition (multiplied by the number of sub-blocks of the partition of the domain), restricted to events of type $E = \{\max\{X_{\mathbf{t}} : \mathbf{t} \in A\} \leqslant u\}$. However, this dependence condition is very difficult to check. For instance, the only example presented in [12], for which it is shown how to compute the extremal index of stationary random fields. Those authors verify the Cw-mixing condition in order to validate the theoretical results, nevertheless the special random field used which fits this mixing property is extremely specific. A similar situation, although the results are more generals, is presented in [14]. Note also that, [16] does not provide example for which this condition holds. In fact, bounding the decay rates of the Cw-mixing coefficients is suggested a further independent development.

A natural idea to reduce those limitations could be the use of classical mixing theory results for random fields, *e.g.* [4]. Nevertheless, this goes on being a heavy restriction for several reasons. The first reason is that the mixing assumptions are difficult to check, especially for random fields; *e.g.*, Doukhan (1994) [4] provides several explicit bounds of the decay for mixing sequences but analogue result only hold for very few classes of examples of random fields.

The second reason is that mixing is quite restrictive for both random processes and fields. For instance, the autoregressive process of order 1:

$$X_i = \frac{1}{2}(X_{i-1} + \xi_i), \tag{2}$$

with innovations $(\xi_i)_{i\in\mathbb{Z}}$ iid such that $\mathbb{P}(\xi_i = 0) = \mathbb{P}(\xi_i = 1) = 1/2$, and the random field defined in (15), are not mixing (see [2] and [6], respectively). However,

these random fields are weakly dependent in the sense of Doukhan and Louhichi (1999) [5], as shown in [5,6], respectively. This last dependence condition is very useful because it writes as covariance functions. In fact, note that an important property of associated random variables is that zero correlation implies independence (see [13]). Therefore, one may hope that dependence will appear in this case only through the covariance structure, which is much easier to compute that a mixing coefficient. Additionally, under weak assumptions, weak-dependence in the sense of Doukhan and Louhichi (1999) [5], includes models like Bernoulli shifts, Markov, associated, mixing, etc.

In view of the above remarks, we will focus on obtaining an "extremal types theorem" for the maxima $M_{(n_1,\ldots,n_d)} = \max\left\{X_{\mathbf{t}} \ : \ \mathbf{t} \in \prod_{i=1}^{d}\{1,2,\ldots,n_i\}\right\}$ of stationary random fields $(X_{\mathbf{t}})_{\mathbf{t}\in\mathbb{Z}^d}$, under conditions of dependence much easier to verify than Cw-mixing condition and less restrictive than mixing conditions *i.e.*; we use *weak-dependence* conditions, such as defined in [5].

This paper is organised as follows. In Sect. 2, we recall a general result on extremal types theorem provided in [16]. In Sect. 3, we precisely settle the useful weak-dependence conditions for stationary random fields, and we provide some examples. For example, if the parameter space is a regular grid, such as $\mathbb{Z}^d$, we precisely check the weak dependence properties of linear fields, well adapted to model, for example, radiography and image data (see [9]). Besides, we exhibit the example of a stationary and weakly dependent random field which is not strongly mixing. In Sect. 4, we introduce assumptions on the weak dependence coefficients of stationary random fields which enable to prove that the maximum for these fields may be asymptotically rewritten as the maximum of approximately independent sequences of sub-maxima, although there may be high local dependence leading to clustering of high values. This is the preliminary result (Lemma 1), which will be the heart of all other results. In Sect. 5, we provide the extremal types theorem and we discuss domain of attraction criteria. A simulation study is included in Sect. 6 using a stationary weakly dependent random field, which is not strongly mixing.

2 A General Result on Extremal Types

The following general extremal types theorem provides a general property to ensure that the maximum of dependent processes and fields admit asymptotic distribution of the extreme value type.

Proposition 1. *([16, Proposition 2.1]) Let M_T, with $T > 0$, be random variables such that*

$$\mathbb{P}\left(a_T^{-1}(M_T - b_T) \leqslant x\right) \underset{T\to\infty}{\longrightarrow} G(x), \tag{3}$$

where G is a non-degenerate distribution and $a_T > 0$, $b_T \in \mathbb{R}$ are normalisation constants. Suppose now that for each real x, $u_T = a_T x + b_T$ and

$$\mathbb{P}\left(M_{kT} \leqslant u_{kT}\right) - \mathbb{P}^k(M_{\phi_k(T)} \leqslant u_{kT}) \underset{T\to\infty}{\longrightarrow} 0 \tag{4}$$

holds for each $k = 1, 2, \ldots$ *and some continuous strictly increasing functions* $\phi_k(T) \underset{T\to\infty}{\longrightarrow} \infty$. *Then* G *is of extreme value type.*

Remark 1. Let $(X_{\mathbf{i}})_{\mathbf{i}\in\mathbb{Z}^d}$ be a stationary random field. Note that (4) is obviously true if $T = n$, $\phi_k(T) = T$ and M_{kn} is the maximum of $k-$independent random blocks $Y_j = (X_i,\ i \in B_j,\ |B_j| = n)$, $j = 1, \ldots, k$. Moreover, observe that the approximation (4) suggests decay of global dependence. We thus can suppose high local dependence and consider small long-range dependence conditions in order to verify such condition (4) and develop the extremal types theorem for stationary random fields. In particular, with these assumptions, we can develop the results under weak-dependence conditions.

3 Weak Dependence of a Random Field

3.1 Definitions

Let $\Lambda^u(E)$ be the set of $\mathbb{R}$-valued functions defined on E^u with $u \in \mathbb{N}$ and $E \subseteq \mathbb{R}$, that are bounded by 1 and have a finite Lipschitz modulus $\mathrm{Lip}(\cdot)$, *i.e.*,

$$\Lambda^u(E) = \{f : E^u \longrightarrow \mathbb{R} \quad | \quad \mathrm{Lip}(f) < \infty \quad \textit{and} \quad ||f||_\infty \leqslant 1\},$$

where

$$\mathrm{Lip}(f) = \sup_{(x_1,\ldots,x_u)\neq(y_1,\ldots,y_u)} \frac{|f(x_1,\ldots,x_u) - f(y_1,\ldots,y_u)|}{\delta_u\left((x_1,\ldots,x_u),(y_1,\ldots,y_u)\right)},$$

with

$$\delta_r((x_1,\ldots,x_r),(y_1,\ldots,y_r)) = \sum_{i=1}^{r} |x_i - y_i|. \tag{5}$$

We will consider E-valued random fields over $\mathbb{Z}^d$, for some $d \in \mathbb{N}$ fixed. In this case, if we set the norm $||(k_1,\ldots,k_d)|| = \delta_d((k_1,\ldots,k_d),\mathbf{0})$ in $\mathbb{Z}^d$, we say that two finite sequences $\mathbf{I} = (\mathbf{i}_1,\ldots,\mathbf{i}_u)$ and $\mathbf{J} = (\mathbf{j}_1,\ldots,\mathbf{j}_v)$ in $\mathbb{Z}^d$ are $l-$distant if

$$\min\{||\mathbf{i}_s - \mathbf{j}_t||\ :\ s = 1,\ldots,u\ ;\ t = 1,\ldots,v\} = l.$$

Definition 1. Let $\psi : \mathbb{N}^2\times(\mathbb{R}^+)^2 \longrightarrow \mathbb{R}^+$ be a function and let $(\epsilon(l))_{l\geqslant 0}$ be a real positive sequence tending to zero. The random field $X = \{X_{\mathbf{t}} : \mathbf{t} \in \mathbb{Z}^d\}$ is (ϵ,ψ)-**weakly dependent** if for any pair of $l-$distant finite sequences $\mathbf{I} = (\mathbf{i}_1,\ldots,\mathbf{i}_u)$, $\mathbf{J} = (\mathbf{j}_1,\ldots,\mathbf{j}_v)$; and any pair of functions $(f,g) \in \Lambda^u(E)\times\Lambda^v(E)$:

$$|\mathrm{Cov}(f(X_{\mathbf{i}_1},\ldots,X_{\mathbf{i}_u}), g(X_{\mathbf{j}_1},\ldots,X_{\mathbf{j}_v}))| \leqslant \psi\left(u,v,\mathrm{Lip}(f),\mathrm{Lip}(g)\right)\epsilon(l). \tag{6}$$

In particular,
if $\psi(u,v,x,y) = vy$, this is called θ-weakly dependent and $\epsilon(l)$ will be denoted by $\theta(l)$,

if $\psi(u,v,x,y) = ux + vy$, this is called η-weakly dependent and $\epsilon(l)$ will be denoted by $\eta(l)$,
if $\psi(u,v,x,y) = uvxy$, this is called κ-weakly dependent and $\epsilon(l)$ will be denoted by $\kappa(l)$,
if $\psi(u,v,x,y) = ux + vy + uvxy$, this is called λ-weakly dependent and $\epsilon(l)$ will be denoted by $\lambda(l)$.

3.2 Examples

In this section we give a non-exhaustive list of examples of weakly dependent random fields. In the sequel, $X = (X_{\mathbf{i}})_{\mathbf{i}\in\mathbb{Z}^d}$ denote a weakly dependent stationary random field (the conditions of the stationarity will not be specified) and $(\xi_{\mathbf{i}})_{\mathbf{i}\in\mathbb{Z}^d}$ denote a centred unit variance independent identically distributed random field.

Example 1 (Bernoulli shifts). Consider a function $H : \mathbb{R}^{\mathbb{Z}^d} \longrightarrow \mathbb{R}$ and define $(X_{\mathbf{i}})_{\mathbf{i}\in\mathbb{Z}^d}$ as

$$X_{\mathbf{i}} = H\left(\xi_{\mathbf{i}-\mathbf{j}},\ \mathbf{j} \in \mathbb{Z}^d\right).$$

For each $l > 0$, define now $X_{\mathbf{i},l} = H\left(\xi_{\mathbf{j}-\mathbf{i}},\ \|\mathbf{j}\| \leqslant l\right)$. Observe that $\Delta_p(l) := \|X_{\mathbf{i}} - X_{\mathbf{i},l}\|_p$ does not depend on $\mathbf{i}$. In this case, $(X_{\mathbf{i}})_{\mathbf{i}\in\mathbb{Z}^d}$ is $\eta-$weakly dependent with $\eta(l) = 2\Delta_1(l/2 - 1)$.

Remark 2.(1) For $d = 1$, if the random field is causal: $X_i = H(\xi_{i-j},\ j \geqslant 0)$, then $\eta(l) = \theta(l) = \Delta_1(l-1)$ and the weak dependence function takes the simple form $\psi(u,v,x,y) = vy$. In particular, the AR(1) process (2) mentioned in the introduction is θ-weakly dependent with $\theta(l) = \mathcal{O}(2^{-l})$. For more details on the dependence properties of causal time series, see [3, § 3.1.4, § 3.1.5], and in the context of extreme value theory, see [11, § 3.1].

(2) For $d > 1$, note that models are inherently non-causal in the sense that, unlike time series, they are not defined with respect to some order relation on $\mathbb{Z}^d$. For particular cases, definitions of causality and semi-causality for random fields and spatio-temporal models are provided in [8] and [9] respectively.

In order to briefly explain the technique for demonstrating weak-dependence in stationary random fields, we will use the following application. For the general case, the exact same steps can be applied, see [3,5,6].

Application 1 (Linear fields). Define $X = (X_{\mathbf{i}})_{\mathbf{i}\in\mathbb{Z}^d}$ as

$$X_{\mathbf{i}} = \sum_{\mathbf{j}\in\mathbb{Z}^d} b_{\mathbf{j}}\xi_{\mathbf{i}-\mathbf{j}}, \tag{7}$$

where $\sum_{\mathbf{j}} b_{\mathbf{j}}^2 < \infty$. Then X is η-weakly dependent with

$$\eta^2(2l) = 4\sum_{\|\mathbf{j}\|>l} b_{\mathbf{j}}^2. \tag{8}$$

Indeed, let $l > 0$ and $X_{\mathbf{i},l} = \sum_{\|\mathbf{j}\| \leqslant l} b_{\mathbf{j}} \xi_{\mathbf{i}-\mathbf{j}}$. Then, we have that

$$\Delta_2^2(l) := \mathbb{E}\left(X_{\mathbf{i},l} - X_{\mathbf{i}}\right)^2 = \mathbb{E}\left(\sum_{\|\mathbf{j}\|>l} b_{\mathbf{j}} \xi_{\mathbf{i}-\mathbf{j}}\right)^2 = \sum_{\|\mathbf{j}\|>l} b_{\mathbf{j}}^2 \, \mathbb{E}\xi_{\mathbf{i}-\mathbf{j}}^2 = \sum_{\|\mathbf{j}\|>l} b_{\mathbf{j}}^2, \quad (9)$$

which does not depend on $\mathbf{i}$ as we expected.

On the other hand, let $\mathbf{X_I} = (X_{\mathbf{i}_1}, \ldots, X_{\mathbf{i}_u})$ and $\mathbf{X_J} = (X_{\mathbf{j}_1}, \ldots, X_{\mathbf{j}_v})$, and consider their truncated versions $\mathbf{X}_{\mathbf{I},s} = (X_{\mathbf{i}_1,s}, \ldots, X_{\mathbf{i}_u,s})$ and $\mathbf{X}_{\mathbf{J},s} = (X_{\mathbf{j}_1,s}, \ldots, X_{\mathbf{j}_v,s})$. Then, if $\mathbf{I} = (\mathbf{i}_1, \ldots, \mathbf{i}_u)$ and $\mathbf{J} = (\mathbf{j}_1, \ldots, \mathbf{j}_v)$ are $l-$distant with $l > 2s$,

$$\mathrm{Cov}\left(f(\mathbf{X}_{\mathbf{I},s}), g(\mathbf{X}_{\mathbf{J},s})\right) = 0, \quad \forall (f,g) \in \Lambda^u(\mathbb{R}) \times \Lambda^v(\mathbb{R}).$$

Thus,

$$\begin{aligned}
|\mathrm{Cov}\left(f(\mathbf{X_I}), g(\mathbf{X_J})\right)| &\leqslant |\mathrm{Cov}\left(f(\mathbf{X_I}) - f(\mathbf{X}_{\mathbf{I},s}), g(\mathbf{X_J})\right)| \\
&\quad + |\mathrm{Cov}\left(f(\mathbf{X}_{\mathbf{I},s}), g(\mathbf{X_J}) - g(\mathbf{X}_{\mathbf{J},s})\right)| \\
&\leqslant 2\|g\|_\infty \mathbb{E}\,|f(\mathbf{X_I}) - f(\mathbf{X}_{\mathbf{I},s})| + 2\|f\|_\infty \mathbb{E}\,|g(\mathbf{X_J}) - g(\mathbf{X}_{\mathbf{J},s})| \\
&\leqslant 2\mathrm{Lip}(f)\mathbb{E}\delta_u(\mathbf{X_I}, \mathbf{X}_{\mathbf{I},s}) + 2\mathrm{Lip}(g)\mathbb{E}\delta_v(\mathbf{X_J}, \mathbf{X}_{\mathbf{J},s}) \\
&= 2\mathrm{Lip}(f)\sum_{k=1}^{u} \mathbb{E}|X_{\mathbf{i}_k} - X_{\mathbf{i}_k,s}| + 2\mathrm{Lip}(g)\sum_{k=1}^{v} \mathbb{E}|X_{\mathbf{j}_k} - X_{\mathbf{j}_k,s}| \\
&\leqslant 2(u\mathrm{Lip}(f) + v\mathrm{Lip}(g))\Delta_p(s), \quad (10)
\end{aligned}$$

for any $p \geqslant 1$. Therefore, it is enough to choose $\eta(l) = 2\Delta_1(l/2-1)$. In particular, as the series b_k is square summable, we can take $p = 2$ to obtain from (9) that

$$\eta^2(2l) = 4\Delta_2^2(l) = 4\sum_{\|\mathbf{j}\|>l} b_{\mathbf{j}}^2. \quad (11)$$

Application 2 (Markovian fields). Let $\mathbf{v} \in \mathbb{Z}^d$, define a shift operator $B_{\mathbf{v}}$ in the fields on $\mathbb{Z}^d$ as $(B_{\mathbf{v}} \cdot X)_{\mathbf{t}} = X_{\mathbf{t}-\mathbf{v}}$. Now, consider a finite sequence of real numbers $(a_j)_{j=1,\ldots,D}$ and a finite sequence $(\mathbf{v}_1, \ldots, \mathbf{v}_D) \in (\mathbb{Z}^d)^D$. A Markovian field is defined by the neighbour regression formula:

$$X_{\mathbf{i}} = \sum_{j=1}^{D} a_j X_{\mathbf{i}-\mathbf{v}_j} + \xi_{\mathbf{i}} = (A \cdot X)_{\mathbf{i}} + \xi_{\mathbf{i}}, \quad (12)$$

where $A = \sum_{j=1}^{D} a_j B_{\mathbf{v}_j}$. Assume that $a = \sum_{j=1}^{D} |a_j| < 1$, then there exists an integrable stationary solution to (12):

$$X_{\mathbf{i}} = \sum_{p=0}^{\infty} (A^p \cdot \xi)_{\mathbf{i}} = \sum_{p=0}^{\infty} \sum_{\substack{0 \leqslant j_1, \ldots, j_D \\ j_1 + \cdots + j_D = p}} \frac{p!}{j_1! \cdots j_D!} a_1^{j_1} \cdots a_D^{j_D} \xi_{\mathbf{i}-(j_1\mathbf{v}_1 + \cdots + j_D\mathbf{v}_D)}. \quad (13)$$

Note that

$$\sum_{\substack{0\leqslant j_1,\dots,j_D\\ j_1+\cdots+j_D=p}} \frac{p!}{j_1!\cdots j_D!}|a_1^{j_1}\cdots a_D^{j_D}| = a^p,$$

therefore the process (12) can be rewritten as (7) with absolutely summable coefficients, i.e.

$$X_{\mathbf{i}} = \sum_{\mathbf{k}\in\mathbb{Z}^d} b_{\mathbf{k}}\xi_{\mathbf{i}-\mathbf{k}},$$

where the series $(b_{\mathbf{k}})$ *is, by definition, such that*

$$|b_{\mathbf{k}}| \leqslant \sum_{p=0}^{\infty} \sum_{(j_1,\dots,j_D)\in V_{\mathbf{k},p}} \frac{p!}{j_1!\cdots j_D!}|a_1^{j_1}\cdots a_D^{j_D}|,$$

$$\text{with}\quad V_{\mathbf{k},p} := \{(j_1,\dots j_D)\in\mathbb{N}^D \ :\ j_1+\cdots+j_D = p,\ j_1\mathbf{v}_1+\cdots j_D\mathbf{v}_D = \mathbf{k}\}.$$

Finally, denoting $v = \max\{\|\mathbf{v}_1\|_\infty,\dots,\|\mathbf{v}_D\|_\infty\}$, *observe that* $V_{k,p}$ *is empty if* $p < \|\mathbf{k}\|_\infty/v$, *and* $|b_{\mathbf{k}}| \leqslant (1-a)^{-1}a^{\|\mathbf{k}\|_\infty/v}$. *So that* $(b_{\mathbf{k}})$ *is square summable and* X $\eta-$*weakly dependent such that (8) holds.*

Example 2 (Chaotic Volterra fields). Assume that $(\xi_{\mathbf{i}})_{\mathbf{i}\in\mathbb{Z}^d}$ has finite moments of any order. Define $(X_{\mathbf{i}})_{\mathbf{i}\in\mathbb{Z}^d}$ as

$$X_{\mathbf{i}} = \sum_{s=1}^{\infty} X_{\mathbf{i}}^{(s)} \quad \text{where} \quad X_{\mathbf{i}}^{(s)} = \sum_{\mathbf{j}_1,\dots,\mathbf{j}_s\in\mathbb{Z}^d} a^{(s)}_{\mathbf{j}_1,\dots,\mathbf{j}_s}\xi_{\mathbf{i}-\mathbf{j}_1}\cdots\xi_{\mathbf{i}-\mathbf{j}_s}. \tag{14}$$

If the real series $a^{(s)}_{\mathbf{j}_1,\dots,\mathbf{j}_s}$ is absolutely summable, the field $(X_{\mathbf{i}})$ is then $\eta-$weakly dependent with

$$\eta(2l) \leqslant \sum_{s=1}^{\infty}\sum_{t=1}^{s} \sum_{\substack{\|\mathbf{j}_t\|>l\\ \mathbf{j}_1,\dots,\mathbf{j}_s\in\mathbb{Z}^d}} |a^{(s)}_{\mathbf{j}_1,\dots,\mathbf{j}_s}|\ \mathbb{E}|\xi_{\mathbf{0}}|^s.$$

Note that if the random field is causal, that is, if the indices $\mathbf{j}_l$ are all in $\mathbb{N}^d$, then the bound holds for $\eta(l)$.

Application 3 (LARCH$(\infty)-$*fields). Let a be a real positive,* $(b_{\mathbf{j}})_{\mathbf{j}\in T^+}$ *a nonnegative sequence and* $(\varsigma_{\mathbf{j}})_{\mathbf{j}\in T^+}$ *an i.i.d nonnegative random field where* $T^+ = \mathbb{N}^d\backslash\{\mathbf{0}\}$. *Define* $X = (X_{\mathbf{i}})_{\mathbf{i}}$ *through the recurrence relation*

$$X_{\mathbf{i}} = \left(a + \sum_{\mathbf{j}\in T^+} b_{\mathbf{j}}X_{\mathbf{i}-\mathbf{j}}\right)\varsigma_{\mathbf{i}}.$$

If $c = \mathbb{E}(\varsigma_{\mathbf{0}}) \sum_{\mathbf{j} \in T^+} b_{\mathbf{j}} < 1$, *in* [7] *is proved that such models have a stationary representation with the chaotic expansion:*

$$X_{\mathbf{i}} = a\varsigma_{\mathbf{i}} + a \sum_{l=1}^{\infty} X_{\mathbf{i}}^{l}, \qquad \textit{where}$$
$$X_{\mathbf{i}}^{l} = \sum_{\mathbf{j}_1 \in T^+} \cdots \sum_{\mathbf{j}_l \in T^+} b_{\mathbf{j}_1} \cdots b_{\mathbf{j}_l} \varsigma_{\mathbf{i}-\mathbf{j}_1} \cdots \varsigma_{\mathbf{i}-(\mathbf{j}_1+\cdots+\mathbf{j}_l)}.$$

Note that the convergence of the series over l *comes from* $\mathbb{E}|X_{\mathbf{i}}^{l}| \leqslant c^l$. *Now, let*

$$X_{\mathbf{i}}^{L,m} = a\varsigma_{\mathbf{i}} + a \sum_{l=1}^{L} \sum_{\mathbf{j}_1 \in [0:m]^d} \cdots \sum_{\mathbf{j}_l \in [0:m]^d} b_{\mathbf{j}_1} \cdots b_{\mathbf{j}_l} \varsigma_{\mathbf{i}-\mathbf{j}_1} \cdots \varsigma_{\mathbf{i}-(\mathbf{j}_1+\cdots+\mathbf{j}_l)},$$

with the notation $[0:m] = \{0, 1, \ldots, m\}$. *Observe that this field* $(X_{\mathbf{i}}^{L,m})_{\mathbf{i}}$ *is an approximation of* X, *such that*

$$\mathbb{E}\left|X_{\mathbf{i}} - X_{\mathbf{i}}^{L,m}\right| \leqslant \sum_{l>L} c^l + \rho(m),$$

where $\rho(m) = \sum_{\mathbf{j} \notin [0:m]^d} |b_{\mathbf{j}}|$. *Then,* X *is* $\eta-$*weakly dependent with coefficient*

$$\eta(l) = \min_{Lm \leqslant l} \left(\frac{c^{L+1}}{1-c} + \rho(m) \right).$$

In particular, for standard LARCH models with delay p, $b_{\mathbf{j}} = 0$ *for all* $\|\mathbf{j}\| > p$,*the coefficient* $\eta(l) = c^{l/p}/(1-c)$. *In the arithmetic decay case,* $b_{\mathbf{j}} = C\|\mathbf{j}\|^{-a}$, $\eta(l) = Const. \left(\frac{l}{\log l}\right)^{1-a}$. *In the geometric case,* $b_{\mathbf{j}} = C\exp(-b\|\mathbf{j}\|)$, $\eta(l) = Const.\gamma^{\sqrt{l}}$, *for* $\gamma = \exp(-\sqrt{-b\log c})$.

Example 3 (Associated random fields). Associated random fields are $\kappa-$weakly dependent, with $\kappa(l) = \sum_{\|\mathbf{j}\|>l} \mathrm{Cov}(X_{\mathbf{0}}, X_{\mathbf{j}})$. See [5].

For other examples of weakly dependent random fields, we defer a reader to [3,6,8]. Those references also include some examples of weakly dependent random fields with weakly dependent inputs $\xi = (\xi_{\mathbf{i}})_{\mathbf{i} \in \mathbb{Z}^d}$.

3.3 An Example of Non-mixing Weakly Dependent Linear Field

Let $\xi = (\xi_i)_{i \in \mathbb{Z}}$ and $\varsigma = (\varsigma_j)_{j \in \mathbb{Z}}$ be two i.i.d Bernoulli sequences with the same parameter $p = 1/2$ such that ξ and ς are independents. Define $(U_i)_{i \in \mathbb{Z}}$ and $(V_j)_{j \in \mathbb{Z}}$ as

$$U_i = \sum_{k=0}^{\infty} \frac{\xi_{i-k}}{2^k}, \qquad V_j = \sum_{k=0}^{\infty} \frac{\varsigma_{j-k}}{2^k}.$$

The sequences $(U_i)_i$ and $(V_j)_j$ are stationaries with uniform common marginal distribution over $[0,1]$. Moreover, these are weakly dependent (see [5]) but non-mixing (because U_0 is a deterministic function of U_i for any $i>0$. The same for V, see [2]).

Consider now the random field

$$X_{(i,j)} = U_i V_j = \sum_{k\geqslant 0, l\geqslant 0} \frac{\xi_{i-k}\varsigma_{j-l}}{2^{k+l}}. \tag{15}$$

This is clearly a stationary linear random field with innovations $\xi_i\varsigma_j$.
Note that

$$\Delta_2^2(r) = \mathbb{E}\left| X_{(i,j)} - \sum_{0\leqslant k+l\leqslant r} \frac{\xi_{i-k}\varsigma_{j-l}}{2^{k+l}}\right|^2 = \sum_{\substack{k+l>r\\ k\geqslant 0, l\geqslant 0}} 4^{-(k+l)},$$

therefore $\Delta_2(r)\leqslant \sqrt{2r}2^{-r}$. Using (8), we can easy prove that the random field $(X_{(i,j)})_{i,j\in\mathbb{Z}}$ is $\eta-$weakly dependent with $\eta(r)=2^{2-r/2}\sqrt{r}$. However, this field is non-mixing, as shown in [6, § 2.2.6].

4 Asymptotic Max-Independence in Stationary Weakly Dependent Random Fields

In this section we provide the main preliminary results in order to prove extremal types theorem in the next section. For this, consider the following notations, definitions and conditions, which we will use throughout this paper.

Let $X=\{X_{\mathbf{t}}:\mathbf{t}\in\mathbb{Z}^d\}$ be a stationary random field. For subsets B of $\mathbb{Z}^d$, we denote

$$M(B)=\sup\{X_{\mathbf{t}}:\mathbf{t}\in B\},$$

and we will write

$$M_{\mathbf{n}}=M(E_{\mathbf{n}}),$$

where $\mathbf{n}=(n_1,n_2,\ldots,n_d)\in\mathbb{N}^d$ and $E_{\mathbf{n}}=\prod_{i=1}^d[n_i]$. Here $\prod_{i=1}^d A_i$ denote the cartesian product $A_1\times\cdots\times A_d$ and $[k]:=[1:k]$, where $[i:j]$ denote the subset $\{i,i+1,\ldots,j-1,j\}$ of $\mathbb{N}$.

Weak dependence conditions will be given here in such a way that the extremal types theorem holds for $M_{\mathbf{n}}$, *i.e.* such that any non-degenerate limit G for $M_{\mathbf{n}}$ normalised as in (3) must be of extreme value type. A similar result is developed in [16] but under mixing conditions (more precisely, under "Cw-mixing" conditions).

We will use a spatial version of the Bernstein block technique for the proof of Lemma 1. In order to develop this technique here, it is necessary to consider for each $i\in[d]$, a sequence $r_i:=r_{n_i}=o(n_i)$ to build the lengths of sides of $d-$blocks (or $d-$lattices):

$$B_{j_1j_2\ldots j_d}:=\prod_{i=1}^d[(j_i-1)r_i+1:j_ir_i],$$

which will be used for subdivision of $E_{\mathbf{n}} = \prod_{i=1}^{d}[n_i]$. Moreover, we denote $m_i = m_{n_i} = \lceil n_i/r_i \rceil$ for $i \in [d]$ with $\lceil x \rceil := \max\{j \in \mathbb{N} : j \leqslant x\}$. Then, it is clear that $E_{\mathbf{n}}$ contains $m_{\mathbf{n}} = m_1 m_2 \cdots m_d$ complete blocks, and no more that $(m_1 + m_2 + \cdots + m_d - d + 1)$ incomplete ones.

Spatial weak dependence (SWD) conditions. Let $l_i := l_{n_i} \underset{n_i \to \infty}{\longrightarrow} \infty$ be a sequence such that $l_i = o(r_i)$ for each $i \in [d]$. We say that a random field X satisfy at least one SWD condition if X satisfy at least one of the following dependence conditions:

1. θ-weakly dependent such that for each i-direction, with $i \in [d]$,

$$\frac{n_i^{\alpha_i} n_{i-1}^{\alpha_{i-1}} \cdots n_1^{\alpha_1}}{m_{i-1} \cdots m_2 m_1} \theta(l_i) \longrightarrow 0, \quad \text{as} \quad (n_1, \ldots, n_i) \to \infty; \tag{16}$$

   for some $(\alpha_1, \ldots, \alpha_i) \in [1, \infty)^i \setminus \{(1, \ldots, 1)\}$.
2. η-weakly dependent such that for each i-direction, with $i \in [d]$,

$$\frac{n_i^{\alpha_i} n_{i-1}^{\alpha_{i-1}} \cdots n_1^{\alpha_1} m_i^{\beta}}{m_{i-1} \cdots m_2 m_1} \eta(l_i) \longrightarrow 0, \quad \text{as} \quad (n_1, \ldots, n_i) \to \infty; \tag{17}$$

   for some $(\alpha_1, \ldots, \alpha_i, \beta) \in [1, \infty)^{i+1} \setminus \{(1, \ldots, 1)\}$.
3. κ-weakly dependent such that for each i-direction, with $i \in [d]$,

$$\left(\frac{n_i^{\alpha_i} n_{i-1}^{\alpha_{i-1}} \cdots n_1^{\alpha_1}}{m_{i-1} \cdots m_2 m_1} \right)^2 \kappa(l_i) \longrightarrow 0, \quad \text{as} \quad (n_1, \ldots, n_i) \to \infty; \tag{18}$$

4. λ-weakly dependent such that for each i-direction, with $i \in [d]$,

$$\left(\frac{n_i^{\alpha_i} n_{i-1}^{\alpha_{i-1}} \cdots n_1^{\alpha_1}}{m_{i-1} \cdots m_2 m_1} \right)^2 \lambda(l_i) \longrightarrow 0, \quad \text{as} \quad (n_1, \ldots, n_i) \to \infty; \tag{19}$$

   for some $(\alpha_1, \ldots, \alpha_i) \in [1, \infty)^i \setminus \{(1, \ldots, 1)\}$.
   In these items, $\mathbf{n} = (n_1, n_2, \ldots, n_d) \to \infty$ means that $n_j \to \infty$ for each $j \in [d]$. Besides, we set the convention $m_0 = 1$.

Lemma 1. *Let $X = \{X_{\mathbf{t}} : \mathbf{t} \in \mathbb{Z}^d\}$ be a stationary random field that satisfy at least one SWD condition. Let $(u_{\mathbf{n}})_{\mathbf{n} \in \mathbb{N}^d}$ be a family of levels such that $\mathbb{P}(M(B) = u_{\mathbf{n}}) = 0$, for all $B \subset \mathbb{Z}^d$ and all $\mathbf{n} \in \mathbb{N}^d$. Then if $B_{\mathbf{n}} = \prod_{i=1}^{d}[m_i r_i]$ (this is, $B_{\mathbf{n}} = \bigcup_{j_1, \cdots, j_d} B_{j_1, \cdots, j_d}$),*

$$\mathbb{P}(M(B_{\mathbf{n}}) \leqslant u_{\mathbf{n}}) = \mathbb{P}^{m_{\mathbf{n}}}(M(\mathbf{J}) \leqslant u_{\mathbf{n}}) + o(1) \tag{20}$$

as $\mathbf{n} \to \infty$, where $\mathbf{J} := B_{11\ldots1} = \prod_{i=1}^{d}[r_i]$.

Proof. W.l.o.g., we will consider $d = 2$ to reduce notations and simplify the proof.

Let $\mathbf{J}_i = [(i-1)r_1 + 1 : ir_1] \times [m_2 r_2]$, $\mathbf{J}'_i = [(i-1)r_1 + l_1 : ir_1 - l_1] \times [m_2 r_2]$ and $\mathbf{J}^*_i = \mathbf{J}_i \setminus \mathbf{J}'_i$. It is evident then that $B_{\mathbf{n}} = \bigcup_{i=1}^{m_1} \mathbf{J}_i$ and

$$\begin{aligned}\mathbb{P}\left(M\left(\bigcup_{i=1}^{m}\mathbf{J}_i\right) \leqslant u_{\mathbf{n}}\right) &= \mathbb{P}\left(M\left(\bigcup_{i=1}^{m-1}\mathbf{J}_i\right) \leqslant u_{\mathbf{n}}\ ,\ M(\mathbf{J}_m) \leqslant u_{\mathbf{n}}\right)\\ &\leqslant \mathbb{P}\left(M\left(\bigcup_{i=1}^{m-1}\mathbf{J}_i\right) \leqslant u_{\mathbf{n}}\ ,\ M(\mathbf{J}'_m) \leqslant u_{\mathbf{n}}\right), \qquad (21)\end{aligned}$$

for any $m \in [2 : m_1]$. Thus

$$\begin{aligned}0 &\leqslant \mathbb{P}\left(M\left(\bigcup_{i=1}^{m-1}\mathbf{J}_i\right) \leqslant u_{\mathbf{n}}\ ,\ M(\mathbf{J}'_m) \leqslant u_{\mathbf{n}}\right) - \mathbb{P}\left(M\left(\bigcup_{i=1}^{m}\mathbf{J}_i\right) \leqslant u_{\mathbf{n}}\right)\\ &= \mathbb{P}\left(M\left(\bigcup_{i=1}^{m-1}\mathbf{J}_i\right) \leqslant u_{\mathbf{n}}\ ,\ M(\mathbf{J}'_m) \leqslant u_{\mathbf{n}}\ ,\ M\left(\bigcup_{i=1}^{m}\mathbf{J}_i\right) > u_{\mathbf{n}}\right)\\ &\leqslant \mathbb{P}(M(\mathbf{J}^*_m) > u_{\mathbf{n}}) = \mathbb{P}(M(\mathbf{J}^*_1) > u_{\mathbf{n}}). \qquad (22)\end{aligned}$$

On the other hand, observe that

$$0 \leqslant \mathbb{P}(M(\mathbf{J}'_m) \leqslant u_{\mathbf{n}}) - \mathbb{P}(M(\mathbf{J}_m) \leqslant u_{\mathbf{n}}) \leqslant \mathbb{P}(M(\mathbf{J}^*_m) > u_{\mathbf{n}}) = \mathbb{P}(M(\mathbf{J}^*_1) > u_{\mathbf{n}}). \qquad (23)$$

Using stationarity together with (22)–(23), it follows that

$$\begin{aligned}&\left|\mathbb{P}\left(M\left(\bigcup_{i=1}^{m}\mathbf{J}_i\right) \leqslant u_{\mathbf{n}}\right) - \mathbb{P}\left(M\left(\bigcup_{i=1}^{m-1}\mathbf{J}_i\right) \leqslant u_{\mathbf{n}}\right)\mathbb{P}\left(M(\mathbf{J}_1) \leqslant u_{\mathbf{n}}\right)\right|\\ &\leqslant \left|\mathbb{P}\left(M\left(\bigcup_{i=1}^{m}\mathbf{J}_i\right) \leqslant u_{\mathbf{n}}\right) - \mathbb{P}\left(M\left(\bigcup_{i=1}^{m-1}\mathbf{J}_i\right) \leqslant u_{\mathbf{n}}\ ,\ M(\mathbf{J}'_m) \leqslant u_{\mathbf{n}}\right)\right|\\ &+ \left|\mathbb{P}\left(M\left(\bigcup_{i=1}^{m-1}\mathbf{J}_i\right) \leqslant u_{\mathbf{n}}\ ,\ M(\mathbf{J}'_m) \leqslant u_{\mathbf{n}}\right)\right.\\ &\quad \left.-\mathbb{P}\left(M\left(\bigcup_{i=1}^{m-1}\mathbf{J}_i\right) \leqslant u_{\mathbf{n}}\right)\mathbb{P}\left(M(\mathbf{J}'_m) \leqslant u_{\mathbf{n}}\right)\right|\\ &+ \mathbb{P}\left(M\left(\bigcup_{i=1}^{m-1}\mathbf{J}_i\right) \leqslant u_{\mathbf{n}}\right)\left|\mathbb{P}\left(M(\mathbf{J}'_m) \leqslant u_{\mathbf{n}}\right) - \mathbb{P}\left(M(\mathbf{J}_m) \leqslant u_{\mathbf{n}}\right)\right|, \qquad (24)\end{aligned}$$

which implies that

$$\begin{aligned}
&\delta_{\mathbf{n}} := |\mathbb{P}(M(B_{\mathbf{n}}) \leqslant u_{\mathbf{n}}) - \mathbb{P}^{m_1}(M(\mathbf{J}_1) \leqslant u_{\mathbf{n}})| \\
&= \left| \sum_{j=1}^{m_1-1} \left[\mathbb{P}\left(M\left(\bigcup_{i=1}^{m_1-j+1} \mathbf{J}_i \right) \leqslant u_{\mathbf{n}} \right) - \mathbb{P}\left(M\left(\bigcup_{i=1}^{m_1-j} \mathbf{J}_i \right) \leqslant u_{\mathbf{n}} \right) \mathbb{P}(M(\mathbf{J}_{m_1-j+1}) \leqslant u_{\mathbf{n}}) \right] \right. \\
&\qquad \left. \times \mathbb{P}^{j-1}(M(\mathbf{J}_1) \leqslant u_{\mathbf{n}}) \right| \\
&\qquad \leqslant 2(m_1-1)\mathbb{P}\{M(J_1^*) > u_{\mathbf{n}}\} + \sum_{j=1}^{m_1-1} \Delta_{j+1,x}, \qquad (25)
\end{aligned}$$

where

$$\begin{aligned}
\Delta_{j,x} := \Bigg| &\mathbb{P}\left(M\left(\bigcup_{i=1}^{j-1} \mathbf{J}_i \right) \leqslant u_{\mathbf{n}} \ , \ M(\mathbf{J}'_j) \leqslant u_{\mathbf{n}} \right) \\
&- \mathbb{P}\left(M\left(\bigcup_{i=1}^{j-1} \mathbf{J}_i \right) \leqslant u_{\mathbf{n}} \right) \mathbb{P}\left(M(\mathbf{J}'_j) \leqslant u_{\mathbf{n}} \right) \Bigg| .
\end{aligned}$$

Note that $\Delta_{j,x}$ is the absolute value of the covariance of two indicators functions, which are not Lipschitz. An approximation by Lipschitz functions is used in order to bound the expressions of interest by using such weak-dependence conditions.

Let $\mathbb{M} := \mathbb{M}_{a\times b}(\mathbb{R})$ be the set of real-matrices with a rows and b columns. For $u > 0$, define the real function $f_{u,a,b} : \mathbb{M} \longrightarrow \mathbb{R}$, such that

$$f_{u,a,b}((x_{ij})_{ij}) = \mathbb{I}_{\{\max\{x_{ij} \ : \ (i,j)\in[a]\times[b]\}\leqslant u\}}. \qquad (26)$$

This function has jump discontinuities at the points (matrices) $(x_{ij})_{ij} \in \mathbb{M}$ such that

$$\max\{x_{ij} \ : \ (i,j) \in [a] \times [b]\} = u,$$

i.e., at all the discontinuity points of $f_{u,a,b}$:

$$D(f_{u,a,b}) := \{(x_{ij})_{ij} \in \mathbb{M} \ : \ \max\{x_{ij} \ : \ (i,j) \in [a] \times [b]\} = u\}.$$

Consider now the polyhedral function $K_{a,b} : \mathbb{M} \longrightarrow \mathbb{R}$, defined as:

$$\begin{aligned}
K_{a,b}((x_{ij})_{ij}) := &\mathbb{I}_{\{\max\{x_{ij} \ :(i,j)\in[a]\times[b]\}\leqslant -1\}} \\
&- \frac{1}{2}\sum_{i=1}^{a}\sum_{j=1}^{b}(x_{ij}-1)\mathbb{I}_{\{|x_{ij}|<1,\ x_{ij}\geqslant x_{kl}\ \forall(k,l)\neq(i,j)\}} + \cdots .
\end{aligned}$$

Consider a positive sequence h_n converging to zero, as $n \to \infty$, then the sequence of Lipschitz functions

$$K_{n,a,b}(A) := K_{a,b}(h_n^{-1}(A-U)), \qquad (27)$$

converges uniformly to $f_{u,a,b}$ on $\mathbb{M} \setminus D(f_{u,a,b})$, where $U := (u_{ij})_{ij} \in \mathbb{M}$ is such that $u_{ij} = u$ for all $(i,j) \in [a] \times [b]$.

Moreover,

$$\frac{1}{abh_n} \leqslant \mathrm{Lip}(K_{n,a,b}) = \frac{1}{h_n}. \tag{28}$$

We can thus use the approximation $\Delta_{j,x} = \Delta_{\mathbf{n},j,x} + o(1)$, where

$$\Delta_{\mathbf{n},j,x} := \left| \mathrm{Cov}\left(K_{\mathbf{n},m_2r_2,(j-1)r_1}\left(\bigcup_{i=1}^{j-1} \mathbf{J}_i \right), K_{\mathbf{n},m_2r_2,r_1-2l_1}\left(\mathbf{J}'_j \right) \right) \right|.$$

We will show now that $\delta_{\mathbf{n}} \longrightarrow 0$ as $\mathbf{n} \to \infty$.

- Decay of the dependence term.

  First, if we suppose that X is η-weakly dependent, we obtain for $j \in [2 : m_1]$ that

$$\begin{aligned} &\Delta_{\mathbf{n},j,x} \\ &\leqslant \psi\left(\mathrm{Lip}(K_{\mathbf{n},m_2r_2,(j-1)r_1}), \mathrm{Lip}(K_{\mathbf{n},m_2r_2,r_1-2l_1}), m_2(j-1)r_2r_1, m_2r_2(r_1-2l_1)\right)\eta(l_1) \\ &\leqslant \frac{m_2r_2r_1(j-2l_1/r_1)}{h_{\mathbf{n}}}\eta(l_1), \end{aligned}$$

  thus

$$\begin{aligned} \sum_{j=1}^{m_1-1} \Delta_{\mathbf{n},j+1,x} &\leqslant \frac{n_2n_1m_1}{h_n}\left[\frac{1-m_1^{-1}}{2} + \left(\frac{1}{m_1} - \frac{2l_1}{r_1m_1}\right)\left(1-\frac{1}{m_1}\right)\right]\eta(l_1) \\ &\leqslant 2^{-1}n_2n_1^{\alpha_1}m_1^{\beta}\left[1 + \frac{2}{m_1}\left(1-\frac{2l_1}{r_1}\right)\right]\eta(l_1), \end{aligned} \tag{29}$$

  where we have set here $h_{\mathbf{n}} = n_1^{1-\alpha_1}m_1^{1-\beta}$, for some $(\alpha_1, \beta) \in [1,\infty)^2 \setminus \{(1,1)\}$

  In the same way, if we suppose that X is respectively κ, λ or $\theta-$weakly dependent, we obtain respectively that

$$\sum_{j=1}^{m_1-1} \Delta_{\mathbf{n},j+1,x} \leqslant \frac{(n_2n_1^{\alpha_1})^2}{2}(1-\frac{2l_1}{r_1})\kappa(l_1), \tag{30}$$

$$\sum_{j=1}^{m_1-1} \Delta_{\mathbf{n},j+1,x} \leqslant \frac{(n_2n_1^{\alpha_1})^2}{2}\left[\frac{h_nm_1}{n_2n_1}\left(1 + \frac{2(1-2\frac{l_1}{r_1})}{m_1}\right) + (1-\frac{2l_1}{r_1})\right]\lambda(l_1), \tag{31}$$

$$\sum_{j=1}^{m_1-1} \Delta_{\mathbf{n},j+1,x} \leqslant n_2n_1^{\alpha_1}\left(1-\frac{2l_1}{r_1}\right)\left(1-\frac{1}{m_1}\right)\theta(l_1), \tag{32}$$

  where we have set $h_{\mathbf{n}} = n_1^{1-\alpha_1}$ for κ, λ and θ cases, for some $\alpha_1 \in (1,\infty)$.

- Decay of the expectation of the excesses in the remaining blocks J^*.
  We only need consider that $\mathbb{P}^{m_1}\{M(J_1^*) \leqslant u_{\mathbf{n}}\} \longrightarrow \rho$ for some $\rho \in [0,1]$. In fact, if $\rho = 1$, then $m_1 \log \mathbb{P}\{M(\mathbf{J}_1^*) \leqslant u_{\mathbf{n}}\} \longrightarrow 0$ as $\mathbf{n} \to \infty$.

  Therefore, $m_1 \mathbb{P}\{M(\mathbf{J}_1^*) > u_{\mathbf{n}}\} \longrightarrow 0$ as $\mathbf{n} \to \infty$. Using this and the respective SWD condition (16)–(19), it follows that $\delta_{\mathbf{n}} \longrightarrow 0$ by (25).

  Now, we suppose that $\rho < 1$. Note that for $\mathbf{n}$ sufficiently large, we can choose $k_1 = k_{1,\mathbf{n}}$ rectangles $\{\mathbf{J}_{1j}^*\}_{j\in[k_1]}$ inside $\mathbf{J}_1'$, congruents to $\mathbf{J}_1^*$, such that all are mutually separated by at least l_1 points in the 1-direction. Using similar arguments to (22)–(25), but in the (22) and (23) cases we only use the left hand inequality, we obtain

$$\mathbb{P}(M(\mathbf{J}_1) \leqslant u_{\mathbf{n}}) \leqslant \mathbb{P}^{k_1}(M(\mathbf{J}_1^*) \leqslant u_{\mathbf{n}}) + \sum_{j=1}^{k_1-1} \tilde{\Delta}_{\mathbf{n},j+1,x} + o(1) \tag{33}$$

  where $\tilde{\Delta}_{\mathbf{n},j,x} = \left|\mathrm{Cov}\left(K_{n,m_2r_2,2(j-1)l_1}\left(\bigcup_{i=1}^{j-1}\mathbf{J}_{1i}^*\right), K_{n,m_2r_2,2l_1}(\mathbf{J}_{1j}^*)\right)\right|$.
  Thus, for $\mathbf{n}$ large, we get

$$\begin{aligned}\mathbb{P}^{m_1}(M(\mathbf{J}_1) \leqslant u_{\mathbf{n}}) &\leqslant \mathbb{P}^{k_1 m_1}(M(\mathbf{J}_1^*) \leqslant u_{\mathbf{n}}) + \sum_{j=1}^{k_1-1} \Delta_{\mathbf{n},j+1,x} + o(1) \\ &= (\rho + o(1))^{k_1} + o(1) \longrightarrow 0,\end{aligned} \tag{34}$$

  because $\rho < 1$. Hence $\delta_{\mathbf{n}} = \mathbb{P}(M(B_{\mathbf{n}}) \leqslant u_{\mathbf{n}}) + o(1)$. However, it follows similarly that

$$\mathbb{P}(M(B_{\mathbf{n}}) \leqslant u_{\mathbf{n}}) \leqslant \mathbb{P}^{m_1}\{M(\mathbf{J}_1') \leqslant u_{\mathbf{n}}\} + \sum_{j=1}^{m_1-1} \Delta_{\mathbf{n},j+1,x} + o(1)$$

  which tends to zero since (33) and hence (34) also apply with $M(\mathbf{J}_1')$ in place of $M(\mathbf{J}_1)$. In consequence, $\delta_{\mathbf{n}} \longrightarrow 0$ as $\mathbf{n} \to \infty$ whenever $\rho < 1$. Therefore, (25) again holds.

Now to prove that $\mathbb{P}(M(B_{\mathbf{n}}) \leqslant u_{\mathbf{n}}) = \mathbb{P}^{m_1 m_2}(M(\mathbf{J}) \leqslant u_{\mathbf{n}}) + o(1)$, it will be suffice to show that:

$$\mathbb{P}^{m_1}(M(\mathbf{J}_1) \leqslant u_{\mathbf{n}}) - \mathbb{P}^{m_1 m_2}(M(\mathbf{J}) \leqslant u_{\mathbf{n}}) \longrightarrow 0.$$

However, this follows by entirely similar reasoning, splitting the rectangle $\mathbf{J}_1$ into rectangles $H_i = [1 : r_1] \times [(i-1)r_2 + 1 : r_2]$, for $i \in [m_2]$, and bearing in mind that the Lipschitz approximations are $K_{\mathbf{n},r_1,(j-1)r_2}$ and $K_{\mathbf{n},r_1,r_2-2l_2}$, defined in smaller domains of size $u = r_2 r_1(j-1)$ and $v = r_1(r_2 - 2l_2)$, respectively. □

Replacing [16, Lemma 3.1] by our Lemma 1 in the proof of Proposition 3.2 in [16], we prove the following result.

Proposition 2. *Let $X = \{X_{\mathbf{t}} : \mathbf{t} \in \mathbb{Z}^d\}$ be a stationary random field that satisfy at least one SWD condition. Suppose that $(u_{\mathbf{n}})_{\mathbf{n}\in\mathbb{N}^d}$ is a family of levels such that for all $B \subset \mathbb{Z}^d$, and all $\mathbf{n} \in \mathbb{N}^d$: $\mathbb{P}(M(B) = u_{\mathbf{n}}) = 0$.*

Let $\mathbf{I} = \prod_{i=1}^{d} [\lceil a_i n_i \rceil]$ (with $a_i \in (0, 1]$ for all $i \in [d]$) be a d-sub-grid of $E_{\mathbf{n}}$, where a_i may change with $\mathbf{n}$ for all $i \in [n]$ but $a_1 a_2 \cdots a_d \longrightarrow a > 0$ as $\mathbf{n} \to \infty$. Then, under the above notation

1. $\mathbb{P}\left(M(\mathbf{I}) \leqslant u_{\mathbf{n}}\right) - \mathbb{P}^{am_{\mathbf{n}}}\left(M(\mathbf{J}) \leqslant u_{\mathbf{n}}\right) \longrightarrow 0.$
2. $\mathbb{P}\left(M(\mathbf{I}) \leqslant u_{\mathbf{n}}\right) - \mathbb{P}^{a}\left(M_{\mathbf{n}} \leqslant u_{\mathbf{n}}\right) \longrightarrow 0.$

Note that if $\mathbb{P}\left(M_{\mathbf{n}} \leqslant u_{\mathbf{n}}\right)$ has a limit $G(x)$, with $u_{\mathbf{n}} = a_{\mathbf{n}} x + b_{\mathbf{n}}$, then, from 2. in the previous proposition, $\mathbb{P}(M(\mathbf{I}) \leqslant u_{\mathbf{n}})$ has the limit $G^a(x)$, which is used to show max-stability of G in the proof of Theorem 1.

5 Extremal Types Theorem and Domain of Attraction Criteria

As before, $\mathbf{n} \to \infty$ means $n_i \to \infty$ for each $i \in [d]$. However, the extremal types theorem for stationary random fields can be reformulated in terms of the limiting distribution of $M_{\mathbf{n}}$ as $\mathbf{n} \to \infty$ along a monotone path on the grid $\mathbb{N}^d$, *i.e.*, along $\mathbf{n} = (n, \lceil \vartheta_1(n) \rceil, \ldots, \lceil \vartheta_{d-1}(n) \rceil)$ for some strictly increasing continuous functions $\vartheta_j : [1, \infty) \longrightarrow [1, \infty)$, with $j \in [d-1]$, such that $\vartheta_j(T) \longrightarrow \infty$ as $T \to \infty$, for all $j \in [d-1]$.

Theorem 1. *Let $X = \{X_{\mathbf{t}} : \mathbf{t} \in \mathbb{Z}^d\}$ be a stationary random field and suppose that $\mathbb{P}\left(a_{\mathbf{n}}^{-1}(M_{\mathbf{n}} - b_{\mathbf{n}}) \leqslant x\right) \longrightarrow G(x)$, non-degenerate as $\mathbf{n} \to \infty$ along the monotone path $\mathbf{n} = (n, \lceil \vartheta_1(n) \rceil, \ldots, \lceil \vartheta_{d-1}(n) \rceil)$ defined above.*

Then, if X satisfies at least one SWD condition and $\mathbb{P}(M(B) = a_{\mathbf{n}} x + b_{\mathbf{n}}) = 0$, for all $(x, n) \in \mathbb{R} \times \mathbb{N}$ and all $B \subset \mathbb{Z}^d$, G is of extreme value type.

Proof. Let $f : [1, \infty) \longrightarrow [1, \infty)$ the continuous strictly increasing function defined by

$$f(T) = T\vartheta_1(T) \cdots \vartheta_{d-1}(T),$$

and define now

$$\phi_k(T) = f^{-1}\left(\frac{1}{k} f(kT)\right), \quad k = 1, 2, 3, \ldots.$$

If we set $T^* = \phi_k(\frac{T}{k})$, note that $k f(T^*) = f(T)$.

Then, for $\mathbf{n} = (n, \lceil \vartheta_1(n) \rceil, \ldots, \lceil \vartheta_{d-1}(n) \rceil)$, we set

$$\mathbf{I} = [\lceil a_1 n \rceil] \times [\lceil a_2 \vartheta_1(n) \rceil] \times \cdots \times [\lceil a_d \vartheta_{d-1}(n) \rceil],$$

where $a_1 = \dfrac{n^*}{n}$ and $a_{i+1} = \dfrac{\vartheta_i(n^*)}{\vartheta_i(n)}$ for $i = 1, \ldots, d-1$.

Clearly $a_i \in (0,1]$ for all $i \in [d]$, because ϑ_i is strictly increasing for all $i \in [d]$ and $n^* = f^{-1}\left(\frac{1}{k} f(n)\right) \leqslant n$ for all $(n,k) \in \mathbb{N}^2$.

Besides,

$$a_1 \cdots a_d = \frac{f(n^*)}{f(n)} = \frac{1}{k} =: a$$

Then the assumptions of Proposition 1 hold for this $\mathbf{I}$ with these $(a_i)_{i\in[d]}$ and a.

Thus,

$$\mathbb{P}\left(M_{\mathbf{n}} \leqslant u_{\mathbf{n}}\right) - \mathbb{P}^k\left(M(\mathbf{I}) \leqslant u_{\mathbf{n}}\right) \underset{\mathbf{n}\to\infty}{\longrightarrow} 0. \tag{35}$$

By writing $M_{\mathbf{n}} = M_n$, $u_{\mathbf{n}} = u_n$ and $M(\mathbf{I}) = M_{n^*}$, expression (35) is equivalent to

$$\mathbb{P}\left(M_n \leqslant u_n\right) - \mathbb{P}^k\left(M_{n^*} \leqslant u_n\right) \underset{n\to\infty}{\longrightarrow} 0, \tag{36}$$

which verifies (4) by setting $n^* = \phi_k(n/k)$ and replacing n by kT. Finally the result now follows from Proposition 1. □

5.1 On Domain of Attraction Criteria

The purpose of this section is to provide a characterisation of the maximum domain of attraction of a extreme value distribution G.

Since the random field is stationary, this characterisation will be an analogous version to the one already made for the cases: i.i.d. random variables, stationary sequences with non-zero extremal index, etc. (See [15,16]). Namely, under SWD conditions, we obtain that the type of limiting distribution for maxima is also determined by the tail behaviour of the common marginal distribution function for each either term or each maximum over (fixed) sub-blocks.

Note that from Lemma 1, we can deduce that the random variables

$$\left(M(B_{j_1,j_2,\ldots,j_d})\right)_{j_1,\ldots,j_d\in[d]}$$

have extremal index 1. Therefore, using this "max-block asymptotic independence", we can provide the characterisation through the tail distribution function of $M(\mathbf{J})$, where, by stationarity, $\mathbf{J} := B_{11\ldots1} = \prod_{i=1}^{d}[r_i]$ denotes "the first block" or a generic block. Let now $\gamma_{\mathbf{n}}$ be the $(1-m_{\mathbf{n}}^{-1})-$percentile of $M(\mathbf{J})$, *i.e.* $\mathbb{P}\left(M(\mathbf{J}) > \gamma_{\mathbf{n}}\right) = m_{\mathbf{n}}^{-1}$.

With the above notation, we state the following proposition.

Proposition 3. *Let $X = \{X_{\mathbf{t}} : \mathbf{t} \in \mathbb{Z}^d\}$ be a stationary random field that satisfies at least one SWD condition.*

Suppose that

(TD) $\mathbb{P}\left(M(\mathbf{J}) > \gamma_{\mathbf{n}} + a_{\mathbf{n}}x\right) / \mathbb{P}\left(M(\mathbf{J}) > \gamma_{\mathbf{n}}\right) \longrightarrow H(x)$ *for some constants*

$a_{\mathbf{n}} > 0$ *and some non-increasing function* $H(x)$ *such that* $H(x) \underset{x\to-\infty}{\longrightarrow} \infty$ *and* $H(x) \underset{x\to\infty}{\longrightarrow} 0$*; and that additionally* $\mathbb{P}(M(B) = a_{\mathbf{n}}x + \gamma_{\mathbf{n}}) = 0$*, for all* $B \in \mathbb{Z}^d$ *and all* $(x, \mathbf{n}) \in \mathbb{R} \times \mathbb{N}^d$.

Then

$$\mathbb{P}\left(a_{\mathbf{n}}^{-1}(M_{\mathbf{n}} - \gamma_{\mathbf{n}}) \leqslant x\right) \underset{\mathbf{n}\to\infty}{\longrightarrow} G(x) = \exp(-H(x)). \tag{37}$$

Proof. It follows from Proposition 2 - 1 (with $a_i = 1$ for all $i \in [d]$) and the assumption (TD) that

$$\begin{aligned}
\mathbb{P}\left(M_{\mathbf{n}} \leqslant a_{\mathbf{n}}x + \gamma_{\mathbf{n}}\right) &= \mathbb{P}^{m_{\mathbf{n}}}\left(M(\mathbf{J}) \leqslant a_{\mathbf{n}}x + \gamma_{\mathbf{n}}\right) + o(1) \\
&= \left(1 - H(x)\mathbb{P}\left(M(\mathbf{J}) > \gamma_{\mathbf{n}}\right)(1 + o(1))\right)^{m_{\mathbf{n}}} + o(1) \\
&= \left(1 - \frac{H(x)}{m_{\mathbf{n}}}(1 + o(1))\right)^{m_{\mathbf{n}}} + o(1) \\
&\underset{\mathbf{n}\to\infty}{\longrightarrow} \exp(-H(x)),
\end{aligned}$$

where the last equality is because $\mathbb{P}\left(M(\mathbf{J}) > \gamma_{\mathbf{n}}\right) = m_{\mathbf{n}}^{-1}$. □

Remark 3. Note that if F is the distribution function of $M(\mathbf{J})$ for fixed $\mathbf{J}$ and if the assumptions of the previous proposition are satisfied, then the relation (37) implies that F belongs to the domain of attraction of G.

6 Numerical Experiment

In this section, we estimate the generalised extreme value (GEV) distribution function of the non-mixing stationary random field (15).

First, we generate a data $\left(x_{(i_1,i_2)}\right)_{(i_1,i_2)\in D}$ from the random field (15), restricted to the domain $D = [n_1] \times [n_2]$, with $n_1 = n_2 = n = 5000$. In order to build the samplings[1], we divide the domain into $m_1 m_2 = 81$ blocks of size $[r_1] \times [r_2]$, where $r_1 = r_2 = n^{0.73}$. We take the maximum on each block and we use the L-moments (LM) estimators for the GEV distribution in order to obtain its estimated parameters (μ, σ, γ). The values resulting from this estimation are shown in Table 1. In the same table, applying a parametric bootstrap with 502 iterations, we provide also the confidence intervals for these parameters with a 95% confidence level. For these calculations we have used the `R` package "`extRemes`", see [10]. Currently, for L-moments, the only method available to calculate confidence intervals in this software is to apply a parametric bootstrap. The implementation of the method is explained on page 16 of the tutorial: https://cran.r-project.org/web/packages/extRemes/extRemes.pdf.

[1] In the context of this work, if we want to simulate a real situation, we would only have a data matrix of $n_1 \times n_2$ real values. The case would be different if we would had a space-time data with either independence or weak dependence over time because the division of the domain would be at least in the time.

Table 1. Estimation and confidence intervals for the parameters (μ, σ, γ) of the GEV distribution function fitted to the maximum of the random field (15) on the blocks of size $[r_1] \times [r_2]$.

	2.5%	Estimate	97.5%
location (μ)	0.98576	0.98739	0.98910
scale (σ)	0.00589	0.00708	0.00839
shape (γ)	−0.75382	−0.55442	−0.38049

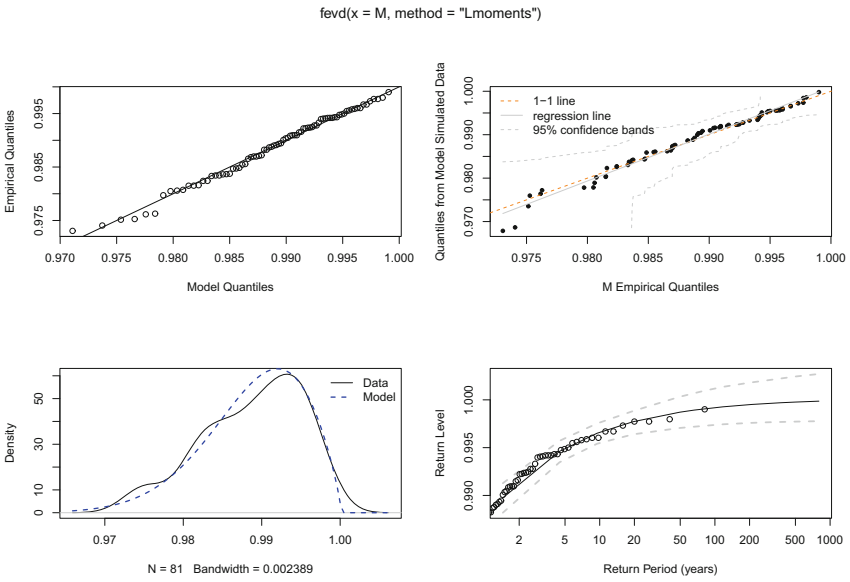


Fig. 1. Diagnostics from the GEV df fitted to the maximum of the random field (15) on the blocks of size $[r_1] \times [r_2]$. Quantile-quantile plot (top left), quantiles from a sample drawn from the fitted GEV df against the empirical data quantiles with 95% confidence bands (top right), density plots of empirical data and fitted GEV df (bottom left), and return level plot with 95% point-wise normal approximation confidence intervals (bottom right).

On the other hand, in Fig. 1, observe that the GEV distribution function $G(x) = G(x; \mu, \sigma, \gamma)$, with the estimated parameters (μ, σ, γ) shown in Table 1, is well fitted to the simulated data of the maximum of the random field (15) taken in this experiment. Note that G is here of Weibull type with parameter $\alpha = -\gamma^{-1} = 1.803677$.

Acknowledgement. This work has been developed within the MME-DII center of excellence (ANR-11-LABEX-0023-01) and with the help of PAI-CONICYT MEC No. 80170072 and the Normandy Region.

Very special thanks are due to Patrice Bertail, his remarks on the numerical experiment were very useful and helped us to improve this section of the manuscript. Thanks also to the anonymous referees, their remarks and comments helped to ameliorate this work to make it more adequate for publication in this volume.

References

1. Adler RA (1981) The Geometry of Random Fields. Wiley, New York
2. Andrews D (1984) Non strong mixing autoregressive processes. J Appl Probab 21:930–934
3. Dedecker J, Doukhan P, Lang G, León JR, Louhichi S, Prieur C (2007) Weak dependence: with examples and applications. In: Lecture notes in statistics. Springer-Verlag, p 190
4. Doukhan P (1994) Mixing: properties and examples. In: Lecture notes in statistics. Springer-Verlag, p 85
5. Doukhan P, Louhichi S (1999) A new weak dependence condition and applications to moment inequalities. Stochast Process Appl 84:313–342
6. Doukhan P, Lang G (2002) Rates in the empirical central limit theorem for stationary weakly dependent random fields. Stat Infer Stoch Process 5:199–228
7. Doukhan P, Teyssiere G, Winant P (2006) A LARCH(∞) vector valued process. In: Bertail P, Doukhan P, Soulier P (eds) Lecture notes in statistics, dependence in probability and statistics, vol 187, pp 245 – 258
8. Doukhan P, Truquet L (2007) A fixed point approach to model random fields. Alea 3:111–132
9. Gaetan C, Guyon X (2010) Spatial statistics and modeling. Springer series in statistics. Springer-Verlag, New York
10. Gilleland E, Katz RW (2016) extRemes 2.0: an extreme value analysis package in *R*. J Stat Softw 72(8):1–39. https://doi.org/10.18637/jss.v072.i08
11. Gómez JG (2018) Dependent lindeberg central limit theorem for the fidis of empirical processes of cluster functionals. Statistics 52(5):955–979. https://doi.org/10.1080/02331888.2018.1470630
12. Ferreira H, Pereira L (2008) How to compute the extremal index of stationary random fields. Stat Probab Lett 78(11):1301–1304
13. Newman M (1984) Asymptotic independence and limit theorems for positively and negatively dependent random variables. In: Tong YL (ed) Inequalities in statistics and probability. IMS lecture notes-monograph series, vol 5, pp 127–140. IMS, Hayward
14. Pereira L, Martins AP, Ferreira H (2017) Clustering of high values in random fields. Extremes 20:807–838. https://doi.org/10.1007/s10687-017-0291-7
15. Leadbetter MR, Lindgren G, Rootzén H (1983) Extremes and related properties of random sequences and processes. Springer, New York
16. Leadbetter MR, Rootzén H (1998) On extreme values in stationary random fields. In: Stochastic processes and related topics, Trends Mathematics Birkhauser, Boston, pp 275–285
17. Rosenblatt M (1956) A central limit theorem and a strong mixing condition. Proc Nat Acad Sci USA 42:43–47

Subordinated Processes with Infinite Variance

Aleksandra Grzesiek and Agnieszka Wyłomańska(✉)

Faculty of Pure and Applied Mathematics, Hugo Steinhaus Center, Wrocław University of Science and Technology, Wrocław, Poland
{aleksandra.grzesiek,agnieszka.wylomanska}@pwr.edu.pl

Abstract. In this paper we consider three models of subordinated processes. A subordinated process, called also a time-changed process, is defined as a superposition of two independent stochastic processes. To construct such stochastic system we replace the time of a given process (called also an external process) by another process which becomes the "operational time". In the literature one can find different models that are constructed as a superposition of two stochastic processes. The most classical example is the Laplace motion, also known as variance gamma process, is stated as a Brownian motion time-changed by the gamma subordinator. In this paper the considered systems are constructed by replacing the time of the symmetric α-stable Lévy motion with another stochastic process, namely the α_S-stable, tempered α_T-stable and gamma subordinator. We discuss the main characteristics of each introduced processes. We examine the characteristic function, the codifference, the probability density function, asymptotic tail behaviour and the fractional order moments. To make the application of these processes possible we propose a simulation procedure. Finally, we demonstrate how to estimate the tail index of the external process, i.e. alpha-stable Levy motion and by using Monte Carlo method we show the efficiency of the proposed estimation method.

Keywords: α-stable Lévy motion · Subordination · α-stable subordinator · Tempered α-stable subordinator · Gamma subordinator · Infinite variance

1 Introduction

A subordinated process, called also a time-changed process, is defined as a superposition of two independent stochastic processes. To construct such stochastic system we replace the time of a given process (called also an external process) by another process which becomes the "operational time". The internal process replacing the time is called subordinator and in general it is an increasing Lévy process with independent, stationary increments and cadlag sample paths. The subordinated processes were introduced in 1949 by Bochner in [1] and expounded by the same author in [2]. In the recent years their theoretical properties have

F. Chaari et al. (Eds.): CSTA 2017, ACM 16, pp. 111–135, 2020.
https://doi.org/10.1007/978-3-030-22529-2_6

been widely considered. Moreover, the time-changed processes become very useful in the real data analysis. It is related to the fact that after subordination certain characteristics of the external process are retained and at the same time certain properties of the subordinated process change.

In the real data modelling very often there is a situation when the time series demonstrates visible jumps. The fact that we can observe large outliers may suggest that a stochastic process describing the data has infinite variance. One of the most useful model of this situation is the α-stable Lévy motion introduced in 1925 by Lévy in [15]. The α-stable Lévy motion is a stochastic process starting at zero with independent and stationary increments. The increments of this process are distributed according to the α-stable distribution which is an extension of the Gaussian distribution and for most of the cases (except the Gaussian case) has diverging variance. The important property of the α-stable random variables is the role they play in the Generalized Central Limit Theorem, namely the stable probability laws attract distributions of sums of random variables with diverging variance, similarly to the Gaussian law that attracts distributions with finite variance. Another important property of the α-stable distribution (expect the Gaussian case) are the so-called heavy-tails, i.e. the fact that the tail of the α-stable distributed random variable decays as a power function. The α-stable Lévy motion and other stable-based processes have found many applications in various areas of interest, for example in finance [21,22,26], physiology [19], electrical engineering [18], biology [3] and economics [17]. Although the stable processes are ubiquitous in nature, for certain phenomena with stable-kind behaviour (like for example the power law property in the tail) they do not appear to be the appropriate models. Because of that in the literature there are considered many extensions of the classical α-stable Lévy motion.

In the literature one can find different models that are constructed as a superposition of two stochastic processes. The most classical example is the Laplace motion [24]. The Laplace motion, also known as variance gamma process, is stated as a Brownian motion time-changed by the gamma subordinator. This process found many applications especially in finance, namely in option pricing [8,14,16]. The subordinated processes were also studied in other areas of interest, for example in physics [4,25], biology [7] and mechanical systems [6]. In this paper we introduce three models of such subordinated processes. We take the classical α-stable Lévy motion, but we replace the time with another stochastic process. We consider the α_S-stable, the tempered α_T-stable and the gamma subordinator. The introduced models can be understood as the extension of the Laplace motion mentioned above. Instead of the Browian motion, which appears in variance gamma process, we take under consideration its extension, namely the α-stable Levy motion. Moreover, we consider not only gamma, but also other subordinators.

As the result of time-changing we obtain three non-stationary stochastic systems. In this paper we focus on the theoretical properties of these processes to compare which characteristics of the external process retain the same and which ones change due to the subordination. For each model we calculate the characteristic function and the autocodifference of the process. Moreover, we write the

formula for the probability density function and we examine the asymptotic tail behaviour. We also calculate the fractional order moments. To make the application of these processes possible, we propose a simulation procedure and a tail index estimation method.
The paper is organized as follows. In Sect. 2 some notations and definitions concerning the α-stable distribution and the α-stable Lévy motion are presented. We discuss the main properties of the stochastic process which is taken as the external process for the models provided in the paper. In Sect. 3 the main properties of the stochastic processes considered as the subordinators are presented. In Sect. 4 we discuss three models of subordinated processes by presenting the lemmas and proofs of their theoretical properties. Next, in Sect. 5, we propose a method of simulations. In Sect. 6, we present an estimation procedure. Last section contains conclusions.

2 The α-stable Lévy Motion

In this section we present the definition and the properties of the α-stable Lévy motion which we consider as the external process for the subordinated stochastic systems considered in this paper. We also recall the definition of the α-stable distribution (called also stable) and the definition of the α-stable process.

Definition 1. *[23] The random variable X has α-stable distribution if its characteristic function has the following form:*

$$\varphi(t) = E(e^{itX}) = \begin{cases} exp\{i\mu t - |\sigma t|^{\alpha}(1 - i\beta \text{sign(t)} \tan\frac{\pi\alpha}{2})\} & \text{if } \alpha \neq 1, \\ exp\{i\mu t - |\sigma t|^{\alpha}(1 + i\beta \text{sign(t)} \frac{2}{\pi}\log(t)\} & \text{if } \alpha = 1, \end{cases} \tag{1}$$

where $0 < \alpha \leq 2$, $\sigma > 0$, $-1 \leq \beta \leq 1$, $\mu \in R$ are called stability, scale, skewness and location parameters, respectively.

The parameters $\alpha, \sigma, \beta, \mu$ uniquely determine the α-stable distribution. The following expression:

$$X \sim S_{\alpha}(\sigma, \beta, \mu) \tag{2}$$

denotes that X is an α-stable distributed random variable with scale σ, skewness β and location parameter μ. Moreover, if we take $\beta = 0$ and $\mu = 0$ the random variable X is symmetric and for $0 < \alpha < 1$ and $\beta = 1$ the corresponding α-stable distribution is totally right skewed [23]. For $\alpha \neq 2$ the tails of the α-stable distribution behave asymptotically as power functions [23]. The first order moment exists only for $\alpha > 1$ and the second moment is infinite for all $\alpha < 2$ [23]. The probability density function of an α-stable random variable can be expressed analytically only in three special cases, i.e. for Gaussian, Cauchy and Lévy distributions [23].

Definition 2. *[10] A stochastic process $\{X(t)\}$, $t \in T$ is α-stable if all its dimensional distributions*

$$\big(X(t_1),\ X(t_2),\ X(t_3), \ldots,\ X(t_n)\big), \qquad t_1,\ t_2,\ t_3, \ldots,\ t_n \in T, \qquad n \geq 1$$

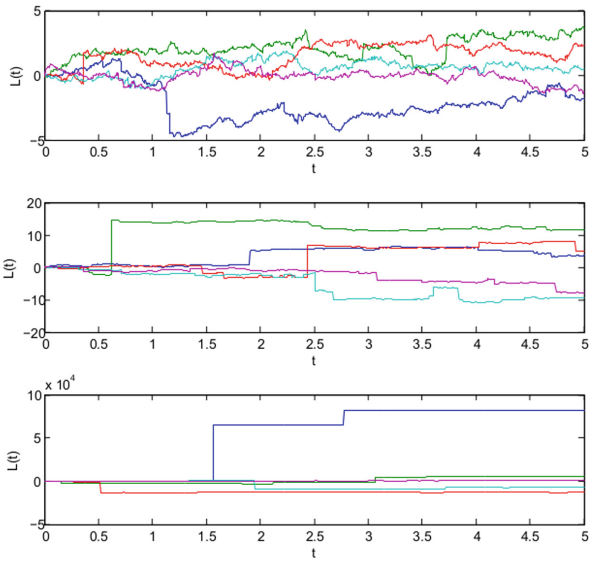


Fig. 1. Sample trajectories of the symmetric α-stable Lévy motion in the case of $\alpha = 1.8$ (top panel), $\alpha = 1.2$ (middle panel) and $\alpha = 0.6$ (bottom panel).

are α-stable. Moreover if all its dimensional distributions are strictly/symmetric α-stable, a stochastic process $\{X(t)\}$ is called to be strictly/symmetric α-stable.

The most famous example of an α-stable process is the α-stable Lévy motion discussed in this section.

Definition 3. *[23] The α-stable Lévy motion $\{L(t)\}$, $t \geq 0$ is a stochastic process satisfying:*

- $L(0) = 0$ *almost surely,*
- $\{L(t)\}$ *has independent increments,*
- $L(t) - L(s) \sim S_\alpha((t-s)^{\frac{1}{\alpha}}, \beta, 0)$.

Properties of the α-stable Lévy motion [10,20,23]:

- for $\alpha = 2$ the 2-stable Lévy motion is simply the Brownian motion multiplied by a constant, i.e. $L(t) \stackrel{d}{=} \sqrt{2}B(t)$;
- for $\beta = 0$ the corresponding α-stable Lévy motion is symmetric;
- the increments of the α-stable Lévy motion are stationary, i.e. for any $0 \leq t < t + h < \infty$ the distribution of $L(t+h) - L(t)$ depends only on h;
- the increments of the α-stable Lévy motion with $\alpha \neq 2$ are distributed according to a heavy-tailed distribution, it means that for each t the following property is satisfied [23]:

$$\begin{aligned} \lim_{x\to\infty} P(L(t) > x)\, x^\alpha &= \frac{1}{2} C_\alpha (1+\beta) t, \\ \lim_{x\to\infty} P(L(t) < -x)\, x^\alpha &= \frac{1}{2} C_\alpha (1-\beta) t, \end{aligned} \tag{3}$$

where $C_\alpha = (1-\alpha)/(\Gamma(2-\alpha)\cos(\pi\alpha/2))$ for $\alpha \neq 1$ and $C_\alpha = \frac{2}{\pi}$ for $\alpha = 1$. This property impacts on the trajectories of the α-stable Lévy motion: the smaller values the parameter α takes, the bigger "jumps" of the trajectories we can observe [10]. Sample paths of the symmetric α-stable Lévy motion for different values of the parameters α are presented in Fig. 1.

– the α-stable Lévy motion is H-sssi, i.e. it is a self-similar process with index H and stationary increments; the self-similarity index H is equal to $1/\alpha \in [1/2, \infty)$ [23], that is for any $c > 0$ the finite dimensional distributions of the processes $\{L(ct)\}$ and $\{c^{1/\alpha}L(t)\}$ are the same.
– from the formula for the fractional moments of the α-stable random variables, given in [23], follows the expression for the fractional moments of the α-stable Lévy motion. Let $\alpha \neq 2$ and $\beta = 0$ in the case $\alpha = 1$. Then the moment of order $0 < p < \alpha$ for each t is given by the following formula [23]:

$$(E|L(t)|^p)^{\frac{1}{p}} = c_{\alpha,\beta}(p)\, t^{\frac{1}{\alpha}}, \tag{4}$$

where the constant $c_{\alpha,\beta}(p)$ equals $(E|L(1)|^p)^{\frac{1}{p}}$ and $L(1) \sim S_\alpha(1, \beta, 0)$. It follows from the fact that $L(t) \stackrel{d}{=} t^{\frac{1}{\alpha}}L(1)$ and

$$L(t) \sim S_\alpha(t^{\frac{1}{\alpha}}, \beta, 0). \tag{5}$$

The constant $c_{\alpha,\beta}(p)$ is given by [23]:

$$\begin{aligned}(c_{\alpha,\beta}(p))^p \;=\; & \frac{2^{p-1}\Gamma(1-\frac{p}{\alpha})}{p\int_0^\infty u^{-p-1}\sin^2(u)du} \\ & \Big(1+\beta^2\, tan^2\Big(\frac{\alpha\pi}{2}\Big)\Big)^{\frac{p}{2\alpha}} \cos\Big(\frac{p}{\alpha}\, arctan\Big(\beta\, tan\Big(\frac{\alpha\pi}{2}\Big)\Big)\Big).\end{aligned} \tag{6}$$

– the α-stable Lévy motion is an example of a stochastic process with diverging variance. It follows from the fact that the second moment is infinite for all $\alpha < 2$ and the first moment is infinite for all $\alpha < 1$. Because of this property, to measure the interdependence we cannot apply the classical measures like correlation or covariance which are based on the second moment [23,27]. Instead of them we apply the measure called codifference (CD) which is an alternative method of measuring the dependence between the α-stable random variables because it is well-defined for all $0 < \alpha \leq 2$. The codifference of a stochastic process $\{X(t)\}, t \in T$ is called the autocodifference and it is given by the following expression [20,23,27]:

$$\begin{aligned}&CD(X(t), X(s)) \\ &\quad = \log\left(E(e^{iX(t)})\right) + \log\left(E(e^{-iX(s)})\right) - \log\left(E(e^{i(X(t)-X(s))})\right),\end{aligned} \tag{7}$$

where $s, t \in T$. The significant advantage is the fact that the codifference is defined based on the characteristic function which always exists and completely defines the distribution. For the α-stable Lévy motion it takes the form [27]:

$$CD(L(t), L(s)) = -2\min(\mathrm{s}, \mathrm{t}). \tag{8}$$

3 Different Types of Subordinators

In this section we introduce three processes which we consider as the subordinators for the process $\{L(t)\}$ presented in Sect. 2. We explain the main properties of the α-stable subordinator $\{S_\alpha(\tau)\}$, the tempered α-stable subordinator $\{T_{\alpha,\lambda}(\tau)\}$ and the gamma subordinator $\{G_{k,\theta}(\tau)\}$.

Since a subordinator is a stochastic process replacing time in an external process, it has to be increasing almost surely. A subordinator has to have cadlag sample paths (right-continuous with left limits) [9,23]. We assume for an infinitely divisible subordinator $\{U(\tau)\}$, $\tau \geq 0$ that the Laplace transform is given by [9]:

$$Ee^{-zU(\tau)} = e^{-\tau\phi(z)}, \tag{9}$$

where $\phi(z)$ is called the Lévy exponent and it is expressed as:

$$\phi(z) = \int_0^\infty (1 - e^{-zx})\ v(dx) \tag{10}$$

and $v(dx)$ is an appropriate Lévy measure [9].

3.1 The α-stable Subordinator

We introduce the α-stable subordinator $\{S_\alpha(\tau)\}$, $\tau \geq 0$ which is a Lévy process with α-stable increments. The characteristic function and the properties of the α-stable distribution are presented in Sect. 2 where the α-stable Lévy motion is considered. Since a subordinator has to be an increasing process, we need to take $0 < \alpha < 1$, $\beta = 1$. We assume also $\mu = 0$ and $\sigma = \cos(\pi\alpha/2)^{1/\alpha}$. The Lévy measure for $\{S_\alpha(\tau)\}$ is defined as [9]:

$$v(dx) = x^{-(1+\alpha)}\mathbf{1}_{x>0}\ dx. \tag{11}$$

The Lévy exponent in Formula (9) is expressed in the following form [9]:

$$\phi(z) = z^\alpha, \tag{12}$$

what gives the corresponding Laplace transform:

$$E(e^{-zS_\alpha(\tau)}) = e^{-\tau z^\alpha}. \tag{13}$$

Sample paths of the α-stable subordinator for exemplary values of the parameter α are presented in Fig. 2.

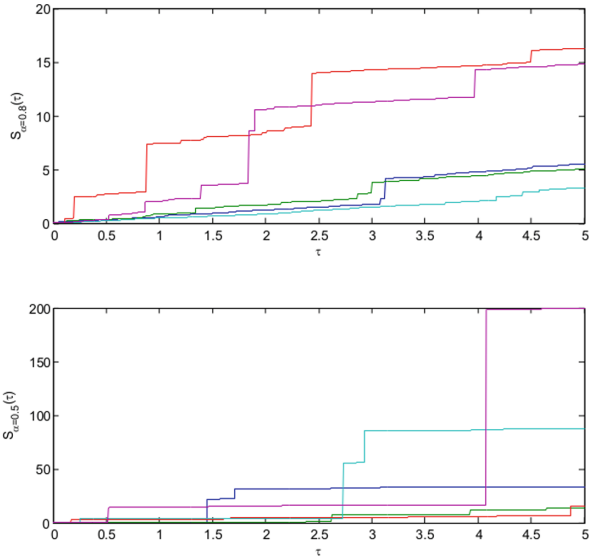


Fig. 2. Sample trajectories of the α-stable subordinator in the case of $\alpha = 0.8$ (top panel) and $\alpha = 0.5$ (bottom panel).

3.2 The Tempered α-stable Subordinator

A positive random variable $T_{\alpha,\lambda}$ with a stability parameter $0 < \alpha < 1$ and a tempering parameter $\lambda > 0$ has the tempered α-stable distribution if it is defined via the following Laplace transform [5,9]:

$$Ee^{-uT_{\alpha,\lambda}} = e^{-((u+\lambda)^{\alpha}-\lambda^{\alpha})}. \tag{14}$$

Let us notice that for $\lambda = 0$ we obtain the Laplace transform of an α-stable random variable, more precisely we obtain a random variable with one-sided α-stable distribution [5,9]. For such definition of the tempered α-stable distribution the probability density function of a random variable $T_{\alpha,\lambda}$ takes the following form [5,9]:

$$f^{T_{\alpha,\lambda}}(x) = Ce^{-\lambda x} f^{S_{\alpha}}(x), \tag{15}$$

where the quantity C is a normalizing constant and $f^{S_{\alpha}}(x)$ is the probability density function corresponding to the one-sided α-stable distribution [5,9]. Furthermore, for the tempered α-stable distribution moments of all orders are finite [9].

The tempered α-stable subordinator $\{T_{\alpha,\lambda}(\tau)\}$, $\tau \geq 0$ is an increasing Lévy process which increments have the tempered α-stable distribution, i.e. the corresponding Lévy measure is given by [9]:

$$v(dx) = e^{-\lambda x} x^{-1-\alpha} \mathbf{1}_{x>0}\, dx. \tag{16}$$

For this subordinator the function $\phi(z)$ appearing in Formula (9) takes the form [9]:

$$\phi(z) = (z+\lambda)^{\alpha} - \lambda^{\alpha} \tag{17}$$

what leads to the following form of the Laplace transform:

$$E(e^{-zT_{\alpha,\lambda}(\tau)}) = e^{-\tau((z+\lambda)^{\alpha}-\lambda^{\alpha})}. \tag{18}$$

Let us notice that for $\lambda = 0$ the considered subordinator becomes an α-stable one. Moreover, the moment generating function of the tempered α-stable subordinator is given by:

$$M_{T_{\alpha,\lambda}(\tau)}(z) = E(e^{zT_{\alpha,\lambda}(\tau)}) = e^{-\tau((\lambda-z)^{\alpha}-\lambda^{\alpha})}. \tag{19}$$

Next property of this subordinator is the fact that the probability density function of $\{T_{\alpha,\lambda}(\tau)\}$ can be written in the following way [9]:

$$f^{T_{\alpha,\lambda}}(x,\tau) = \; e^{-\lambda x+\lambda^{\alpha}\tau} \; f^{S_{\alpha}}(x,\tau), \tag{20}$$

where $f^{S_{\alpha}}(x,\tau)$ represents the probability density function of a totally skewed α-stable Lévy process.

In [12] it was proven that for $q > 0$ the q-th order moment of the tempered α-stable subordinator $\{T_{\alpha,\lambda}(\tau)\}$ behaves asymptotically in the following way:

$$E(T_{\alpha,\lambda}(\tau))^q \sim (\alpha\lambda^{\alpha-1}\tau)^q, \tag{21}$$

as $\tau \to \infty$. Sample paths of the tempered α-stable subordinator for exemplary values of the parameters are presented in Fig. 3.

3.3 The Gamma Subordinator

A gamma-distributed random variable $G_{k,\theta}$ with a shape parameter $k > 0$ and a scale parameter $\theta > 0$ has the probability density function denoted as follows [9]:

$$f^{G_{k,\theta}}(x) = \frac{1}{\Gamma(k)\theta^k} x^{k-1} e^{-\frac{x}{\theta}}, \tag{22}$$

where $x \geq 0$ and $\Gamma(k)$ given by:

$$\Gamma(k) = \int_0^{\infty} t^{k-1} e^{-t} dt \tag{23}$$

is the Gamma function. The Laplace transform of the gamma distribution takes the following form [9]:

$$E(e^{-uG_{k,\theta}}) = (1+\theta u)^{-k}. \tag{24}$$

When the parameter k takes only integer values, then this distribution is called an Erlang distribution and a random variable $G_{k,\theta}$ is a sum of k independent identically distributed random variables and each of them is exponentially

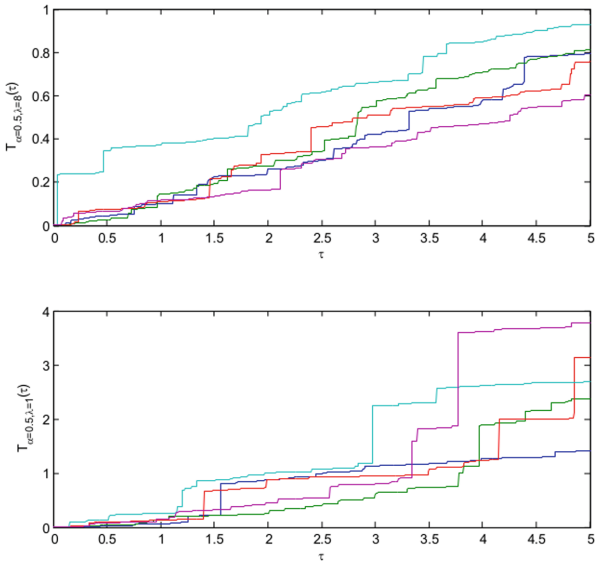


Fig. 3. Sample trajectories of the tempered α-stable subordinator in the case of $\alpha = 0.5$, $\lambda = 8$ (top panel) and $\alpha = 0.5$, $\lambda = 1$ (bottom panel).

distributed with mean equal to θ. Furthermore, the gamma distribution is an example of an infinitely divisible distribution, i.e. for n independent random variables X_i having gamma distribution with parameters k_i and θ a random variable $\sum_{i=1}^{n} X_i$ is also gamma-distributed with parameters $\sum_{i=1}^{n} k_i$ and θ [9].

The gamma subordinator $\{G_{k,\theta}(\tau)\}$, $\tau \geq 0$ is the Lévy process with independent gamma-distributed increments. The trajectories of the process $\{G_{k,\theta}(\tau)\}$ are strictly increasing with jumps [13]. In this case the Lévy measure is defined as [9,13]:

$$v(dx) = k\frac{e^{-\theta x}}{x}\mathbf{1}_{x>0}\, dx \tag{25}$$

and the Lévy exponent $\phi(z)$ is given by [9,13]:

$$\phi(z) = k\log(1+\theta z). \tag{26}$$

The expressions (25) and (26) lead to the following Laplace transform:

$$E(e^{-zG_{k,\theta}(\tau)}) = (1+\theta z)^{-k\tau}. \tag{27}$$

Moreover, the probability density function of the gamma subordinator $\{G_{k,\theta}(\tau)\}$ is given in an analytical form [9]:

$$f^{G_{k,\theta}}(x,\tau) = \frac{1}{\Gamma(k\tau)\theta^{k\tau}}\, x^{k\tau-1}e^{-\frac{x}{\theta}}. \tag{28}$$

It was proved in [13] that for $q > 0$ the q-th order moment of the gamma subordinator $\{G_{k,\theta}(\tau)\}_{\tau \geq 0}$ behaves in the following way:

$$E(G_{k,\theta}(\tau))^q = \theta^q \frac{\Gamma(q + k\tau)}{\Gamma(k\tau)} \sim (k\theta\tau)^q, \tag{29}$$

as $\tau \to \infty$. Sample paths of the gamma subordinator for exemplary values of the parameters are presented in Fig. 4.

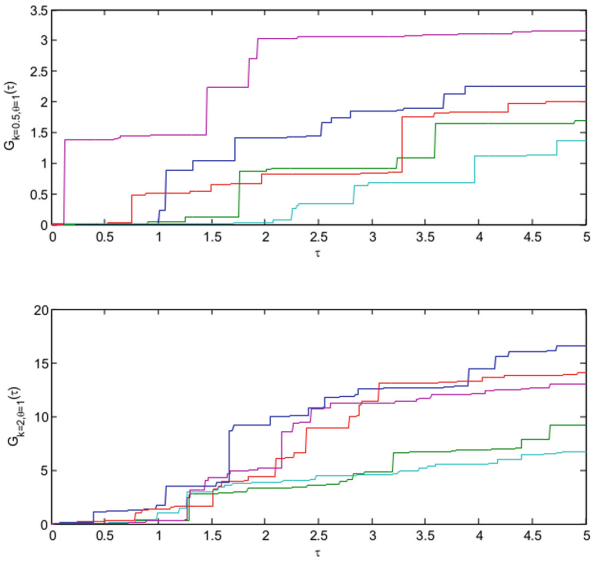


Fig. 4. Sample trajectories of the gamma subordinator in the case of $k = 0.5$, $\theta = 1$ (top panel) and $k = 2$, $\theta = 1$ (bottom panel).

4 Subordinated α-stable Lévy Motion

In this section we introduce the time-changed α-stable Lévy motion for which we replace the time with another process presented in Sect. 3. The subordinated process takes the following form:

$$X(t) := L(Y(t)), \tag{30}$$

where $L(\cdot)$ denotes the α-stable Lévy motion and $Y(\cdot)$ is its subordinator. In the paper we consider the special case of the α-stable Lévy motion, namely the symmetric process, i.e. we take the skewness parameter $\beta = 0$. The characteristic function of $L(t)$ for the symmetric case is given by:

$$\phi_{L(t)}(u) = E\left(e^{iuL(t)}\right) = e^{-t|u|^\alpha}, \tag{31}$$

where $0 < \alpha < 2$, $t > 0$ and $u \in R$. We also assume that the external and internal processes are independent. In our consideration we take into account three subordinators: α_S-stable, tempered α_T-stable and gamma process. Since all the subordinators and the external process are Lévy processes, $\{X(t)\}$ is also Lévy (i.e. has independent and stationary increments).

Lemma 4.1. The subordinated α-stable Lévy motion $\{X(t)\}$ has stationary increments, i.e. for any $0 \leq t < t + h < \infty$ the distribution of an increment $X(t + h) - X(t)$ depends only on h.

Proof. Let us take t and h such that $0 \leq t < t + h < \infty$. Then:

$$\begin{aligned} X(t+h) - X(t) &= L(Y(t+h)) - L(Y(t)) \stackrel{d}{=} \\ (*) &\stackrel{d}{=} L(Y(t+h) - Y(h)) \stackrel{d}{=} \\ (**) &\stackrel{d}{=} L(Y(t+h-t)) = L(Y(h)) = X(h) \end{aligned} \tag{32}$$

In $(*)$ we apply the fact that the increments of the α-stable Lévy motion $\{L(t)\}$ are stationary and in $(**)$ we use the property about the stationarity of the increments of any subordinator $\{Y(t)\}$. □

4.1 The α-stable Lévy Motion Time-Changed by the α_S-stable Subordinator

The α-stable Lévy motion time-changed by the α_S-stable subordinator is defined in the following way:

$$X(t) := L(S_{\alpha_S}(t)), \tag{33}$$

where $L(\cdot)$ denotes the symmetric α-stable Lévy motion presented in Sect. 2 with the skewness parameter $\beta = 0$ and $S_{\alpha_S}(\cdot)$ is the α_S-stable subordinator presented in Sect. 3.1 with the stability parameter $0 < \alpha_S < 1$ and the skewness parameter $\beta_S = 1$.

Lemma 4.2. For the α-stable Lévy motion time-changed by the α_S-stable subordinator, the characteristic function of $X(t)$ is given by:

$$\phi_{X(t)}(u) = e^{-t|u|^{\alpha \cdot \alpha_S}} \quad \text{for} \quad u \in R. \tag{34}$$

Proof. Applying the standard conditioning procedure and using Formula (31) for the characteristic function of $L(t)$ and Formula (13) for the Laplace transform of $S_{\alpha_S}(t)$, we obtain:

$$\begin{aligned} \phi_{X(t)}(u) &= E\big(e^{iuX(t)}\big) = E\big(e^{iuL(S_{\alpha_S}(t))}\big) \\ &= E\big(E\big(e^{iuL(S_{\alpha_S}(t))}|S_{\alpha_S}(t)\big)\big) \\ &= E\big(e^{-S_{\alpha_S}(t)|u|^{\alpha}}\big) = e^{-t|u|^{\alpha \cdot \alpha_S}}. \end{aligned} \tag{35}$$

□

Lemma 4.3. The codifference of the α-stable Lévy motion time-changed by the α_S-stable subordinator is given by:

$$CD(X(t), X(s)) = -2\min(s, t). \tag{36}$$

Proof. We use the general formula for the codifference of a stochastic process given in (7). From Formula (34) for the characteristic function of $X(t)$ we have:

$$E\left(e^{iX(t)}\right) = e^{-t|1|^{\alpha\cdot\alpha_S}} = e^{-t}, \tag{37}$$

$$E\left(e^{-iX(s)}\right) = e^{-s|-1|^{\alpha\cdot\alpha_S}} = e^{-s}, \tag{38}$$

$$E\left(e^{i(X(t)-X(s))}\right) = E\left(e^{i(X(t-s))}\right) = e^{-(t-s)|1|^{\alpha\cdot\alpha_S}} = e^{-(t-s)}. \tag{39}$$

Finally, we get:

$$\begin{aligned} CD(X(t), X(s)) &= \log\left(E(e^{iX(t)})\right) + \log\left(E(e^{-iX(s)})\right) - \log\left(E(e^{i(X(t)-X(s)})\right) \\ &= \log(e^{-t}) + \log(e^{-s}) - \log(e^{-(t-s)}) \\ &= -t - s + t - s = -2s = -2\min(s, t). \end{aligned} \tag{40}$$

□

From the form of the characteristic function follows the fact that the random variable $X(t)$ is stable distributed, namely:

$$X(t) \sim S_{\alpha_S\cdot\alpha}(t^{\frac{1}{\alpha_S\cdot\alpha}}, 0, 0). \tag{41}$$

Moreover, from Formula (34) we can conclude that $\{X(t)\}$ is a self-similar process with $H = (\alpha_S \cdot \alpha)^{-1}$. In general for the stochastic process $\{X(t)\}$ defined in (33) the probability density function $p(x, t)$ takes the following form [11,12]:

$$p(x, t) = \int_0^\infty g(x, r) f(r, t) dr, \tag{42}$$

where $x \in R$ and $t \geq 0$. The function $g(x, t)$ denotes the pdf of the symmetric α-stable Lévy motion and $f(r, t)$ denotes the pdf of the α_S-stable subordinator.

Lemma 4.4. For the α-stable Lévy motion time-changed by the α_S-stable subordinator the following formula holds for each t:

$$P(X(t) > x) \sim \frac{1}{2} C_{\alpha_S\cdot\alpha} x^{-\alpha_S\cdot\alpha} t \quad \text{as} \quad x \to \infty, \tag{43}$$

where the constant $C_{\alpha_S\cdot\alpha}$ is given by:

$$C_{\alpha_S\cdot\alpha} = \begin{cases} \frac{1-\alpha_S\cdot\alpha}{\Gamma(2-\alpha_S\cdot\alpha)\cos(\pi\alpha_S\cdot\alpha/2)} & \text{for} \quad \alpha_S \cdot \alpha \neq 1, \\ \frac{2}{\pi} & \text{for} \quad \alpha_S \cdot \alpha = 1. \end{cases} \tag{44}$$

Proof. We use the fact that $X(t)$ has $\alpha_S \cdot \alpha$-stable distribution and from the properties of the stable random variables we can write that for each t the following limit holds [23]:

$$\lim_{x\to\infty} P(|X(t)| > x)x^{\alpha_S\cdot\alpha} = \frac{1}{2}C_{\alpha_S\cdot\alpha}t, \tag{45}$$

where the constant $C_{\alpha_S\cdot\alpha}$ given by (44) is specified in [23]. Thus, we can write that

$$P(X(t) > x) \sim \frac{1}{2}C_{\alpha_S\cdot\alpha}x^{-\alpha_S\cdot\alpha}t \quad \text{as} \quad x \to \infty. \tag{46}$$

□

Lemma 4.5. For $0 < p < \alpha_S \cdot \alpha < 2$, the fractional order moments of the α-stable Lévy motion time-changed by the α_S-stable subordinator for each t are given by:

$$E|X(t)|^p = (c_{\alpha_S\cdot\alpha,0}(p))^p t^{\frac{p}{\alpha_S\cdot\alpha}}, \tag{47}$$

where the constant $(c_{\alpha,0}(p))^p$ has the form:

$$(c_{\alpha,0}(p))^p = \frac{2^{p-1}\Gamma(1-\frac{p}{\alpha_S\cdot\alpha})}{p\int_0^\infty u^{-p-1}\sin^2(u)du}. \tag{48}$$

Proof. We use the fact that $X(t)$ has $\alpha_S \cdot \alpha$-stable distribution and the general formula for the fractional order moments of the stable random variables presented in [23]. □

4.2 The α-stable Lévy Motion Time-Changed by the Tempered α_T-stable Subordinator

The α-stable Lévy motion time-changed by the tempered α_T-stable subordinator is defined in the following way:

$$X(t) := L(T_{\alpha_T,\lambda}(t)), \tag{49}$$

where $L(\cdot)$ denotes the symmetric α-stable Lévy motion presented in Sect. 2 with the skewness parameter $\beta = 0$ and $T_{\alpha_T,\lambda}(\cdot)$ is the tempered α_T-stable subordinator presented in Sect. 3.2 with the stability parameter $0 < \alpha_T < 1$ and the tempering parameter $\lambda > 0$.

Lemma 4.6. For the α-stable Lévy motion time-changed by the tempered α_T-stable subordinator, the characteristic function of $X(t)$ is given by:

$$\phi_{X(t)}(u) = e^{-t((|u|^\alpha+\lambda)^{\alpha_T}-\lambda^{\alpha_T})} \quad \text{for} \quad u \in R. \tag{50}$$

Proof. Applying the standard conditioning procedure and Formula (31) for the characteristic function of $L(t)$ and Formula (18) for the Laplace transform of $T_{\alpha_T,\lambda}(t)$, we have:

$$\begin{aligned}\phi_{X(t)}(u) &= E\left(e^{iuX(t)}\right) = E\left(e^{iuL(T_{\alpha_T,\lambda}(t))}\right) \\ &= E\left(E\left(e^{iuL(T_{\alpha_T,\lambda}(t))}|T_{\alpha_T,\lambda}(t)\right)\right) \\ &= E\left(e^{-T_{\alpha_T,\lambda}(t)|u|^{\alpha}}\right) = e^{-t((|u|^{\alpha}+\lambda)^{\alpha_T}-\lambda^{\alpha_T})}.\end{aligned} \tag{51}$$

□

Lemma 4.7. The codifference of the α-stable Lévy motion time-changed by the tempered α_T-stable subordinator is given by:

$$CD(X(t), X(s)) = -2\mathrm{min(s,t)}((1+\lambda)^{\alpha_T} - \lambda^{\alpha_T}). \tag{52}$$

Proof. We use the general formula for the codifference of a stochastic process presented in (7). From Formula (50) for the characteristic function of $X(t)$ we have:

$$E\left(e^{iX(t)}\right) = e^{-t((|1|^{\alpha}+\lambda)^{\alpha_T}-\lambda^{\alpha_T})} = e^{-t((1+\lambda)^{\alpha_T}-\lambda^{\alpha_T})}, \tag{53}$$

$$E\left(e^{-iX(s)}\right) = e^{-s((|-1|^{\alpha}+\lambda)^{\alpha_T}-\lambda^{\alpha_T})} = e^{-s((1+\lambda)^{\alpha_T}-\lambda^{\alpha_T})}, \tag{54}$$

$$\begin{aligned}E\left(e^{i(X(t)-X(s))}\right) &= E\left(e^{i(X(t-s))}\right) = e^{-(t-s)((|1|^{\alpha}+\lambda)^{\alpha_T}-\lambda^{\alpha_T})} \\ &= e^{-(t-s)((1+\lambda)^{\alpha_T}-\lambda^{\alpha_T})}.\end{aligned} \tag{55}$$

Finally, we obtain the following expression:

$$\begin{aligned}CD(X(t), X(s)) &= \log\left(E(e^{iX(t)})\right) + \log\left(E(e^{-iX(s)})\right) - \log\left(E(e^{i(X(t)-X(s)})\right) \\ &= \log(e^{-t((1+\lambda)^{\alpha_T}-\lambda^{\alpha_T})}) + \log(e^{-s((1+\lambda)^{\alpha_T}-\lambda^{\alpha_T})}) \\ &- \log(e^{-(t-s)((1+\lambda)^{\alpha_T}-\lambda^{\alpha_T})}) \\ &= -t((1+\lambda)^{\alpha_T} - \lambda^{\alpha_T}) - s((1+\lambda)^{\alpha_T} - \lambda^{\alpha_T}) \\ &+ (t-s)((1+\lambda)^{\alpha_T} - \lambda^{\alpha_T}) = -2s((1+\lambda)^{\alpha_T} - \lambda^{\alpha_T}) \\ &= -2\mathrm{min(s,t)}((1+\lambda)^{\alpha_T} - \lambda^{\alpha_T}).\end{aligned} \tag{56}$$

□

Moreover, from the form of the characteristic function of $\{X(t)\}$ follows that the stochastic process defined in this way is not a self-similar process. For the stochastic process $\{X(t)\}$ defined in (49) the probability density function $p(x,t)$ has the form presented in (42). In this formula, the pdf of the symmetric α-stable Lévy motion is denoted by $g(x,t)$ and the pdf of the tempered α_T-stable subordinator is denoted by $f(r,t)$. From Formula (20) we can write $f(r,t)$ using the probability density function of a totally skewed α_T-stable process. Thus, the pdf $p(x,t)$ can be written as:

$$p(x,t) = \int_0^{\infty} g(x,r)e^{-\lambda r+\lambda^{\alpha}t} f^{S_{\alpha_T}}(r,t)dr, \tag{57}$$

where $f^{S_{\alpha_T}}(r,t)$ represents the pdf of a totally skewed α_T-stable process.

Lemma 4.8. For the α-stable Lévy motion time-changed by the tempered α_T-stable subordinator the following formula holds for each t:

$$P(X(t) > x) \sim \frac{1}{2} C_\alpha x^{-\alpha} (t\alpha_T \lambda^{\alpha_T - 1}) \quad \text{as} \quad x \to \infty, \tag{58}$$

where the constant C_α is given by:

$$C_\alpha = \begin{cases} \frac{1-\alpha}{\Gamma(2-\alpha)\cos(\pi\alpha/2)} & \text{for} \quad \alpha \neq 1, \\ \frac{2}{\pi} & \text{for} \quad \alpha = 1. \end{cases} \tag{59}$$

Proof. Applying the standard conditioning procedure and Formula (3) describing the tail of the α-stable Lévy motion, we obtain:

$$\begin{aligned} P(X(t) > x) &= E(P(L(T_{\alpha_T,\lambda}(t)) > x | T_{\alpha_T,\lambda}(t))) \\ &\sim E\Big(\frac{1}{2} C_\alpha T_{\alpha_T,\lambda}(t) x^{-\alpha}\Big) \\ &= \frac{1}{2} C_\alpha x^{-\alpha} E(T_{\alpha_T,\lambda}(t)) = \\ (*) \ &= \frac{1}{2} C_\alpha x^{-\alpha} (t\alpha_T \lambda^{\alpha_T - 1}), \end{aligned} \tag{60}$$

as $x \to \infty$, where the constant C_α is given in (3). In $(*)$ we use the formula for the expected value of $T_{\alpha_T,\lambda}(t)$. □

Lemma 4.9. For $0 < p < \alpha < 2$, the fractional order moments of the α-stable Lévy motion time-changed by the tempered α_T-stable subordinator are given by:

$$E(|X(t)|^p) \sim (t\alpha_T \lambda^{\alpha_T - 1})^{p/\alpha} (c_{\alpha,0}(p))^p \quad \text{as} \quad t \to \infty, \tag{61}$$

where the constant $(c_{\alpha,0}(p))^p$ has the following form:

$$(c_{\alpha,0}(p))^p = \frac{2^{p-1}\Gamma(1 - \frac{p}{\alpha})}{p \int_0^\infty u^{-p-1} \sin^2(u) du}. \tag{62}$$

Proof. Using the self-similarity of the α-stable Lévy motion and the expression (4) for the fractional order moments of the α-stable Lévy motion, for $0 < p < \alpha < 2$ we obtain:

$$\begin{aligned} E(|X(t)|^p) &= E(|L(T_{\alpha_T,\lambda}(t))|^p) \\ &= E(|T_{\alpha_T,\lambda}(t)^{1/\alpha} L(1)|^p) \\ &= E(T_{\alpha_T,\lambda}(t)^{p/\alpha} |L(1)|^p) \\ &= E(T_{\alpha_T,\lambda}(t)^{p/\alpha}) E(|L(1)|^p) \\ (*) &\sim (t\alpha_T \lambda^{\alpha_T - 1})^{p/\alpha} (c_{\alpha,0}(p))^p \quad \text{as} \quad t \to \infty. \end{aligned} \tag{63}$$

In $(*)$ we apply Formula (21) for the asymptotic behaviour of the moments of the tempered α_T-stable subordinator. Moreover, the constant $(c_{\alpha,\beta}(p))^p$ is specified in (6) and for $\beta = 0$ takes the form presented in (63). □

4.3 The α-stable Lévy Motion Time-Changed by the Gamma Subordinator

The α-stable Lévy motion time-changed by the gamma subordinator is defined in the following way:

$$X(t) := L(G_{k,\theta}(t)), \tag{64}$$

where $L(\cdot)$ denotes the symmetric α-stable Lévy motion presented in Sect. 2 with the skewness parameter $\beta = 0$ and $G_{k,\theta}(\cdot)$ is the gamma subordinator presented in Sect. 3.3 with the shape parameter $k > 0$ and the scale parameter $\theta > 0$.

Lemma 4.10. For the α-stable Lévy motion time-changed by the gamma subordinator, the characteristic function of $X(t)$ is given by:

$$\phi_{X(t)}(u) = \Big(\frac{1}{1+\theta|u|^{\alpha}}\Big)^{kt} \quad \text{for} \quad u \in R. \tag{65}$$

Proof. Using the standard conditioning procedure and applying Formula (31) for the characteristic function of $L(t)$ and Formula (27) for the Laplace transform of $G_{k,\theta}(t)$, we obtain:

$$\begin{aligned}\phi_{X(t)}(u) &= E\big(e^{iuX(t)}\big) = E\big(e^{iuL(G_{k,\theta}(t))}\big)\\ &= E\big(E\big(e^{iuL(G_{k,\theta}(t))}|G_{k,\theta}(t)\big)\big)\\ &= E\big(e^{-G_{k,\theta}(t)|u|^{\alpha}}\big)\\ &= \Big(\frac{1}{1+\theta|u|^{\alpha}}\Big)^{kt}.\end{aligned} \tag{66}$$

□

Lemma 4.11. The codifference of the α-stable Lévy motion time-changed by the gamma subordinator is given by:

$$CD(X(t), X(s)) = -2\mathrm{min(s,t)k}\log\big(1+\theta\big). \tag{67}$$

Proof. We use the general formula for the codifference of a stochastic process presented in (7). From Formula (66) for the characteristic function of $X(t)$ we have:

$$E\big(e^{iX(t)}\big) = \Big(\frac{1}{1+\theta|1|^{\alpha}}\Big)^{kt} = \Big(\frac{1}{1+\theta}\Big)^{kt}, \tag{68}$$

$$E\big(e^{-iX(s)}\big) = \Big(\frac{1}{1+\theta|-1|^{\alpha}}\Big)^{ks} = \Big(\frac{1}{1+\theta}\Big)^{ks}, \tag{69}$$

$$E\big(e^{i(X(t)-X(s))}\big) = E\big(e^{i(X(t-s))}\big) = \Big(\frac{1}{1+\theta|1|^{\alpha}}\Big)^{k(t-s)} = \Big(\frac{1}{1+\theta}\Big)^{k(t-s)}. \tag{70}$$

Finally, we get the formula:

$$\begin{aligned} CD(X(t),X(s)) &= \log\left(E(e^{iX(t)})\right) + \log\left(E(e^{-iX(s)})\right) - \log\left(E(e^{i(X(t)-X(s)})\right) \\ &= \log\left(\frac{1}{1+\theta}\right)^{kt} + \log\left(\frac{1}{1+\theta}\right)^{ks} - \log\left(\frac{1}{1+\theta}\right)^{k(t-s)} \\ &= kt\log\left(\frac{1}{1+\theta}\right) + ks\log\left(\frac{1}{1+\theta}\right) - k(t-s)\log\left(\frac{1}{1+\theta}\right) \\ &= 2ks\log\left(\frac{1}{1+\theta}\right) = -2ks\log\left(1+\theta\right) \\ &= -2\mathrm{min(s,t)k}\log\left(1+\theta\right). \end{aligned} \tag{71}$$

□

Moreover, we conclude from the form of the characteristic function that the stochastic process {X(t)} is not a self-similar process. The probability density function $p(x,t)$ of the stochastic process $\{X(t)\}$ has the form given by Formula (42), where the function $g(x,t)$ denotes the pdf of the symmetric α-stable Lévy motion and $f(r,t)$ denotes the pdf of the gamma subordinator. Since $f(r,t)$ is given by Formula (28) we can write the integral in (42) as:

$$p(x,t) = \int_0^\infty \frac{g(x,r)}{\Gamma(kt)\theta^{kt}}\, r^{kt-1}e^{-\frac{r}{\theta}}dr. \tag{72}$$

Lemma 4.12. The α-stable Lévy motion time-changed by the gamma subordinator is in domain of attraction of a stable law, i.e. the following limit holds:

$$\lim_{t\to\infty}\left(\frac{1}{k\theta}\right)^{\frac{1}{\alpha}}\frac{X(t)}{t^{\frac{1}{\alpha}}} \xrightarrow{d} L(1). \tag{73}$$

Proof. Using the self-similarity of the Lévy motion, for $t\to\infty$, we have:

$$\begin{aligned} \left(\frac{1}{k\theta}\right)^{\frac{1}{\alpha}}\frac{X(t)}{t^{\frac{1}{\alpha}}} &= \left(\frac{1}{k\theta}\right)^{\frac{1}{\alpha}}\frac{L(G_{k,\theta}(t))}{t^{\frac{1}{\alpha}}} \\ &\overset{d}{\sim} \left(\frac{1}{k\theta}\right)^{\frac{1}{\alpha}}\frac{G_{k,\theta}(t)^{\frac{1}{\alpha}}L(1)}{t^{\frac{1}{\alpha}}} \\ &= \left(\frac{1}{k\theta}\right)^{\frac{1}{\alpha}}\left(\frac{G_{k,\theta}(t)}{t}\right)^{\frac{1}{\alpha}}L(1) \\ (*) &\xrightarrow{d} \left(\frac{1}{k\theta}\right)^{\frac{1}{\alpha}}\left(E(G_{k,\theta}(1))\right)^{\frac{1}{\alpha}}L(1) = L(1), \end{aligned} \tag{74}$$

where the convergence $(*)$ results form the following property of a subordinator $\{Y(t)\}$ [11]:

$$\lim_{t\to\infty}\frac{Y(t)}{t} \xrightarrow{a.s.} E\big(Y(1)\big). \tag{75}$$

We also use the fact that $E(G_{k,\theta}(1)) = k\theta$ and the result that if $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{d} c$ for some constant c, then $X_n Y_n \xrightarrow{d} cX$ s $n \to \infty$ [11]. Summarizing, we obtain:

$$\lim_{t\to\infty} \left(\frac{1}{k\theta}\right)^{\frac{1}{\alpha}} \frac{X(t)}{t^{\frac{1}{\alpha}}} \xrightarrow{d} L(1). \tag{76}$$

□

Lemma 4.13. For the α-stable Lévy motion time-changed by the gamma subordinator the following formula holds for $t \to \infty$:

$$P(X(t) > x) \sim \frac{1}{2} C_\alpha x^{-\alpha} (k\theta t) \quad \text{as} \quad x \to \infty, \tag{77}$$

where the constant C_α is given by:

$$C_\alpha = \begin{cases} \frac{1-\alpha}{\Gamma(2-\alpha)\cos(\pi\alpha/2)} & \text{for} \quad \alpha \neq 1, \\ \frac{2}{\pi} & \text{for} \quad \alpha = 1. \end{cases} \tag{78}$$

Proof. Using the standard conditioning procedure and Formula (3) describing the tail of the α-stable Lévy motion, we get that for $x \to \infty$:

$$\begin{aligned} P(X(t) > x) &= E(P(L(G_{k,\theta}(t)) > x | G_{k,\theta}(t))) \\ &\sim E\left(\frac{1}{2} C_\alpha G_{k,\theta}(t) x^{-\alpha}\right) \\ &= \frac{1}{2} C_\alpha x^{-\alpha} E(G_{k,\theta}(t)) \\ &= \frac{1}{2} C_\alpha x^{-\alpha} \theta \frac{\Gamma(1+kt)}{\Gamma(kt)} \\ (**) &\sim \frac{1}{2} C_\alpha x^{-\alpha} (k\theta t) \quad \text{as} \quad t \to \infty. \end{aligned} \tag{79}$$

In $(**)$ we apply Formula (29) which describes the asymptotic behaviour of the moments of the gamma process. Moreover, the constant C_α is given in (3). □

Lemma 4.14. For $0 < p < \alpha < 2$, the fractional order moments of the α-stable Lévy motion time-changed by the gamma subordinator are given by:

$$E(|X(t)|^p) \sim (k\theta t)^{p/\alpha} (c_{\alpha,0}(p))^p \quad \text{as} \quad t \to \infty, \tag{80}$$

where the constant $(c_{\alpha,\beta}(p))^p$ has the following form:

$$(c_{\alpha,0}(p))^p = \frac{2^{p-1}\Gamma(1-\frac{p}{\alpha})}{p\int_0^\infty u^{-p-1}\sin^2(u)du}. \tag{81}$$

Proof. Applying the self-similarity of the α-stable Lévy motion and the expression (4) for the fractional order moments of the α-stable Lévy motion and using Formula (29) for the asymptotic behaviour of q-th order moment of $\{G_{k,\theta}(t)\}$, we get that for $0 < p < \alpha < 2$:

$$\begin{aligned} E(|X(t)|^p) &= E(|L(G_{k,\theta}(t))|^p) \\ &= E(|G_{k,\theta}(t)^{1/\alpha} L(1)|^p) \\ &= E(G_{k,\theta}(t)^{p/\alpha} |L(1)|^p) \\ &= E(G_{k,\theta}(t)^{p/\alpha}) E(|L(1)|^p) \\ &= \theta^{p/\alpha} \, \frac{\Gamma(p/\alpha + kt)}{\Gamma(kt)} (c_{\alpha,0}(p))^p \\ &\sim (k\theta t)^{p/\alpha} (c_{\alpha,0}(p))^p \quad \text{as} \quad t \to \infty, \end{aligned} \tag{82}$$

where the constant $(c_{\alpha,\beta}(p))^p$ is specified in (6) and for $\beta = 0$ takes the form presented in (83). □

The final summary of the results presented in this section is included in Table 1, where all calculated characteristics are shown.

Table 1. Comparison of the calculated characteristics of the stochastic processes considered in the paper.

	$X(t) := L(S_{\alpha_S}(t))$	$X(t) := L(T_{\alpha_T,\lambda}(t))$	$X(t) := L(G_{k,\theta}(t))$
$\phi_{X(t)}(u)$	$e^{-t\lvert u\rvert^{\alpha\cdot\alpha_S}}$	$e^{-t((\lvert u\rvert^{\alpha}+\lambda)^{\alpha_T}-\lambda^{\alpha_T})}$	$(1+\theta\lvert u\rvert^{\alpha})^{-kt}$
$CD(X(t), X(s))$	$-2\min(s,t)$	$-2\min(s,t)((1+\lambda)^{\alpha_T} - \lambda^{\alpha_T})$	$-2\min(s,t)k\log(1+\theta)$
$P(X(t) > x)$	$\frac{1}{2} C_{\alpha_S\cdot\alpha} x^{-\alpha_S\cdot\alpha} t$	$\frac{1}{2} C_{\alpha} x^{-\alpha} (t\alpha_T \lambda^{\alpha_T - 1})$	$\frac{1}{2} C_{\alpha} x^{-\alpha} (k\theta t)$
$E\lvert X(t)\rvert^p$	$(c_{\alpha_S\cdot\alpha,0}(p))^p t^{\frac{p}{\alpha_S\cdot\alpha}}$	$(t\alpha_T \lambda^{\alpha_T-1})^{p/\alpha} (c_{\alpha,0}(p))^p$	$(k\theta t)^{p/\alpha} (c_{\alpha,0}(p))^p$

5 Simulations

In this section we present a simulation method of the time-changed symmetric α-stable Lévy motion defined in Sect. 4. The main idea is to simulate independent trajectories of the external and internal processes and to take their superposition. We suppose we want to simulate a trajectory of $\{X(t)\}$ on the interval $[0, T]$, where T denotes the time horizon. We introduce the following grid: $t_i = ih$, where $h = T/n$ and $i = 0, 1, \ldots, n$. A sketch of the algorithm is presented below.

(I) In the first step we generate a trajectory of the α_S-stable subordinator by summing up the increments of $\{Y(t)\}$:

$$\begin{aligned} &Y(t_0) = 0 \\ &Y(t_i) = Y(t_{i-1}) + \delta_i \quad \text{for} \quad i = 1, \ldots, n, \end{aligned} \tag{83}$$

where δ_i are independent random variables having the Laplace transform given by the formula:

$$E(e^{-z\delta_i}) = e^{-hz^\alpha} \tag{84}$$

for the first model with α_S-stable subordinator, by the formula:

$$E(e^{-z\delta_i}) = e^{-h((z+\lambda)^\alpha - \lambda^\alpha)}. \tag{85}$$

for the second model with tempered α_T-stable subordinator and by the formula:

$$E(e^{-x\delta_i}) = (1+\theta z)^{-kh} \tag{86}$$

for the third model with gamma subordinator.
(II) In the second step we generate a trajectory of the subordinated process, i.e. we simulate the symmetric α-stable Lévy motion but we replace the time step with the increments of the subordinator generated in the first step:

$$\begin{aligned} &X(t_0) = 0 \\ &X(t_i) = X(t_{i-1}) + (Y(t_i) - Y(t_{i-1}))^{1/\alpha}\xi_i \quad \text{for} \quad i = 1, \ldots, n, \end{aligned} \tag{87}$$

where $\xi_i \sim S_\alpha(1,0,0)$ are independent random variables.

In Fig. 5 we present sample trajectories of the subordinators (top panels) and the corresponding symmetric α-stable Lévy motion time-changed by these subordinators (bottom panels). Left panels correspond to the first model with

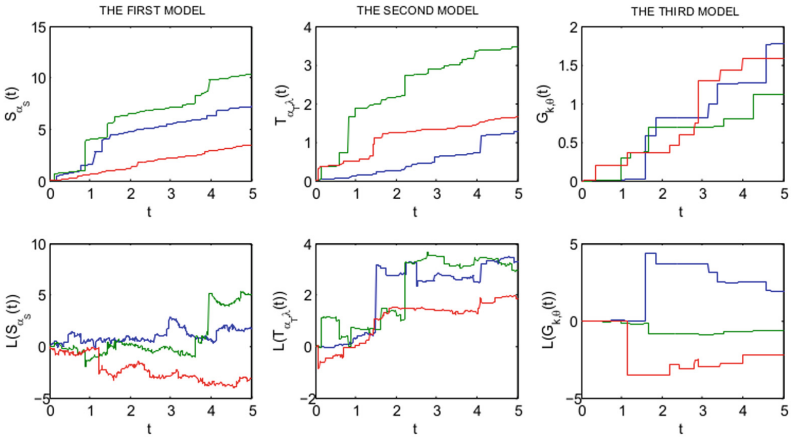


Fig. 5. Sample trajectories of the subordinators (top panels) and corresponding subordinated processes (bottom panels) for the symmetric α-stable Lévy motion time-changed by the α_S-stable subordinator with $\alpha = 1.8$ and $\alpha_S = 0.8$ (left panels), for the symmetric α-stable Lévy motion time-changed by the tempered α_T-stable subordinator with $\alpha = 1.5$, $\alpha_T = 0.5$ and $\lambda = 1$ (middle panels) and for the symmetric α-stable Lévy motion time-changed by the gamma subordinator with $\alpha = 1.6$, $k = 0.5$ and $\theta = 1$ (right panels).

α_S-stable subordinator, middle panels correspond to the second model with tempered α_T-stable subordinator and right panels correspond to the third model with gamma subordinator. The processes are generated on the interval $[0, 5]$ with $h = 0.005$ and the parameters are: $\alpha = 1.8$ and $\alpha_S = 0.8$ for the first model, $\alpha = 1.5$, $\alpha_T = 0.5$ and $\lambda = 1$ for the second model, $\alpha = 1.6$, $k = 0.5$ and $\theta = 1$ for the third model.

6 Estimation

In this section we propose an estimation procedure for the parameters of the processes presented in Sect. 4. We focus on the tail index estimation, which we denote as β, i.e. we want to estimate the parameter describing the rate at which the tail of the distribution converges to zero. In the proposed method we use the fact that the considered processes are Lévy and therefore their increments are stationary and independent. Thus, to estimate the parameters we calculate the increments of the processes $\Delta X_i = X(t_i) - X(t_{i-1})$ for $i = 1, \ldots, n$ and we examine their tails. For each discussed process, using the formulas describing the asymptotic behaviour of the tail of $X(t)$, we can write the corresponding formulas for the increments. Therefore, from the expression (43), for the increments of the symmetric α-stable Lévy motion time-changed by the α_S-stable subordinator we obtain:

$$\begin{aligned} P(\Delta X_i > x) &= P(X(t_i) - X(t_{i-1}) > x) = P(X(t_i - t_{i-1}) > x) \\ &= P(X(h) > x) \sim \frac{1}{2} C_{\alpha_S \cdot \alpha} x^{-\alpha_S \cdot \alpha} h \\ &\text{as} \quad x \to \infty \quad \text{for} \quad i = 1, \ldots, n, \end{aligned} \tag{88}$$

where the constant $C_{\alpha_S \cdot \alpha}$ is specified in (44) and h denotes the time step. Analogously, using Formula (59) for the increments of the symmetric α-stable Lévy motion time-changed by the tempered α_T-stable subordinator we have:

$$\begin{aligned} P(\Delta X_i > x) &= P(X(t_i) - X(t_{i-1}) > x) = P(X(t_i - t_{i-1}) > x) \\ &= P(X(h) > x) \sim \frac{1}{2} C_\alpha x^{-\alpha} (h \alpha_T \lambda^{\alpha_T - 1}) \\ &\text{as} \quad x \to \infty \quad \text{for} \quad i = 1, \ldots, n, \end{aligned} \tag{89}$$

where the constant C_α is specified in (60) and h denotes the time step. For the third model, from Formula (82), for the increments of the symmetric α-stable Lévy motion time-changed by the gamma subordinator we obtain:

$$\begin{aligned} P(\Delta X_i > x) &= P(X(t_i) - X(t_{i-1}) > x) = P(X(t_i - t_{i-1}) > x) \\ &= P(X(h) > x) \sim \frac{1}{2} C_\alpha x^{-\alpha} (k \theta h) \\ &\text{as} \quad x \to \infty \quad \text{for} \quad i = 1, \ldots, n, \end{aligned} \tag{90}$$

where the constant C_α is specified in (83) and h denotes the time step.

The estimation method is based on the idea of fitting the theoretical formulas presented above to the empirical tails of the increments of $\{X(t)\}$. The main steps of this procedure are described below.

1. Calculate the increments of the process $\{X(t)\}$:
$$\Delta X_i = X(t_i) - X(t_{i-1}) \quad \text{for} \quad i = 1, \ldots, n.$$
2. Estimate the empirical cumulative distribution function $F_n(x)$ of the increments:
$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\Delta X_i \leq x},$$
where $\mathbf{1}_A$ denotes the indicator of event A.
3. Estimate the right tail of the distribution $\hat{G}_n(x) = 1 - \hat{F}_n(x)$.
4. Plot $\hat{G}_n(x)$ in the log-log scale.
5. Fit the straight line:
$$y = ax + b \tag{91}$$
to the second linear part of the plot.
6. Put the tail index β equal to $-a$.

More precisely, in order to estimate the tail index β we use the least squares estimator. In the third step of the procedure described above, we calculate the right tail at points $x_1, x_2, \ldots, x_n$ and then, in the fourth and fifth step, we fit $\log(\hat{G}_n(x_i))$ by the linear function which form follows directly from the formulas (88), (89) and (90), and exemplary for the first model is given by $\log(\frac{1}{2} C_{\alpha_S \cdot \alpha} h) - \alpha_S \alpha \log(x_i)$ for $i = 1, 2, \ldots, n$. Since for a linear regression model,

$$y_i = ax_i + b + \epsilon_i, \quad i = 1, 2, \ldots, n,$$

the least squares estimator of the a parameter is given by

$$\hat{a} = \frac{\sum_{i=1}^{n} (x_i - \bar{x}) y_i}{\sum_{i=1}^{n} (x_i - \bar{x})^2},$$

thus in the considered case we obtain the following estimator of the tail index

$$\hat{\beta} = -\frac{\sum_{i=1}^{n} (\log(x_i) - \frac{1}{n} \sum_{i=1}^{n} \log(x_i)) \log(\hat{G}_n(x_i))}{\sum_{i=1}^{n} (\log(x_i) - \frac{1}{n} \sum_{i=1}^{n} \log(x_i))^2}.$$

To check the properties of the estimator we apply the method to the simulated data. In Fig. 6 one can see the bias $E(\hat{\beta}) - \beta$ of the considered estimator, obtained via Monte Carlo simulations, as a function of the length of trajectory. Top panel corresponds to the first model with α_S-stable subordinator, middle panel corresponds to the second model with tempered α_T-stable subordinator and bottom panel corresponds to the third model with gamma subordinator. In all three cases, the bias tends to zero as the number of observed points increases. Moreover, the efficiency of the procedure is also presented in Fig. 7. Here, we again use

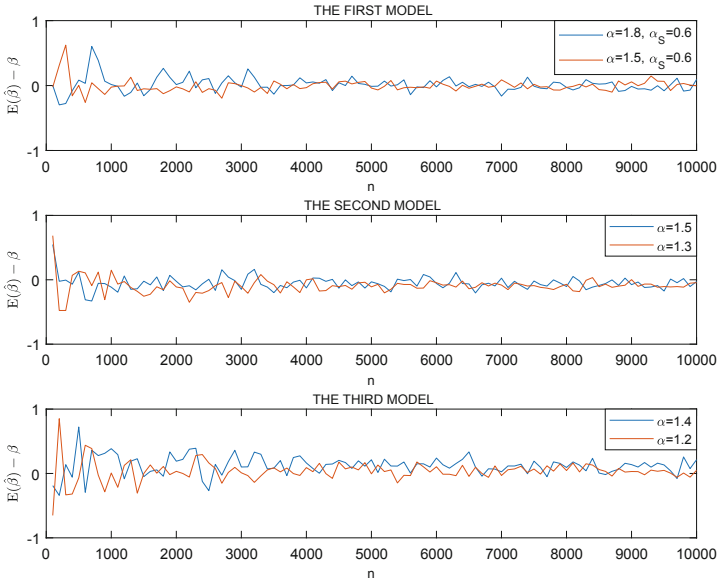


Fig. 6. Bias of the tail index estimator for the symmetric α-stable Lévy motion time-changed by the α_S-stable subordinator (top panel), for the symmetric α-stable Lévy motion time-changed by the tempered α_T-stable subordinator (middle panel) and for the symmetric α-stable Lévy motion time-changed by the gamma subordinator (bottom panel).

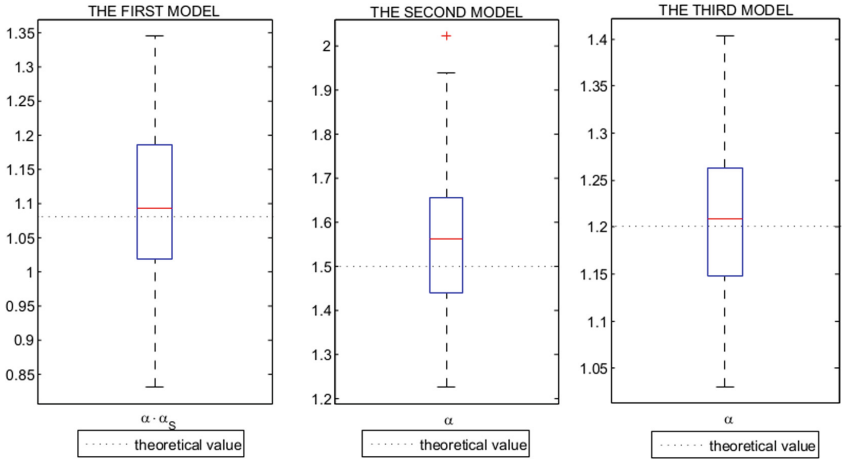


Fig. 7. Boxplots of the estimated tail index values for the symmetric α-stable Lévy motion time-changed by the α_S-stable subordinator (left panel), for the symmetric α-stable Lévy motion time-changed by the tempered α_T-stable subordinator (middle panel) and for the symmetric α-stable Lévy motion time-changed by the gamma subordinator (right panel).

the Monte Carlo method, i.e. we generate one trajectory of the stochastic process $\{X(t)\}$, we estimate the parameter using the procedure introduced above and then we repeat the experiment 100 times. We obtain the results corresponding to 100 realizations of the process $\{X(t)\}$ and we compare the median of these outcomes with the theoretical parameters of the process. The results of these experiments are presented in Fig. 7. Left panel corresponds to the first model with α_S-stable subordinator, middle panel corresponds to the second model with tempered α_T-stable subordinator and right panel corresponds to the third model with gamma subordinator. In all cases, the median of estimated values is close to the theoretical value of the tail parameter.

7 Conclusions

In this paper we have considered the α-stable Lévy motion subordinated by three different subordinators, namely α-stable, tempered α-stable and gamma. The new processes can be especially important in modelling of real data when they exhibit some properties of the external process (α-stable Lévy motion) however their some properties are different. In this case the pure α-stable Lévy motion can not be used. In this case we propose to replace the time in the external process by other process. This replacement can help to introduce a new process with very specific properties. From the theoretical point of view we have considered main properties of new processes and indicated their connections with external process. Moreover, we have also present the new estimation procedure and check its efficiency for simulated data.

Acknowledgements. This paper is supported by National Center of Science Opus Grant No. 2016/21/B/ST1/00929 "Anomalous diffusion processes and their applications in real data modelling".

References

1. Bochner S (1949) Diffusion equation and stochastic processes. Proc Nat Acad Sci USA 35(7):368–370
2. Bochner S (1955) Harmonic analysis and the theory of probability. University of California Press, Berkeley
3. Burnecki K, Weron A (2010) Fractional Lévy stable motion can model subdiffusive dynamics. Phys Rev E 82:021130
4. Failla R, Grigolini P, Ignaccolo M, Schwettmann A (2004) Random growth of interfaces as a subordinated process. Phys Rev E 70(1):010101
5. Gajda J, Magdziarz M (2010) Fractional Fokker-Planck equation with tempered alpha-stable waiting times Langevin picture and computer simulation. Phys Rev E 82:011117
6. Gajda J, Wyłomańska A, Zimroz R (2016) Subordinated continuous-time AR processes and their application to modeling behavior of mechanical system. Physica A 464(15):123–137
7. Golding I, Cox EC (2006) Physical nature of bacterial cytoplasm. Phys Rev Lett 96(9):098102

8. Hirsa A, Madan DB (2003) Pricing American options under variance gamma. J Comput Financ 7(2):63–80
9. Janczura J, Wyłomańska A (2012) Anomalous diffusion models: different types of subordinator distribution. Acta Phys Polon B 43(5):1001–1016
10. Janicki A, Weron A (1994) Simulation and chaotic behaviour of α-stable stochastic processes. Marcel Dekker, New York
11. Kumar A, Wyłomańska A, Gajda J (2017) Stable Lévy motion with inverse Gaussian subordinator. Physica A 482:486–500
12. Kumar A, Wyłomańska A, Połoczański R, Gajda J (2019) Fractional Brownian motion delayed by tempered and inverse tempered stable subordinators. Methodol Comput Appl Probab 21(1):185–202
13. Kumar A, Wyłomańska A, Połoczański R, Sundar S (2017) Fractional Brownian motion time-changed by gamma and inverse gamma process. Physica A 468:648–667
14. Lemmens D, Liang LZ, Tempere J, De Schepper A (2010) Pricing bounds for discrete arithmetic Asian options under Lévy models. Phys A Stat Mech Appl 389(22):5193–5207
15. Lévy PP (1925) Calcul des Probabilités, Gauthier Villars, Paris. Théorie del'Addition des Variables Aléatoires, 2nd edn. Gauthier Villars, Paris (1937)
16. Madan D, Carr P, Chang E (1998) The variance gamma process and option pricing. Rev. Financ. 2(1):79–105
17. Mantegna RN, Stanley SE (2000) An introduction to econophysics. Cambridge University Press, Cambridge
18. Nikias CL, Shao M (1995) Signal processing with alpha-stable distributions and applications. Wiley, New York
19. Peng C-K, Mietus J, Hausdorff JM, Havlin S, Stanley HE, Goldberger AL (1993) Long-range anticorrelations and non-gaussian behavior of the heartbeat. Phys Rev Lett 70:1343
20. Pipiras V, Taqqu MS (2017) Stable non-Gaussian self-similar processes with stationary increments. Springer, Cham
21. Rachev ST (ed) (2003) Handbook of heavy tailed distributions in finance. Elsevier bScience B.V., Amsterdam
22. Rachev S, Mittnik S (2000) Stable paretian models in finance. Wiley, New York
23. Samorodnitsky G, Taqqu MS (1994) Stable non-Gaussian random processes: stochastic models with infinite variance. Chapman & Hall, London
24. Seneta E (2000) The early years of the variance-gamma process. In: Fu MC, Jarrow RA, Yen J-YJ. Elliott RJ (ed.) Advances in Mathematical Finance, Birkhauser, Boston
25. Stanislavsky A, Weron K (2008) Two-time scale subordination in physical processes with long-term memory. Ann. Phys 323(3):643–653
26. Tankov P, Cont R (2003) Financial modelling with jump processes. Financial Mathematics Series. Chapman and Hall/CRC, London
27. Wyłomańska A, Chechkin A, Gajda J, Sokolov IM (2015) Codifference as a practical tool to measure interdependence. Physica A 421:412–429

Influence of Signal to Noise Ratio on the Effectiveness of Cointegration Analysis for Vibration Signal

Anna Michalak[2], Jacek Wodecki[2], Agnieszka Wyłomańska[1(✉)], and Radosław Zimroz[2]

[1] Faculty of Pure and Applied Mathematics, Hugo Steinhaus Center, Wroclaw University of Science and Technology, Na Grobli 15, 50-421 Wroclaw, Poland
agnieszka.wylomanska@pwr.edu.pl
[2] Faculty of Geoengineering, Mining and Geology, Diagnostics and Vibro-Acoustic Science Laboratory, Wroclaw University of Science and Technology, Na Grobli 15, 50-421 Wroclaw, Poland
{anna.michalak,jacek.wodecki}@pwr.edu.pl

Abstract. In this paper we consider the problem of local damage detection in rotating machines. Usually vibration signal is used for the analysis. Unfortunately classical methods are often not effective enough to properly detect cyclic components in real-life signal, especially for heavy-duty machinery used in mining industry. Therefore it is proposed to design analytical algorithms focused on the evaluation of the periodicity features of the signal. In this article we apply the cointegration approach to vibration signal. Although the method itself has been already proposed in the previous work, the main goal of this article is evaluation of the detection effectiveness with respect to changing Signal-to-Noise Ratio (SNR) of the input data. In the paper authors present the complete procedure for the effectiveness evaluation, which is based on the Monte Carlo simulations. The success rate of the method is assessed based on analyzing the cointegration with respect to SNR value of input signal. It allows to discover confidence level for such methodology considering internal quality of the data.

Keywords: Cointegration · Periodically correlated process · Monte Carlo simulation · Vibration signal · Signal-to-Noise Ratio · Local damage detection

1 Introduction

Local damage detection in rotating machinery has been a challenging problem for a long time. In the literature there are many methods for fault detection in gears and bearings [2,3,15,25]. In such case, local damage is manifested by the cyclic impulses that appears in the vibration signal. In many real cases the time series analysis approach can not be applied. This is the case when cyclic

F. Chaari et al. (Eds.): CSTA 2017, ACM 16, pp. 136–146, 2020.
https://doi.org/10.1007/978-3-030-22529-2_7

impulses are hidden below the noise floor or another components (low-frequency) which are related to normal operation of the machine [22,28,31,32]. For better results in such conditions, researchers often consider the vibration signal in other domains. However, for many real signals, this approach seems to not sufficient and therefore it is required to apply more advanced processing techniques [8,32].

The essence of this paper is an extension of previously developed analytical method [23] but the main goal of this article is to analyze its effectiveness with respect to different levels of signal to noise ratio (SNR) in the input data. In this case signal part considered to be the impulsive component carrying information about the damage. In order to achieve that, simulated signal is prepared incorporating the impulsive component with different amplitudes with respect to the background. After that, detection success rate is analyzed in relation to the increasing SNR [30].

Advantages of presented method come from the fact that it operates in time domain, which allows to omit any transformation to other domains, which is never truly lossless. It does not operate from the frequency viewpoint, which is the basis for most of the vibration-processing methods, and often comes down to informative frequency band (IFB) identification. On the other hand, it successfully identifies components otherwise invisible with any frequency-related processing methods.

The proposed method is based on the property of the signal, called cointegration, that is especially important in time series analysis. The idea of cointegration was introduced by Eangle in 1987 [13] and was applied to financial data [4,12,14]. Currently we can find examples in literature of cointegration in various fields (e.g. structural health monitoring and condition monitoring of wind turbines) [7–9,26,33].

We say that two processes X and Y are cointegrated, when neither of them hovers around a constant value or deterministic trend, but some linear combination of them does, so we can think of cointegration as a feature of long-term equilibrium relationship.

The rest of the paper is organized as follows: after defining the most important features, overall methodology is described. After that simulation methodology and results are presented.

2 Methodology

In this section authors present the theoretical background used in the further analysis.

Definition 1 (Periodically Correlated time series). *A second order process* $\{X(t)\}_{t\in Z}$ *is called periodically correlated with period* T *if for every* $s,t \in Z$ *the following conditions hold:*

$$\begin{cases} m(t) = E\left(X(t)\right) = m(t+T) \\ R(s,t) = Cov(X(t), X(s)) = R(s+T, t+T), \end{cases} \tag{1}$$

and there are no smaller values of $T > 0$ *for which the above conditions hold.*

PC time series are very popular in the literature [1,2,6,16,19,20,24]. In [5] it was proved that there is relation between periodically correlated time series and integrated and cointegrated one.

Definition 2 (Integrated time series). *A series is said to be integrated of order d, denoted* $X(n) \sim I(d)$ *if:*

$$(1-L)^d X(n) = z(n), \tag{2}$$

where $1-L$ *is the difference operator and* $z(n)$ *is stationary signal without trend [17].*

Definition 3 (Cointegrated time series). *The m time series* $\{X(1,n)\},\ldots,\{X(m,n)\}$ *are cointegrated with order d and b* $(X(1,n),\ldots,X(m,n) \sim CI(d,b))$ *if*

1. *all components of X are integrated with order d, we denote* $X \sim I(d)$,
2. *there exists a non-zero vector* $[a_1,\ldots,a_m]$ *such that*

$$a_1X(1,n) + a_2X(2,n) + \cdots + a_mX(m,n), \sim I(d-b), \quad b > 0. \tag{3}$$

The vector $[a_1,\ldots,a_m]$ *is called the cointegrating vector.*

In the literature one can find various methods that can be applied to detect the period T, see for instance [2]. However, in this article we use the algorithm which is based on the coherent statistic (CS) [18]. The CS is defined as:

$$|\gamma(\omega_p,\omega_q,M)|^2 = \frac{|\sum_{m=0}^{M-1} I_N(\omega_{p+m})\overline{I_N(\omega_{q+m})}|^2}{\sum_{m=0}^{M-1}|I_N(\omega_{p+m})|^2 \sum_{m=0}^{M-1}|I_N(\omega_{q+m})|^2}, \tag{4}$$

where $I_N(\omega) = \sum_{n=1}^{N} X(n)e^{-i\omega(n-1)}$ for $\omega_k = \frac{2\pi(k-1)}{N}$, $k = 1,2,\ldots,N$. Using statistic defined in (4) we can calculate one dimensional coherence $|\gamma(0,\omega_d,N-d)|^2$ which takes values between $[0,1]$. The peaks of CS for arguments $(\omega_d, 2\omega_d, 3\omega_d,\ldots)$ indicate the period T as $T = \frac{1}{\omega_d}$. The procedure is described in [23].

If the given time series is PC, then the appropriate selected subsignals are integrated. Moreover, they are also cointegrated, [5]. Let us assume that given time series $\{X(n)\}$ is PC with period T. In this case the subsignals defined as follows:

$$Y(v,n) = X(nT+v), \quad v = 1,2,\ldots,T \tag{5}$$

are cointegrated with the some order. The idea of constructing $\{Y(v,n)\}$ subsignals from the raw signal is demonstrated in Fig. 1.

In the literature one can find various statistics which can be used in the problem of testing if considered time series is integrated. The classical tests are Dickey-Fuller (DF), Augmented Dickey-Fuller (ADF), Kwiatkowski-Phillips-Schmidt-Shin (KPSS) [10,21]. In this paper we propose to use Durbin-Watson statistic which is given by following formula [11]:

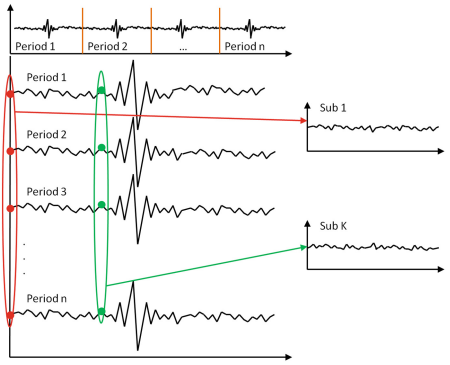


Fig. 1. Functional schematic of subsignals construction procedure, image thanks to courtesy of the Authors of [23].

$$IDW = \frac{\sum (X(n) - X(n-1))^2}{\sum \left(X(n) - \overline{X}(n)\right)^2}, \tag{6}$$

where $\overline{X}(n)$ is the sample mean of $\{X(n)\}$. This statistic always takes value between 0 and 4. The Durbin-Watson test is used to test the null hypothesis that linear regression residuals are uncorrelated, against the alternative that autocorrelation exists. Since null hypothesis cannot be rejected on the given confidence level, we say that time series of the signal is stationary.

Based on IDW test, one can conclude that analyzed time series is integrated with order 0 if IDW statistic for the signal is close to 2. Otherwise, we calculate once again the IDW statistic for differenced signal. If $IDW \sim 2$ for differenced series, then the raw signal is $I(1)$, if not we repeat the procedure unless the IDW statistic takes value close to 2. The number of times the time series was differentiated is equal to the order of integration d.

After checking if appropriate subsignals are integrated with given order, one can check if they are cointegrated. First, the cointegrating vector $[a_1, \ldots, a_m]$ should be estimated. This vector one can calculate using the least squares method in the multivariate regression model:

$$Y(1,n) = a_1 Y(2,n) - \cdots - a_{(m-1)} Y(m,n). \tag{7}$$

After that we calculate the residuals:

$$v(n) = Y(1,n) - a_1 Y(2,n) - \cdots - a_{(m-1)} Y(m,n)$$

As a final step, we check if residuals are integrated. In this case we use the approach based on the IDW statistic.

In the presented case we observe PC behaviour of the simulated signal. First step is to divide the signal by extracting T subsignals based on presented methodology. Next, we calculate IDW statistic and test if the received signals are integrated. Finally, one can calculate the cointegrating vector and examine if it

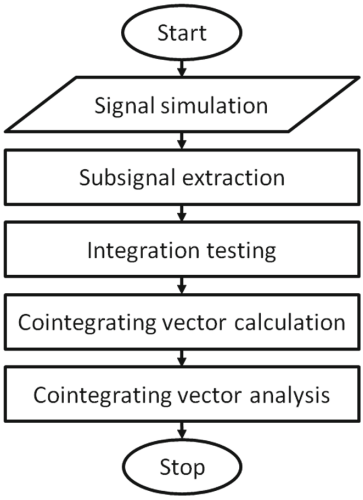


Fig. 2. Block diagram of the applied methodology

exhibits random (chaotic) behavior. In this case we apply the Wald-Wolfowitz test for randomness [27] for distances between nonzero values of the cointegrating vector $[a_1, \ldots, a_m]$. If the cointegrating vector comes from the healthy machine, the Wald-Wolfowitz test does not reject the hypothesis of randomness (H_0). In case of damaged machine – the test rejects the hypothesis. In our analysis we take under consideration the confidence level 0.05 (Fig. 2).

Procedure was performed for a predefined range of damage SNR, and the whole sequence has been repeated in Monte Carlo (MC) simulation. Then, after averaging MC results, the probability of damage detection has been calculated, that leads to determination of success rate of presented analytical method.

It is important to note that the aim of this article is not to present the analytical method already published in [23], but to evaluate its efficiency with respect to Signal-to-Noise Ratio of input data.

3 Simulations

In this section we analyzed the simulated data of the vibration signal from gearbox. The simulated signal consists of the Gaussian noise and the deterministic part given by four sine waves of 500, 1000, 1500 and 2000 Hz, acting as mesh components. The sine waves are frequency modulated with modulation frequency 4 Hz. The frequency of damage is equal to 8 Hz and the sampling frequency is 16,384 Hz. The simulated signals have 819,200 observations, which corresponds to 50 s of data. Thus the simulated model has the following form:

$$\text{Signal} = \text{N}(1 + \text{D}) + \text{C}, \tag{8}$$

where, N is the Gaussian noise, D is a deterministic function which contains information about the damage and C is a deterministic function with the signal characteristics mentioned above. We generate the signal using the procedure described in details [23,29]. Because the simulated signal is a sum of the noise and periodic functions thus it is PC. In [29] the relationship between simulated and real signals characteristics is discussed.

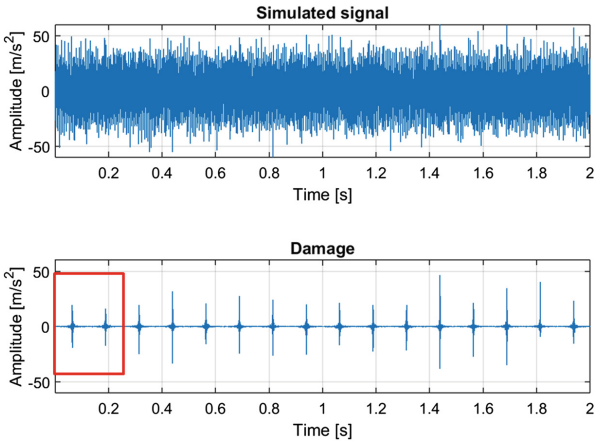


Fig. 3. Simulated signal corresponding to damaged machine (top panel) and cyclic impulses related to damage (bottom panel). Red frame denotes a single period detected in the signal, that is presented closer in Fig. 6

In Fig. 3 on the top panel we observe part of simulated signal corresponding to the damaged machine. In the bottom panel damage impulses are presented for informative purposes, which were added to healthy machine signal. For these impulses the ratio of impulses' amplitude to noise amplitude is 0.91.

3.1 Cointegration - A Numerical Example

First step is to calculate subsignals according to Fig. 1 and formula (5). The number of subsignals is the same as the value of period. The period which the signal has been simulated with, contains 4096 observations, this gives 0.25 s. It means that for period $T = 4096$ we will get 4096 subsignals. In Fig. 4 we present the exemplary subsignals coming from damaged machine.

After the extraction of subsignals, we test the if they are integrated. In this case we apply the Durbin-Watson statistic (see formula (6)). In Fig. 5 we present the IDW statistic for all subsignals. As one can see, this statistics is close to two thus the hypothesis that the subsignals are integrated with order 0 (I(0)) can be accepted.

In Fig. 6 in the top panel a part of simulated signal is presented, that contains the length of one determined period. In the middle panel there is one period of

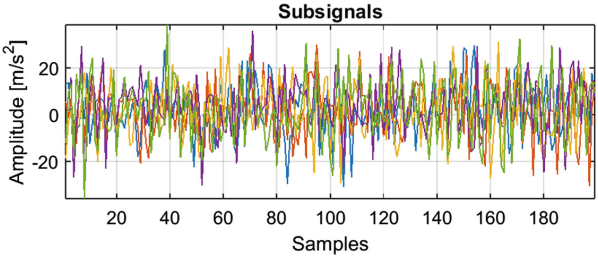


Fig. 4. The exemplary subsignals extracted from simulated signal

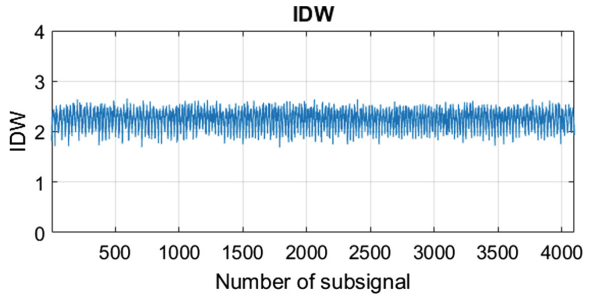


Fig. 5. Durbin-Watson statistic for each of subsignals.

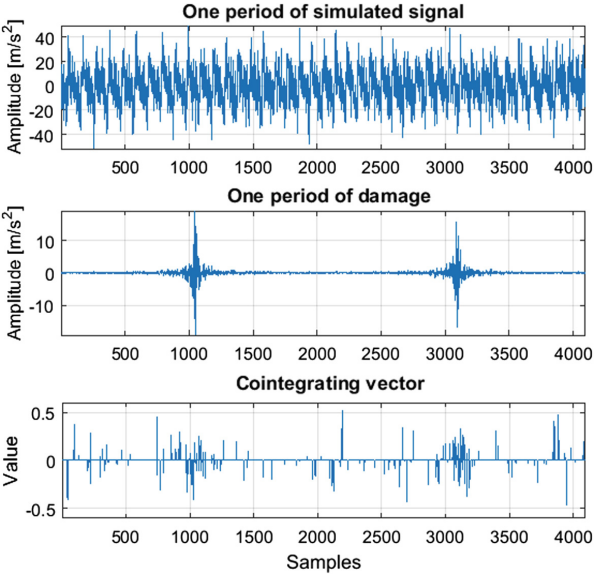


Fig. 6. Signal segment (from damaged machine) containing one period (top panel), part related to damage (middle panel) and estimated cointegrating vector (bottom panel).

damage. Bottom panel presents the cointegrating vector. As we can see, there are two groups in cointegrating vector which correspond with damage impulse in period of damage. To detect not randomness behaviour in cointegrating vector we used Wald-Wolfowitz test. P-value of this test is 4.3205e−07 what means that H_0 hypothesis is rejected and diagnosed machine is damaged.

3.2 Monte Carlo Simulation

In the Monte Carlo simulation signal is simulated with 51 monotonically increasing damage signal amplitudes. The number of MC iterations is 650. In Fig. 7 we present map of results of Wald-Wolfowitz test for randomness of just 100 MC iterations for better visibility. On vertical axis we observe number of MC simulations and on horizontal one the average ratio between damage and noise amplitude. Blue color on the map means that the cointegrating vector has random behaviour and diagnosed machine is not damaged. Yellow color is responsible for non random behaviour of cointegrating vector (we observe groups) and algorithm classified the machine as damaged. For the purpose of the core message of this paper, one can describe the results denoted by yellow color as a logical 1 (H_0 hypothesis is rejected) and blue color as logical 0 (H_0 hypothesis is not rejected at given confidence level). Considering this, the main goal is to test the success rate of the method for different levels of damage SNR.

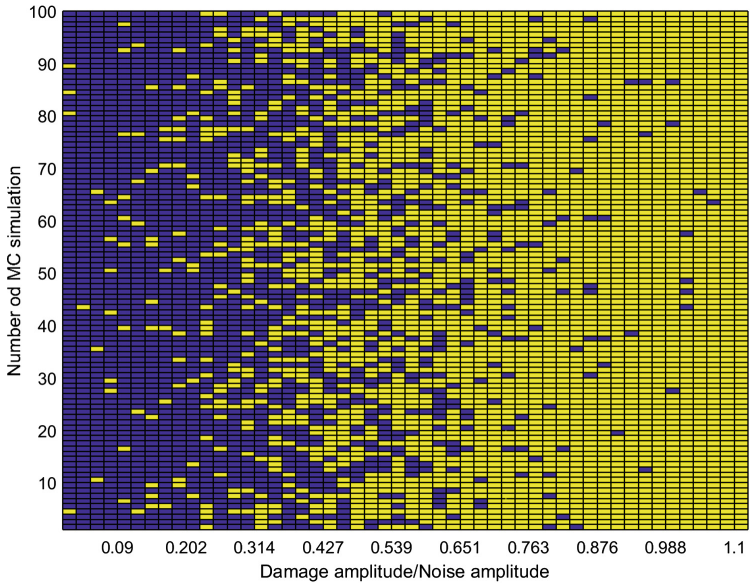


Fig. 7. Map of Wald-Wolfowitz test.

Based on the results of Wald-Wolfowitz test we calculate probability of damage detection with respect to damage SNR, based on 650 Monte Carlo

simulations by averaging the results of MC iterations (see Fig. 8). On 95% confidence level (red line) we detect damage when ratio between damage and noise amplitude is greater than 0.9.

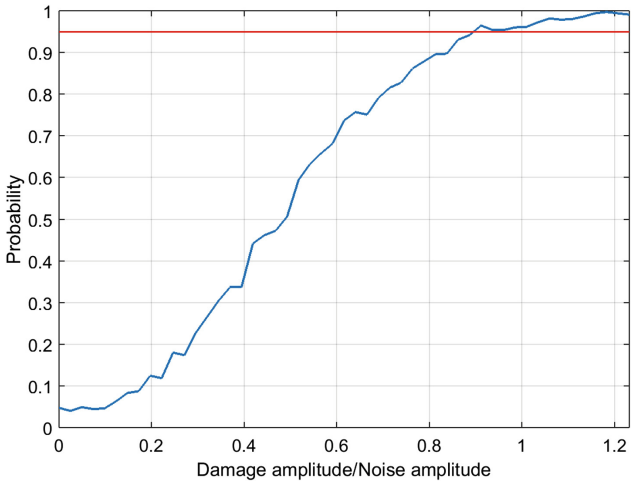


Fig. 8. Probability of damage detection based on 650 Monte Carlo simulations.

4 Conclusions

The integration and cointegration features yield great potential for local damage detection. While previously described method was proven to be useful, it is also very important to be aware of the restrictions related to the character and quality of the input data. In this paper authors evaluate the success rate of diagnostic method based on cointegration analysis with respect to SNR value of input signal. Results can allow to define confidence level for such analysis with respect to internal quality of the data.

References

1. Antoni J (2009) Cyclostationarity by examples. Mech Syst Signal Process 23(4):987–1036
2. Antoni J, Bonnardot F, Raad A, El Badaoui M (2004) Cyclostationary modelling of rotating machine vibration signals. Mech Syst Signal Process 18(6):1285–1314
3. Bartelmus W (2014) Gearbox dynamic multi-body modelling for condition monitoring under the case of varying operating condition. J Coupled Syst Multiscale Dyn 2(4):224–237
4. Broszkiewicz-Suwaj E, Makagon A, Weron R, Wyłomańska A (2004) On detecting and modeling periodic correlation in financial data. Phys A Stat Mech Appl 336(1):196–205

5. Broszkiewicz-Suwaj E, Wyłomańska A (2004) Periodic correlation–integration and cointegration. Technical report HSC/04/4, Wroclaw University of Technology
6. Cioch W, Knapik O, Leśkow J (2013) Finding a frequency signature for a cyclostationary signal with applications to wheel bearing diagnostics. Mech Syst Signal Process 38(1):55–64
7. Cross EJ, Worden K, Chen Q (2011) Cointegration: a novel approach for the removal of environmental trends in structural health monitoring data. Proc R Soc Lond A: Math Phys Eng Sci 467(2133):2712–2732
8. Dao PB (2018) Condition monitoring of wind turbines based on cointegration analysis of gearbox and generator temperature data. Diagnostyka 19(1):63–71
9. Dao PB, Staszewski WJ, Barszcz T, Uhl T (2018) Condition monitoring and fault detection in wind turbines based on cointegration analysis of SCADA data. Renew Energy 116:107–122
10. Dickey DA, Fuller WA (1979) Distribution of the estimators for autoregressive time series with a unit root. J Am Stat Assoc 74(366a):427–431
11. Durbin J, Watson GS (1951) Testing for serial correlation in least squares regression. Biometrika 38(1/2):159–177
12. Engel C (1996) A note on cointegration and international capital market efficiency. J Int Money Finan 15(4):657–660
13. Engle RF, Granger CW (1987) Co-integration and error correction: representation, estimation, and testing. Econom J Econom Soc 55:251–276
14. Engsted T, Tanggaard C (1994) Cointegration and the US term structure. J Bank Finan 18(1):167–181
15. Feng Z, Liang M, Chu F (2013) Recent advances in time-frequency analysis methods for machinery fault diagnosis: a review with application examples. Mech Syst Signal Process 38(1):165–205
16. Gardner WA, Napolitano A, Paura L (2006) Cyclostationarity: half a century of research. Signal Process 86(4):639–697
17. Hamilton JD (1994) Time series analysis, vol 2. Princeton University Press, Princeton
18. Hurd HL, Miamee A (2007) Periodically correlated random sequences: spectral theory and practice, vol 355. Wiley, Hoboken
19. Javorskyj I, Kravets I, Matsko I, Yuzefovych R (2017) Periodically correlated random processes: application in early diagnostics of mechanical systems. Mech Syst Signal Process 83:406–438
20. Kruczek P, Obuchowski J (2016) Cyclic modulation spectrum–an online algorithm. In: 2016 24th mediterranean conference on control and automation (MED), pp 361–365. IEEE
21. Kwiatkowski D, Phillips PC, Schmidt P, Shin Y (1992) Testing the null hypothesis of stationarity against the alternative of a unit root: how sure are we that economic time series have a unit root? J Econom 54(1–3):159–178
22. Makowski R, Zimroz R (2014) New techniques of local damage detection in machinery based on stochastic modelling using adaptive Schur filter. Appl Acoust 77:130–137
23. Michalak A, Wodecki J, Wyłomańska A, Zimroz R (2019) Application of cointegration to vibration signal for local damage detection in gearboxes. Appl Acoust 144:4–10
24. Napolitano A (2016) Cyclostationarity: new trends and applications. Signal Process 120:385–408
25. Samuel PD, Pines DJ (2005) A review of vibration-based techniques for helicopter transmission diagnostics. J Sound Vib 282(1):475–508

26. Tabrizi AA, Al-Bugharbee H, Trendafilova I, Garibaldi L (2016) A cointegration-based monitoring method for rolling bearings working in time-varying operational conditions. Meccanica 52:1–17
27. Wald A, Wolfowitz J (1940) On a test whether two samples are from the same population. Ann Math Stat 11(2):147–162
28. Wodecki J, Zdunek R, Wyłomańska A, Zimroz R (2017) Local fault detection of rolling element bearing components by spectrogram clustering with semi-binary NMF. Diagnostyka 18:3–8
29. Wyłomańska A, Obuchowski J, Zimroz R, Hurd H (2014) Periodic autoregressive modeling of vibration time series from planetary gearbox used in bucket wheel excavator. Springer International Publishing, Cham, pp 171–186
30. Wyłomańska A, Żak G, Kruczek P, Zimroz R (2017) Application of tempered stable distribution for selection of optimal frequency band in gearbox local damage detection. Appl Acoust 128:14–22
31. Żak G, Wyłomańska A, Zimroz R (2016) Data-driven vibration signal filtering procedure based on the α-stable distribution. J Vibroengineering 18(2):826–837
32. Żak G, Wyłomańska A, Zimroz R (2019) Periodically impulsive behavior detection in noisy observation based on generalized fractional order dependency map. Appl Acoust 144:31–39
33. Zolna K, Dao PB, Staszewski WJ, Barszcz T (2015) Nonlinear cointegration approach for condition monitoring of wind turbines. Math Probl Eng 2015:1–11

Ornstein-Uhlenbeck Process Delayed by Gamma Subordinator

Paula Poczynek, Piotr Kruczek, and Agnieszka Wyłomańska(✉)

Faculty of Pure and Applied Mathematics, Hugo Steinhaus Center,
Wroclaw University of Science and Technology,
Wybrzeze Wyspianskiego, 27, 50-370 Wroclaw, Poland
ppoczynek@cuprum.wroc.pl, {piotr.kruczek,agnieszka.wylomansk}@pwr.edu.pl

Abstract. The Ornstein-Uhlenbeck (OU) process is one of the most popular stochastic system applied in many different fields of studies. It was introduced in 1930 and can be considered as a continuous extension of the autoregressive model of order one, AR(1). Furthermore, the OU process in finance it is known as the Vasicek model and is mainly used in interest rate modelling. Furthermore, it is deeply studied and its main properties are well known. However, many real data exhibit some properties of the OU process although they cannot be directly modelled with this classical system. This is in case when certain characteristics adequate to the OU process are visible in the data however other properties of the classical model change. In such case the subordination scenario can be considered. In general, the subordination it is a time change of the original process. In this paper we consider the Ornstein-Uhlenbeck process delayed (subordinated) by Gamma subordinator. The Gamma subordinator is Lévy process of Gamma distribution. The main properties are studied, like the influence of the initial condition on the stationarity of the new subordinated process. Moreover, the formulas for the expected value and the autocovariance are derived. Furthermore, the simulation procedures and estimation algorithms are proposed.

Keywords: Ornstein-Uhlenbeck process · Gamma process · Subordination

1 Introduction

The Ornstein-Uhlenbeck (OU) process is one of the most popular stochastic system. It was introduced in 1930, [17]. Interestingly, it can be considered as a continuous extension of the autoregressive model of order one, AR(1), [10]. Furthermore, the OU process in finance it is known as the Vasicek model [18] and is mainly used in interest rate modelling, [7]. On the other hand, its applications to different financial data were widely consider in publications, [3,10,14,21]. It was shown in [15] that this model can be also successfully applied in optical physics. Furthermore, it can be also applied in vibration based diagnostics, [19].

F. Chaari et al. (Eds.): CSTA 2017, ACM 16, pp. 147–165, 2020.
https://doi.org/10.1007/978-3-030-22529-2_8

Many real data exhibit properties adequate to the OU process although they can not be directly modelled with this classical system. This is in case when certain characteristics adequate to the OU process are visible in the data however other properties of the classical model change. In such case it can be beneficial to consider so called subordination scenario. In the literature the subordinated process is also known as a time changed one. In general, it is constructed as a superposition of two independent process. The first one is called external or parent process while the second one is called subordinator. The subordinator is a process which can play a role of time therefore it is non-decreasing and positive. The subordinated process possesses some properties of the external one. On the other hand, some of the features change.

The idea of subordination was introduced in [5]. In the literature one can find different subordinated processes, for instance the popular Variance Gamma (VG) process, [12] arises as the Brownian motion with drift time-changed by Gamma process. The other classical example is the compound Poisson process, [4] or normal inverse Gaussian (NIG) process, [2]. In the literature one can find also different constructions of subordinated processes, like arithmetic Brownian motion, [20], or geometric Brownian motion, [8], subordinated by tempered stable process. In recent years in the theory of subordinated process the special attention is paid on so-called inverse subordinators. Among them the most classical is the inverse stable subordinator, [13]. As the examples of the subordinated processes with inverse subordinators we mention here tempered stable Lévy motion driven by inverse stable subordinator, [9], fractional Lévy stable process driven by inverse tempered stable subordinator, [16] or stable OU process time-changed by inverse stable subordinator, [10].

In this paper we introduce the OU process driven by Gamma subordinator. This process is some sense is an extension of the VG process however the more complicated structure of the external process is considered here. We focus on the main properties of the new process like stationarity or the formulas for expected value and autocovariance. In the application part it is shown how to simulate the trajectories of the new process and demonstrated the algorithms of the parameters estimation.

2 Ornstein-Uhlenbeck Process

Originally introduced by Uhlenbeck and Ornstein, the OU process is well known in finance as the Vasicek model, in which it is used as model for interest rate fluctuations. It has also been used in many applications including mathematics and physics. The model exhibits mean reversion, which means that in a long time period the process will go back to its equilibrium level. The definition of the classical OU process $\{X(t)\}$ is the solution of a stochastic differential equation of the following form [6]

$$dX(t) = \kappa(\mu - X(t))dt + \sigma dB(t), \tag{1}$$

where $\mu \in R$ is a long-term mean, κ is the speed of mean-reversion, σ is the volatility and $\{B(t)\}$ is the Brownian motion. The solution of this equation is given by

$$X(t) = X(0)e^{-\kappa t} + \mu(1 - e^{-\kappa t}) + \sigma \int_0^t e^{-\kappa(t-s)} dB(s).$$

The Fokker-Planck equation for the one-dimensional conditional probability density function (pdf) f(t, x) of an OU process is given by the formula [17]

$$\frac{\partial f(x,t)}{\partial t} = \kappa \frac{\partial}{\partial x}\big[(x-\mu)f(x,t)\big] + \frac{\sigma^2}{2}\frac{\partial^2 f(x,t)}{\partial x^2}.$$

Finally, one can observe that the stationary solution of this equation with $t \to \infty$ has a Gaussian distribution with mean μ and variance $\frac{\sigma^2}{2\kappa}$. Therefore, the pdf of this equation is following

$$f(x,t) = \sqrt{\frac{\kappa}{\pi\sigma^2}} e^{-\frac{\kappa(x-\mu)^2}{\sigma^2}}, \quad x \in \mathbb{R}.$$

2.1 Simulation

In this section the simulation procedure of OU process is presented. The most intuitive methods is based on the definition of the process. Indeed, in order to simulate the OU process the representation in stochastic differential equation can be used, see (1). The following algorithm can be applied:

1. Determine starting point X(0) and duration time T.
2. Create a grid of n equally distant points on interval $[0, T]$ and set $\tau = \frac{T}{n}$.
3. $\forall i \in (1, n)$ set $X(t_i) = \kappa(\mu - X(t_{i-1}))\tau + \sigma\sqrt{\tau}N$ where $N \sim N(0, 1)$.

Three exemplary trajectories of the OU process are presented in Fig. 1. Each of them has different starting point and the same parameters. Clearly, one can observe the ability of mean reversion of this process, all of the trajectories are close to mean $\mu = 1$.

2.2 Properties

In this section main properties of the OU process are presented, namely the moments and the autocovariance. Furthermore, depending on the starting point it can be weakly stationary or non-stationary. Indeed, two different cases have to be considered, namely the fixed and random initial condition. Firstly, let us assume that starting point is constant ($X(0) = \text{const}$).

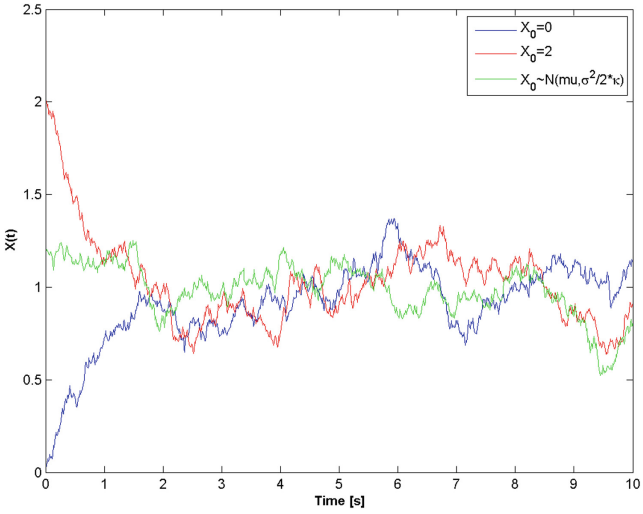


Fig. 1. Three trajectories of OU process with different initial condition and the same parameters $\mu = 1$, $\kappa = 1$, $\sigma = 0.2$.

Lemma 1. *The expected value of the OU process for fixed t with fixed initial condition has following form*

$$\mathrm{E}(X(t)) = X(0)e^{-\kappa t} + \mu(1 - e^{-\kappa t}). \tag{2}$$

Lemma 2. *The autocovariance of the OU process with fixed initial condition has following form*

$$\begin{aligned}\mathrm{Cov}\Big(X(t), X(t+h)\Big) &= \mathrm{E}\Big(X(t) - \mathrm{E}(X(t))\Big)\Big(X(t+h) - \mathrm{E}(X(t+h))\Big) \\ &= \frac{\sigma^2}{2\kappa}\Big(e^{-\kappa h} - e^{-\kappa(2t+h)}\Big).\end{aligned} \tag{3}$$

One can observe that $\mathrm{Var}\, X(t) = \frac{\sigma^2}{2\kappa}(1 - e^{-2t\kappa})$. Then variance converges to $\mathrm{Var}\, X(t) = \frac{\sigma^2}{2\kappa}$ with $t \to \infty$. Clearly the autocovariance depends on time, thus the OU process with such initial condition is non-stationary, however it is asymptotic stationary.

In the second case, let us assume that starting point is a random variable from Gaussian distribution, more precisely $X(0) \sim N(\mu, \frac{\sigma^2}{2\kappa})$.

Lemma 3. *The expected value of the OU process for fixed t with random initial condition from Gaussian distribution with mean μ and variance $\frac{\sigma^2}{2\kappa}$ has following form*

$$\mathrm{E}(X(t)) = \mu e^{-\kappa t} + \mu(1 - e^{-\kappa t}) = \mu. \tag{4}$$

Lemma 4. *The autocovariance of the OU process with random initial condition from Gaussian distribution with mean μ and variance $\frac{\sigma^2}{2\kappa}$ has following form*

$$\mathrm{Cov}\Big(X(t), X(t+h)\Big) = \frac{\sigma^2}{2\kappa}\Big(e^{-\kappa h}\Big).$$

In case on random initial condition from Gaussian distribution with mean μ and variance $\frac{\sigma^2}{2\kappa}$ the autocovariance and mean do not depend on the time. Thus, such OU process is weakly stationary. Finally, using Lemmas 1–4 we can formulate following Lemma.

Lemma 5. *For different starting points the OU process is weakly stationary or non-stationary*

1. *with fixed initial condition it is non-stationary,*
2. *with random initial condition from Gaussian distribution with mean μ and variance $\frac{\sigma^2}{2\kappa}$ it is weakly stationary.*

In order to confirm theoretical results 1000 trajectories of weakly stationary and non-stationary OU process are simulated. Then the empirical moments were computed and compared with theoretical formulas. The results are presented in Figs. 2 and 3, respectively. It is worth to mentioning, that autocovariance is computed for first 300 observations. As we observe the empirical functions coincide with the theoretical ones which indicates the calculation and proposed simulation method work well.

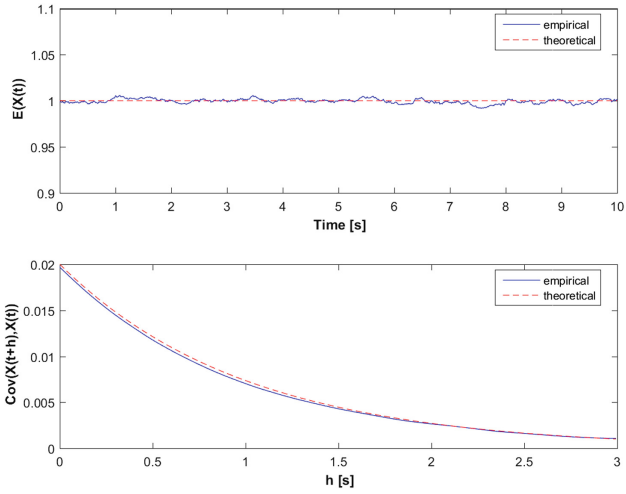


Fig. 2. Expected value (top panel) and autocovariance (bottom panel) of stationary OU process with $\mu = 1$, $\kappa = 1$, $\sigma = 0.2$ and $X(0) \sim N(\mu, \frac{\sigma^2}{2\kappa})$.

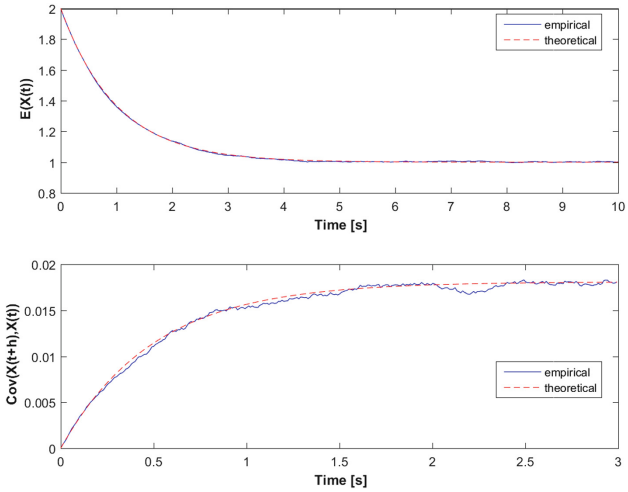


Fig. 3. Expected value (top panel) and autocovariance (bottom panel) of non-stationary OU process with $\mu = 1$, $\kappa = 1$, $\sigma = 0.2$, X(0) = 2, for fixed $h = 0.1$ s

3 Gamma Process

In this section main properties of Gamma process are recalled. Let $\{G_{k,\theta}(t)\}$ be Gamma Lévy process such that the increments are stationary, independent and have Gamma distribution with parameters k and θ [1]. A random variable is said to have Gamma(k, θ) distribution if its pdf is given by

$$f_G(x) = \frac{1}{\Gamma(k)\theta^k} x^{k-1} e^{-\frac{x}{\theta}},$$

where $x > 0$. Lévy process with positive increment can be used as subordinator. The Gamma process is non-negative and non-decreasing, thus it can be used as a subordinator.

3.1 Simulation

It is worth mentioning that, increments of Gamma process have Gamma distribution, thus simulation is straightforward. Indeed, it is sufficient to simulate this increments. Therefore, following algorithm can be applied

1. Set $G(0) = 0$,
2. $\forall i > 0$ set $G(i) = G(i-1) + X$ where $X \sim \Gamma(k, \theta)$.

Three exemplary Gamma process trajectories are presented in Fig. 4. Each of them has different parameters.

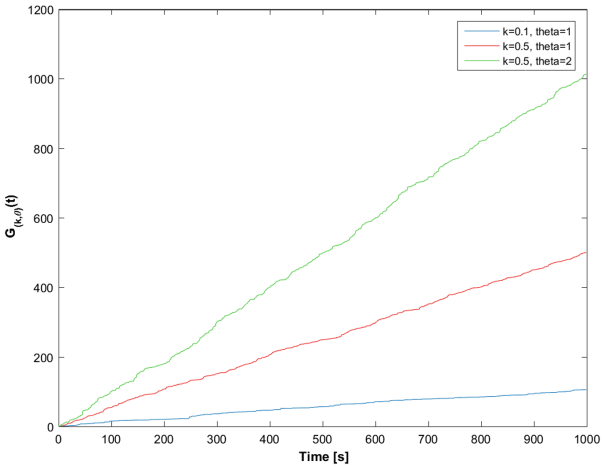


Fig. 4. Three trajectories of Gamma process with different parameters.

3.2 Properties

In this section the moments and the autocovariance of the Gamma process are presented. First of all let us introduce the formula for the moments

$$\mathrm{E}\big(G_{k,\theta}(t)^n\big) = \frac{\Gamma(n+kt)}{\Gamma(kt)\theta^n}.$$

It yields that $\mathrm{E}\big(G_{k,\theta}(t)^n\big) \sim \big(\frac{kt}{\theta}\big)^n$, when $t \to \infty$. From the properties of Gamma distribution the expected value for fixed t values is equal to

$$\mathrm{E}\big(G_{k,\theta}(t)\big) = \frac{kt}{\theta}. \tag{5}$$

Moreover one can easy derive formula for the variance

$$\mathrm{Var}\left(G_{k,\theta}(t)\right) = \frac{kt}{\theta^2}. \tag{6}$$

Finally, applying $\mathrm{E}(XY) = \frac{1}{2}\big(\mathrm{E}(X^2) + \mathrm{E}(Y^2) - \mathrm{E}(X-Y)^2\big)$, we obtain the formula for the autocovariance

$$\mathrm{E}\big(G_{k,\theta}(t), G_{k,\theta}(s)\big) = \frac{kt}{\theta^2}s + \frac{k^2}{\theta^2}st,$$

where $s < t$.

4 Subordinated Ornstein-Uhlenbeck Process

Real data often demonstrate characteristic behaviour of the OU process, namely mean reversion. However, in real signals the jumps can be observed. Such

behaviour is not demonstrated by the OU process. In order to overcome this drawback one can apply subordination approach. For instance, it is proposed to consider the combination of the OU process and Gamma process. In this section the main properties of the examined process are discussed. Furthermore, two cases are thoroughly explored.

Definition 1. *Let $\{G_{k,\theta}(t)\}$ be Gamma process, $\{X(t)\}$ be OU process and both processes are independent. The OU process delayed by Gamma process is defined as*

$$Y(t) = X\big(G_{k,\theta}(t)\big). \tag{7}$$

One can observe that time was replaced by another process, namely by Gamma one.

4.1 Simulation

The main idea of simulation of process $\{Y(t)\}$ given in 7, is to generate independent trajectories of the subordinator $\{G_{k,\theta}(t)\}$ and the OU process $\{X(t)\}$. For the OU process delayed by Gamma subordinator following algorithm can be used

1. Simulate the OU process $X(t)$.
2. Simulate Gamma process $G_{k,\theta}(t)$.
3. $\forall i \in (1, n)$ set $Y(t_i) = X(G_{k,\theta}(t_i))$.

A sample trajectory of weakly stationary and non-stationary OU process delayed by Gamma process are presented in Fig. 5 and in Fig. 6, respectively. Moreover, in order to illustrate the influence of subordination a corresponding weakly stationary and non-stationary OU process are presented in the same figures. For instance, the jumps are visible and mean reversion is preserved.

4.2 Properties

This section contains the information about the main properties of the OU process delayed by Gamma process, namely mean and autocovariance. Furthermore, two cases (stationary and non-stationary) are considered. Let us firstly examine the non-stationary process, when the initial condition is fixed.

Lemma 6. *The expected value of the OU process delayed by Gamma process for fixed t with fixed initial condition has following form*

$$\mathrm{E}\Big(X\big(G_{k,\theta}(t)\big)\Big) = \Big(X(0)\Big(\frac{1}{\theta}+\kappa\Big)^{-kt}\Big)\theta^{-kt} + \mu\Big(1-\Big(\frac{1}{\theta}+\kappa\Big)^{-kt}\theta^{-kt}\Big). \tag{8}$$

Proof 1. *Using Eqs. (2) and (5) and applying the conditional expected value, it can be derived that*

$$\begin{aligned}\mathrm{E}\Big(X\big(G_{k,\theta}(t)\big)\Big) &= \mathrm{E}\Big(\mathrm{E}\big(X(G_{k,\theta}(t))|G_{k,\theta}(t)\big)\Big)\\ &= \mathrm{E}\Big(X(0)e^{-\kappa G_{k,\theta}(t)} + \mu(1-e^{-\kappa G_{k,\theta}(t)})\Big).\end{aligned} \tag{9}$$

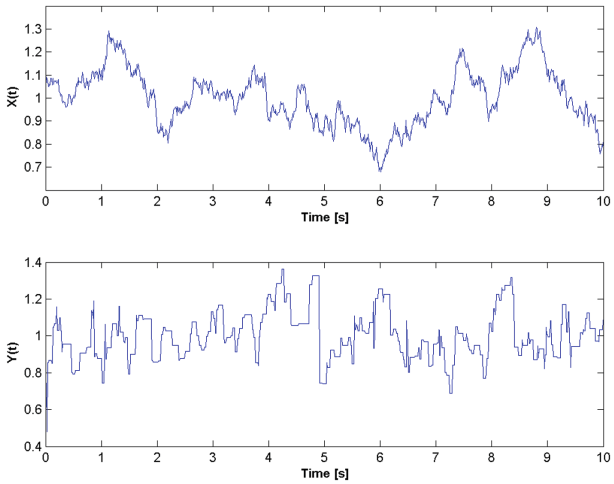


Fig. 5. Trajectory of weakly stationary OU process (top panel) and weakly stationary OU process delayed by Gamma process (bottom panel) with $k = \frac{1}{2}$, $\theta = 1$, both with $\mu = 1$, $\kappa = 1$, $\sigma = 0.2$ and $X(0) \sim N(\mu, \frac{\sigma^2}{2\kappa})$.

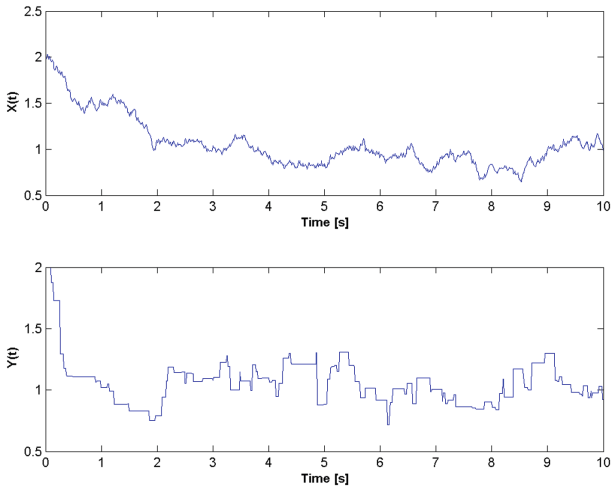


Fig. 6. Trajectory of non-stationary OU process (top panel) and non-stationary OU process delayed by Gamma process (bottom panel) with $k = \frac{1}{2}$, $\theta = 1$, both with $\mu = 1$, $\kappa = 1$, $\sigma = 0.2$ and X(0) = 2.

Furthermore,

$$\mathrm{E}\left(e^{-\kappa G_{k,\theta}(t)}\right) = \int_0^\infty \frac{1}{\Gamma(kt)\theta^{kt}} x^{kt-1} e^{-\frac{x}{\theta}} e^{-\kappa x} dx = \left(\frac{1}{\theta} + \kappa\right)^{-kt} \theta^{-kt}. \qquad (10)$$

Applying Eqs. (9) and (10) we finally obtain the mean value

$$\mathrm{E}\Big(X\big(G_{k,\theta}(t)\big)\Big) = \Big(X(0)\Big(\frac{1}{\theta}+\kappa\Big)^{-kt}\Big)\theta^{-kt} + \mu\Big(1-\Big(\frac{1}{\theta}+\kappa\Big)^{-kt}\theta^{-kt}\Big) \qquad \square$$

Lemma 7. *The autocovariance of the OU process delayed by Gamma process with fixed initial condition has following form*

$$\begin{aligned}
&\mathrm{Cov}\Big(X\big(G_{k,\theta}(t)\big), X\big(G_{k,\theta}(t+h)\big)\Big) \\
&= \frac{\sigma^2}{2\kappa}\Big(p(h)-p(h)v\Big) + X(0)^2 p(h)v + X(0)\mu\Big(p(t)-p(h)v+p(t+h)-p(h)v\Big) \\
&+ \mu^2\Big(1-p(t)-p(t+h)-p(h)v\Big) - p(t)p(t+h)\Big(X(0)-\mu+\frac{1}{\theta}^{-k(t+h)}\mu\Big)^2,
\end{aligned} \tag{11}$$

where $p(h) = \theta^{-kh}\big(\kappa+\frac{1}{\theta}\big)^{-kh}$ and $v = \big(2\kappa+\frac{1}{\theta}\big)^{-kt}\theta^{-kt}$.

Proof 2. *In case of non-stationary OU process it can be derived*

$$\begin{aligned}
\mathrm{E}\big(X(t)X(t+h)\big) &= \mathrm{Cov}(X(t),X(t+h)) + \mathrm{E}X(t)\mathrm{E}X(t+h) \\
&= \frac{\sigma^2}{2\kappa}\big(e^{-\kappa h}-e^{-\kappa(2t+h)}\big) + X(0)^2 e^{-\kappa t}e^{-\kappa(t+h)} \\
&+ X(0)\mu\big(e^{-\kappa t}\big(1-e^{-\kappa(t+h)}\big) + X(0)\mu e^{-\kappa(t+h)} \\
&\quad \big(1-e^{-\kappa t}\big) + \mu^2\big(1-e^{-\kappa t}\big)\big(1-e^{-\kappa(t+h)}\big).
\end{aligned}$$

Furthermore, in order to calculate the autocovariance of subordinated OU process let us compute

$$\begin{aligned}
&\mathrm{E}\Big(X\big(G_{k,\theta}(t)\big)X\big(G_{k,\theta}(t+h)\big)\Big) \\
&= \mathrm{E}\Big(E\big(X(G_{k,\theta}(t))X\big(G_{k,\theta}(t+h)\big)|G_{k,\theta}(t),G_{k,\theta}(t+h)\big)\Big) \\
&= \mathrm{E}\Big(\frac{\sigma^2}{2\kappa}\big(e^{-\kappa G_{k,\theta}(h)}-e^{-\kappa G_{k,\theta}(2t+h)}\big) + X(0)^2 e^{-\kappa G_{k,\theta}(t+h)}e^{-\kappa G_{k,\theta}(t)} \\
&+ X(0)\mu e^{-\kappa G_{k,\theta}(t)}\big(1-e^{-\kappa G_{k,\theta}(t+h)}\big) + X(0)\mu e^{-\kappa G_{k,\theta}(t+h)}\big(1-e^{-\kappa G_{k,\theta}(t)}\big) \\
&+ \mu^2\big(1-e^{-\kappa G_{k,\theta}(t)}-e^{-\kappa G_{k,\theta}(t+h)}+e^{-\kappa G_{k,\theta}(t)}e^{-\kappa G_{k,\theta}(t+h)}\big)\Big).
\end{aligned}$$

Furthermore

$$\begin{aligned}
\mathrm{E}\big(e^{-\kappa G_{k,\theta}(t)}e^{-\kappa G_{k,\theta}(t+h)}\big) &= \mathrm{E}\big(e^{-\kappa G_{k,\theta}(t)}e^{-\kappa(G_{k,\theta}(t+h)-G_{k,\theta}(t)+G_{k,\theta}(t))}\big) \\
&= \mathrm{E}\big(e^{-2\kappa G_{k,\theta}(t)}e^{-\kappa G_{k,\theta}(h)}\big) \\
&= \big(\kappa+\frac{1}{\theta}\big)^{-kh}\theta^{-kh}\big(2\kappa+\frac{1}{\theta}\big)^{-kt}\theta^{-kt}.
\end{aligned}$$

Finally, applying above equation it is easy to see that autocovariance is given by

$$\begin{aligned}
&\mathrm{Cov}\Big(X\big(G_{k,\theta}(t)\big), X\big(G_{k,\theta}(t+h)\big)\Big) = \mathrm{E}\Big(X\big(G_{k,\theta}(t)\big)X\big(G_{k,\theta}(t+h)\big)\Big) \\
&- \mathrm{E}\Big(X\big(G_{k,\theta}(t)\big)\Big)\mathrm{E}\Big(X\big(G_{k,\theta}(t+h)\big)\Big) = \frac{\sigma^2}{2\kappa}\Big(\theta^{-kh}\big(\kappa+\frac{1}{\theta}\big)^{-kh} - \big(\kappa+\frac{1}{\theta}\big)^{-kh} \\
&\theta^{-kh}\big(2\kappa+\frac{1}{\theta}\big)^{-kt}\theta^{-kt}\Big) + X(0)^2\big(\kappa+\frac{1}{\theta}\big)^{-kh}\theta^{-kh}\big(2\kappa+\frac{1}{\theta}\big)^{-kt}\theta^{-kt} \\
&+ X(0)\mu\Big(\theta^{-kt}\big(\kappa+\frac{1}{\theta}\big)^{-kt} - \big(\kappa+\frac{1}{\theta}\big)^{-kh}\theta^{-kh}\big(2\kappa+\frac{1}{\theta}\big)^{-kt}\theta^{-kt}\Big) \\
&+ X(0)\mu\Big(\theta^{-k(t+h)}\big(\kappa+\frac{1}{\theta}\big)^{-k(t+h)} - \big(\kappa+\frac{1}{\theta}\big)^{-kh}\theta^{-kh}\big(2\kappa+\frac{1}{\theta}\big)^{-kt}\theta^{-kt}\Big) \\
&+ \mu^2\Big(1 - \theta^{-kt}\big(\kappa+\frac{1}{\theta}\big)^{-kt} - \theta^{-k(t+h)}\big(\kappa+\frac{1}{\theta}\big)^{-k(t+h)} - \big(\kappa+\frac{1}{\theta}\big)^{-kh}\theta^{-kh} \\
&\big(2\kappa+\frac{1}{\theta}\big)^{-kt}\theta^{-kt}\Big) - \Big(\theta^{-kt}\big(\big(\kappa+\frac{1}{\theta}\big)^{-kt}(X(0)-\mu) + \frac{1}{\theta}^{-kt}\mu\big)\Big) \\
&\Big(\theta^{-k(t+h)}\big(\big(\kappa+\frac{1}{\theta}\big)^{-k(t+h)}(X(0)-\mu) + \frac{1}{\theta}^{-k(t+h)}\mu\big)\Big)
\end{aligned}$$

□

The second case is a weakly stationary process. Therefore, initial condition is following Gaussian distribution $X(0) \sim N(\mu, \frac{\sigma^2}{2\kappa})$.

Lemma 8. *The expected value of the OU process delayed by Gamma process for fixed t with random initial condition from Gaussian distribution with mean μ and variance $\frac{\sigma^2}{2\kappa}$ has following form*

$$\mathrm{E}\Big(X\big(G_{k,\theta}(t)\big)\Big) = \mu. \tag{12}$$

The above Lemma follows from the fundamental properties of the conditional expected value.

Lemma 9. *The autocovariance of the OU process delayed by Gamma process with random initial condition from Gaussian distribution with mean μ and variance $\frac{\sigma^2}{2\kappa}$ has following form*

$$\mathrm{Cov}\Big(X\big(G_{k,\theta}(t)\big), X\big(G_{k,\theta}(t+h)\big)\Big) = \frac{\sigma^2}{2\kappa}\theta^{-kh}\Big(\kappa+\frac{1}{\theta}\Big)^{-kh}. \tag{13}$$

Proof 3. *Let us recall that*

$$\mathrm{E}\big(X(t)X(t+h)\big) = \frac{\sigma^2}{2\kappa}\big(e^{-\kappa h}\big) + \mu^2. \tag{14}$$

Using (12) and (14) we obtain

$$\begin{aligned}
&\mathrm{E}\Big(X\big(G_{k,\theta}(t)\big)X\big(G_{k,\theta}(t+h)\big)\Big) \\
&= \mathrm{E}\Big(\mathrm{E}\Big(X\big(G_{k,\theta}(t)\big)X\big(G_{k,\theta}(t+h)\big)|G_{k,\theta}(t), G_{k,\theta}(t+h)\Big)\Big) \\
&= \mathrm{E}\Big(\frac{\sigma^2}{2\kappa}\big(e^{-\kappa G_{k,\theta}(h)}\big) + \mu^2\Big) = \frac{\sigma^2}{2\kappa}\theta^{-kh}\Big(\kappa+\frac{1}{\theta}\Big)^{-kh} + \mu^2.
\end{aligned}$$

Finally, the autocovariance for subordinated OU process with Gamma subordinator is following

$$\mathrm{Cov}\Big(X\big(G_{k,\theta}(t)\big), X\big(G_{k,\theta}(t+h)\big)\Big) = \mathrm{E}\Big(X\big(G_{k,\theta}(t)\big)X\big(G_{k,\theta}(t+h)\big)\Big)$$
$$- \mathrm{E}\Big(X\big(G_{k,\theta}(t)\big)\Big)E\Big(X\big(G_{k,\theta}(t+h)\big)\Big) = \frac{\sigma^2}{2\kappa}\theta^{-kh}\Big(\kappa + \frac{1}{\theta}\Big)^{-kh} \qquad \square$$

Lemma 10. *For different starting points the OU process delayed by Gamma process is weakly stationary or non-stationary*

1. *with fixed initial condition it is non-stationary,*
2. *with random initial condition from Gaussian distribution with mean μ and variance $\frac{\sigma^2}{2\kappa}$ it is weakly stationary.*

In order to confirm the obtained results 1000 trajectories of weakly stationary OU process and non-stationary OU process delayed by Gamma process are simulated. Moreover, the empirical moments are computed and compared with theoretical formulas. The autocovariance is calculated for first 300 observations. The results are presented in Figs. 7 and 8. One can observe that the empirical functions coincide with the theoretical ones, thus it confirms that simulations and theoretical values are performed properly.

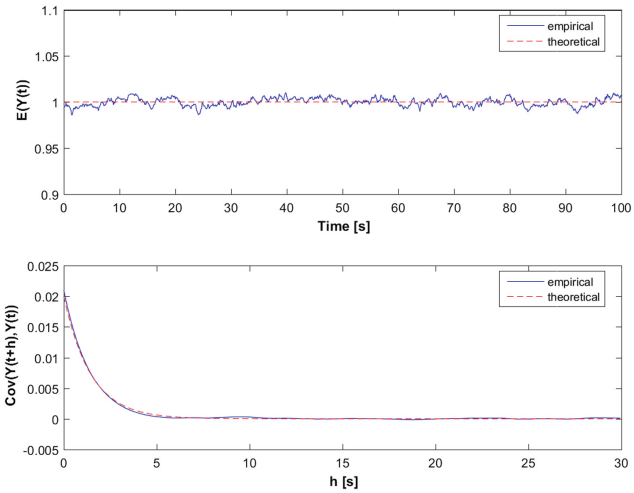


Fig. 7. Expected value (top panel) and autocovariance (bottom panel) of weakly stationary OU process with $\mu = 1$, $\kappa = 1$, $\sigma = 0.2$ delayed by Gamma process with $k = 2$, $\theta = 1$ and $X(0) \sim N(\mu, \frac{\sigma^2}{2\kappa})$.

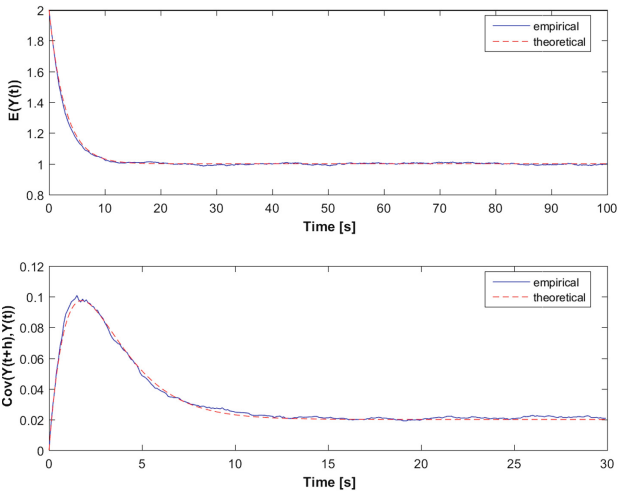


Fig. 8. Expected value (top panel) and autocovariance (bottom panel) of non-stationary OU process with $\mu = 1$, $\kappa = 1$, $\sigma = 0.2$, delayed by Gamma process with $k = 2$, $\theta = 1$ and $X(0) = 2$ for fixed $h = 0.1\,\mathrm{s}$.

4.3 Distribution

In this section a distribution of the OU process delayed by Gamma process is derived. It is known, that the OU process have a Gaussian distribution with appropriate mean and variance. Furthermore, when $X \sim N(m, s)$ then cumulative distribution functions is given by $\Phi(z) = \frac{1}{2}\Big(1 + erf\big(\frac{z-m}{\sqrt{2}s}\big)\Big)$. There are two cases considered, namely weakly stationary and non-stationary.

Lemma 11. *Cumulative distribution function of the OU process delayed by Gamma process with fixed initial condition is given by*

$$\begin{aligned} P\Big(X\big(G_{k,\theta}(t)\big) \le x\Big) &= \mathrm{E}\Big(\mathbb{1}_{X(G_{k,\theta}(t)) \le x}\Big) = \mathrm{E}\Big[\mathrm{E}\big(\mathbb{1}_{X(G_{k,\theta}(t)) \le x} | G_{k,\theta}(t)\big)\Big] \\ &= \mathrm{E}\left[\Phi\Big(\frac{x - (\mu - (X(0) + \mu)e^{-kG_{k,\theta}(t)}}{\frac{\sigma^2}{2\kappa}(1 - e^{-2kG_{k,\theta}(t)})}\Big)\right] \\ &= \mathrm{E}\left[\frac{1}{2}\Big(1 + erf\Big(\frac{x - (\mu - (X(0) + \mu)e^{-kG_{k,\theta}(t)})}{\frac{\sqrt{2}\sigma^2}{2\kappa}(1 - e^{-2kG_{k,\theta}(t)})}\Big)\Big)\right]. \end{aligned}$$

In case with fixed initial condition, it is extremely difficult to derive the formula for probability density function. Therefore, it is not given in the explicit formula.

Lemma 12. *Cumulative distribution function of the OU process delayed by Gamma process with random initial condition from Gaussian distribution with mean* μ *and variance* $\frac{\sigma^2}{2\kappa}$ *is given by*

$$P\Big(X\big(G_{k,\theta}(t)\big) \leq x\Big) = \mathrm{E}\Big(\mathbb{1}_{X(G_{k,\theta}(t))\leq x}\Big) = \mathrm{E}\Big[\mathrm{E}\big(\mathbb{1}_{X(G_{k,\theta}(t))\leq x}|G_{k,\theta}(t)\big)\Big]$$
$$= \mathrm{E}\left[\Phi\Big(\frac{x-\mu}{\frac{\sigma^2}{2\kappa}}\Big)\right] = \mathrm{E}\left[\frac{1}{2}\Big(1+erf\Big(\frac{x-\mu}{\frac{\sqrt{2}\sigma^2}{2\kappa}}\Big)\Big)\right] = \Phi\Big(\frac{x-\mu}{\frac{\sigma^2}{2\kappa}}\Big).$$

It is easy to observe that if initial condition is $X(0) \sim N(\mu, \frac{\sigma^2}{2\kappa})$, then $X\big(G_{k,\theta}(t)\big) \sim N(\mu, \frac{\sigma^2}{2\kappa})$. Finally, 1000 trajectories of weakly stationary subordinated OU process are simulated. Furthermore, empirical cumulative distribution function and probability density function are computed and compared with theoretical formulas. Results are presented in Fig. 9. It can be observed that, the empirical functions coincide with the theoretical ones.

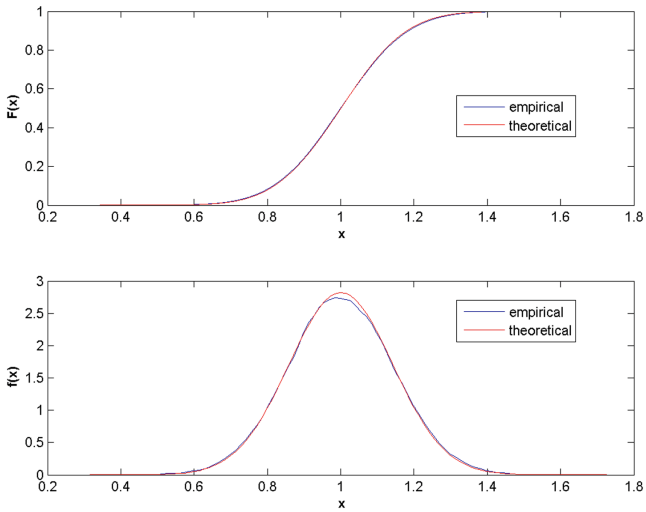


Fig. 9. Probability density function (top panel) and cumulative distribution function (bottom panel) of stationary OU process with $\mu = 1$, $\kappa = 1$, $\sigma = 0.2$ delayed by Gamma process with $k = 2$, $\theta = 1$ and $X(0) \sim N(\mu, \frac{\sigma^2}{2\kappa})$.

5 Estimation

In this section we describe the parameter's estimation of the considered model, namely weakly stationary OU process delayed by Gamma process. In case of real data usually only one trajectory of the data is available, thus we concentrated on the weakly stationary case. In this process there are five parameters. It can be quite complicated to estimate them simultaneously. Therefore, it was decided to

consider two different cases. First of all, let us assume the parameters of Gamma process are known, namely k and θ. In the second case, parameters of OU process are given, namely μ, κ and σ. It is worth mentioning that special case of Gamma distribution is exponential distribution, where $k = 1$. In order to illustrate the effectiveness of the estimation procedure the exponential process is used.

5.1 Estimation Procedure of Stationary OU Parameters

In this subsection it is assumed that, the parameters of Gamma process are known. Then parameters of OU process are estimated using method of moments. In the literature one can find another methods [11]. Let us recall that

$$\gamma(h) = \mathrm{Cov}\Big(X\big(G_{k,\theta}(t)\big)X\big(G_{k,\theta}(t+h)\big)\Big) = \frac{\sigma^2}{2\kappa}\theta^{-kh}\Big(\kappa + \frac{1}{\theta}\Big)^{-kh}$$

and

$$\mathrm{E}\Big(X\big(G_{k,\theta}(t)\big)\Big) = \mu.$$

Therefore, we obtain

$$\begin{cases} \mathrm{E}\Big(X\big(G_{k,\theta}(t)\big)\Big) = \mu \\ \gamma(0) = \frac{\sigma^2}{2\kappa} \\ \gamma(h_1) = \frac{\sigma^2}{2\kappa}\theta^{-kh_1}\Big(\kappa + \frac{1}{\theta}\Big)^{-kh_1}. \end{cases}$$

Hence, the following formulas for the parameters can be derived

$$\begin{cases} \mu = \mathrm{E}\Big(X\big(G_{k,\theta}(t)\big)\Big) \\ \kappa = \frac{\gamma(h_1)}{\gamma(0)}^{-\frac{1}{kh_1}} - \frac{1}{\theta} \\ \sigma = \sqrt{2\kappa\gamma(0)}. \end{cases}$$

Finally, the estimators are easilly obtained

$$\begin{cases} \widehat{\mu} = \widehat{\mathrm{E}}\Big(X\big(G_{k,\theta}(t)\big)\Big) \\ \widehat{\kappa} = \frac{\widehat{\gamma}(h_1)}{\widehat{\gamma}(0)}^{-\frac{1}{kh_1}} - \frac{1}{\theta} \\ \widehat{\sigma} = \sqrt{2\kappa\widehat{\gamma}(0)}, \end{cases}$$

where $\widehat{E}\Big(X\big(G_{k,\theta}(t)\big)\Big)$ is empirical mean and $\widehat{\gamma}(h)$ is empirical autocovariance. In order to check the effectiveness of the described estimation procedure we simulate 1000 trajectories of length 1000 of the considered process, with parameters $\mu = 1$, $\kappa = 1$, $\sigma = 0.2$, $k = 1$ and $\theta = 1$, and for each of them we estimate the parameters μ, κ and σ. Then, we compute the boxplot of the obtained estimators. In Fig. 10 the boxplot of estimators are presented, and the theoretical values are marked in green. The estimated values of parameters are similar to the theoretical ones.

Furthermore in Fig. 11 the histograms which show the deviation between the estimators and theoretical values of the parameters are presented. One can observe that most of the observations are close to 0.

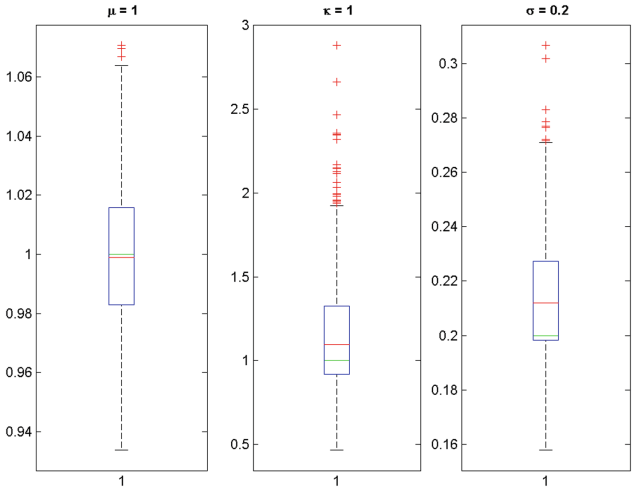


Fig. 10. The boxplots of estimated OU parameters with known parameters of the Gamma process. Theoretical parameters of the OU process are equal $\mu = 1$, $\kappa = 1$, $\sigma = 0.2$.

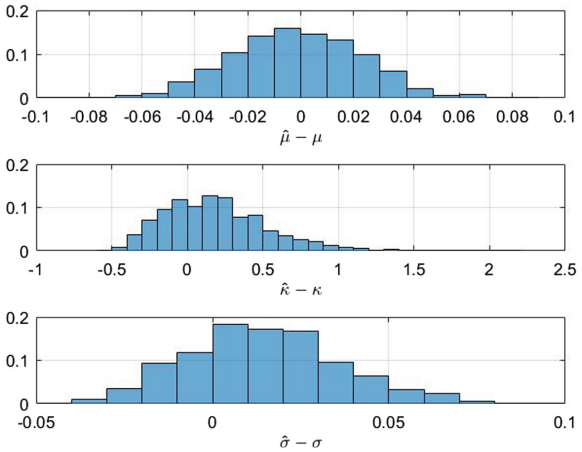


Fig. 11. The histograms of estimated OU parameters minus theoretical known parameters of the Gamma process. Theoretical parameters of the OU process are equal $\mu = 1$, $\kappa = 1$, $\sigma = 0.2$.

5.2 Estimation Procedure of the Gamma Process Parameters

In this subsection let us assume the parameters of OU process are known. Then, the Gamma process parameters are estimated. The approach is based on minimizing the difference between theoretical and empirical autocovariance. For instance, for each trajectory we compute the theoretical and empirical autocovariance and we minimize the difference with respect to the parameters.

In order to check the effectiveness of the described estimation procedure we simulate 1000 trajectories of length 1000 of the considered process, with parameters $\mu = 1$, $\kappa = 1$, $\sigma = 0.2$, $k = 1$ and $\theta = 1$, and for each of them we estimate the parameters k and θ. In Fig. 12 the boxplot of estimated values are presented, and the theoretical values are marked in green. In this case we also see the estimated values of appropriate parameters coincide with the theoretical ones which confirms the effectiveness of the estimation procedure.

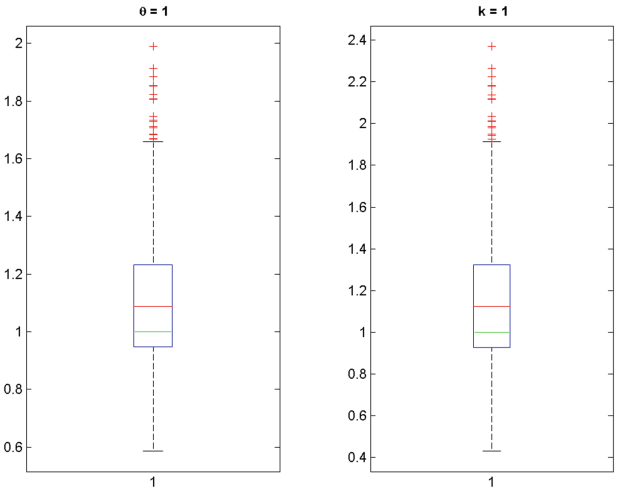


Fig. 12. The boxplots of estimated Gamma parameters with known parameters of the OU process. Theoretical parameters of Gamma process are equal $\theta = 1$, $k = 1$.

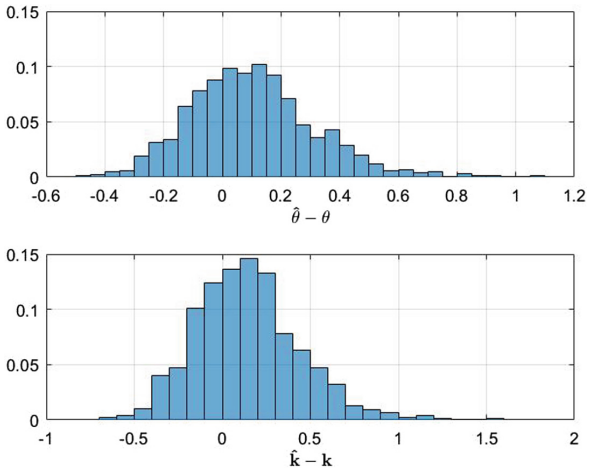


Fig. 13. The histograms of estimated Gamma parameters minus theoretical known parameters of the Gamma process. Theoretical parameters of Gamma process are equal $\theta = 1$, $k = 1$.

Furthermore in Fig. 13 the histograms which show the deviation between the estimators and theoretical values of the parameters for Gamma process are presented. One can observe that most of the observations are close to 0.

6 Conclusions

In the paper the OU process time-changed by Gamma subordinator was studied. Many different properties of this system were analyzed. For instance, it was shown that stationarity of this process depends on the initial condition, similar as in case of the classical OU process. Therefore, easily the stationary and non-stationary process can be obtained through the subordination scenario. Furthermore, the simulation algorithm was provided. Finally, the estimation procedures were discussed. Clearly, the analyzed process is complex and estimation is challenging. Therefore, it is proposed to estimate the Gamma process and OU process parameters separately. By using Monte Carlo simulations we confirmed the effectiveness of the estimation procedures. In the future it is planned to study the convergence of the estimators.

Acknowledgment. This paper is supported by National Center of Science Opus Grant No. 2016/21/B/ST1/00929 "Anomalous diffusion processes and their applications in real data modelling".

References

1. Applebaum D (2009) Levy processes and stochastic calculus. Cambridge University Press, Cambridge
2. Barndorff-Nielsen O (1997) Normal inverse Gaussian distributions and stochastic volatility modelling. Scand J Stat 24:1–13
3. Barndorff-Nielsen OE, Shephard N (2001) Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. J R Stat Soc Ser B (Stat Methodol) 63(2):167–241
4. Beard RE, Pentikaeinen T, Pesonen E (1984) Compound Poisson process. In: Beard RE, Pentikäinen T, Pesonen E (eds) Risk theory, vol 20. Monographs on statistics and applied probability. Springer, Dordrecht
5. Bochner S (1949) Diffusion equation and stochastic processes. Proc Natl Acad Sci 35(7):368–370
6. Cizek P, Hardle WK, Weron R (2005) Statistical tools for finance and insurance. Springer, Heidelberg
7. Chan KC, Karolyi GA, Longstaff FA, Sanders AB (1992) An empirical comparison of alternative models of the short-term interest rate. J Financ 47(3):1209–1227
8. Gajda J, Wyłomańska A (2012) Geometric Brownian motion with tempered stable waiting times. J Stat Phys 148:296–305
9. Gajda J, Wyłomańska A (2013) Tempered stable Levy motion driven by stable subordinator. Physica A 392:3168–3176
10. Janczura J, Orzeł S, Wyłomańska A (2011) Subordinated alpha-stable Ornstein-Uhlenbeck process as a tool of financial data description. Physica A 390:4379–4387

11. Lipster RS, Shiryayev AN (1974) Statistics of random processes II: applications. Springer, Heidelberg
12. Madan D, Carr P, Chang E (1998) The variance Gamma process and option pricing. Eur Financ Rev 2:79–105
13. Magdziarz M, Weron A (2007) Competition between subdiffusion and Levy flights: stochastic and numerical approach. Phys Rev E 75:056702
14. Obuchowski J, Wyłomańska A (2013) Ornstein-Uhlenbeck process with non-Gaussian structure. Acta Phys Pol B 44(5):1123–1136
15. Slezak J, Drobczynski S, Weron K, Masajada J (2014) Moving average process underlying the holographic-optical-tweezers experiments. Appl Opt 53(10):254–258
16. Teuerle M, Wyłomańska A, Sikora G (2013) Modelling anomalous diffusion by subordinated fractional Levy-stable process. J Stat Mech P05016
17. Uhlenbeck GE, Ornstein LS (1930) On the theory of the Brownian motion. Phys Rev 36(5):823–841
18. Vasicek O (1977) An equilibrium characterization of the term structure. J Financ Econ 5(2):177–188
19. Wang X, Makis V (2009) Autoregressive model-based gear shaft fault diagnosis using the Kolmogorov-Smirnov test. J Sound Vib 327(3):413–423
20. Wyłomańska A (2012) Arithmetic Brownian motion subordinated by tempered stable and inverse tempered stable processes. Phys A 391:5685–5696
21. Wyłomańska A (2011) Measures of dependence for Ornstein-Uhlenbeck processes with tempered stable distribution. Acta Phys Pol B 42(10):2049–2062

Combination of Kolmogorov-Smirnov Statistic and Time-Frequency Representation for P-Wave Arrival Detection in Seismic Signal

Jacek Wodecki[1], Anna Michalak[1], Paweł Stefaniak[2], Agnieszka Wyłomańska[3(✉)], and Radosław Zimroz[1]

[1] Faculty of Geoengineering, Mining and Geology, Wroclaw University of Science and Technology, Na Grobli 15, 50-421 Wroclaw, Poland
{jacek.wodecki,anna.michalak}@pwr.edu.pl
[2] KGHM Cuprum Ltd., Research and Development Centre, Sikorskiego 2-8, 53-659 Wroclaw, Poland
[3] Faculty of Pure and Applied Mathematics, Hugo Steinhaus Center, Wroclaw University of Science and Technology, 50-370 Wroclaw, Poland
agnieszka.wylomanska@pwr.edu.pl

Abstract. The problem of determination of P-wave onset moment is elementary in seismic signals analysis. This problem is also widely discussed in the literature. In recent years many methods based on statistical properties of the signal have arisen. From the mathematical perspective the problem reduces to segmentation of the raw signal into parts with different features. Having the knowledge of the particular P-wave onset moment (for couple differently located sensors), the establishment of corresponding event's location and energy is possible. The difference in signals' frequency spectra for the registered event and its preceding noise allows for using time-frequency domain in designating the onset moment. In this paper an innovative method for searching of the P-wave arrival is proposed. The method incorporates using signal's time-frequency representation (namely spectrogram) and Kolmogorov-Smirnov (KS) statistic analysis. We apply two-sample one-sided Kolmogorov-Smirnov statistic to spectra vectors obtained from the spectrogram. On the basis of KS map it is possible to find the regime switching point which indicates the P-wave onset moment. The method is tested on a real life signal originating from underground mine. Proposed methodology is compared with classical Short-Time Average over Long-Time Average (STA/LTA) algorithm and P-wave arrival moment indicated manually by the expert.

Keywords: Kolmogorov-Smirnov statistic · P-wave detection · Seismic data

F. Chaari et al. (Eds.): CSTA 2017, ACM 16, pp. 166–174, 2020.
https://doi.org/10.1007/978-3-030-22529-2_9

1 Introduction

One of the most crucial challenges in terms of seismic activity monitoring is recognition of seismic wave onset moment. Identification of this particular point in time leads to assessment of basic properties for registered event, i.e. its location, energy etc. [3]. These attributes are fundamental from view point of further seismic hazard assessment. Data acquisition system is used for monitoring of rock mass activity. This data should be next evaluated and interpreted (information about energy and localization of the event is required). In typical deep mine one could expect dozens of events per day. In practice this task is technically very difficult and complex, and daily will require extensive analysis of hundreds of signals from multiple different data receivers distributed spatially in wide area. Each signal should be interpreted immediately after measuring. Therefore it is recommended that such procedures for marking the P-wave onset moment should be automatic. In the literature this subject has been widely studied, however still the most commonly used technique is (STA/LTA) for preliminary detection, which is often corrected manually. When considering seismic signal which includes pure noise at the beginning (which is most commonly regarded as white) and actual seismic event subsequently, it is easy to denote a point in which the signal loses its stationary character. This is a consequence of occurrence of sudden instantaneous energy growth which is visible in the signal in form of amplitude variation increase. When considering time-frequency domain, sole change of signal power will be visible in spectrogram (due to Parseval's theorem). During P-wave onset moment the frequencies amplitudes rise as well as their proportions, so the short-time spectrum becomes much more non-uniform. The pure noise spectrum is flat and recorded seismic event has relatively wide frequency range at first, which is becoming narrower in time. Existing algorithms include STA/LTA, AR-based algorithms (finding different models for both event parts and a noise one [5,10]), algorithms using neural networks [13], or multiscale wavelet analysis [18]. The reader is referred to [4,15], where the comparison of some classical algorithms has been investigated. The recent work on this subject includes [7,9,12,19], and their comparison can be found in [11]. First mentioned group of algorithms is based mainly on sole time domain and ignores capability of frequency domain. The potential hidden in time-frequency domain has been presented in [2,17]. Effectiveness of these methods confirms an adequacy of using the spectrogram in term of finding P-wave onset moment. Hence, development of method based on spectrogram in P-wave recognition is justified.

Authors propose to analyze spectrogram of seismic signal using Kolmogorov-Smirnov statistic values arranged into a map [16]. In presented approach it is used to differentiate the structure of spectra vectors from the spectrogram. This enables the separation of spectra vectors from before and after the arrival moment. As a result we can determine arrival moment as a point located between obtained groups of spectra vectors. Such approach is fully data-driven in its analytical process.

2 Methodology

In this section the overall methodology is explained (see Fig. 1).

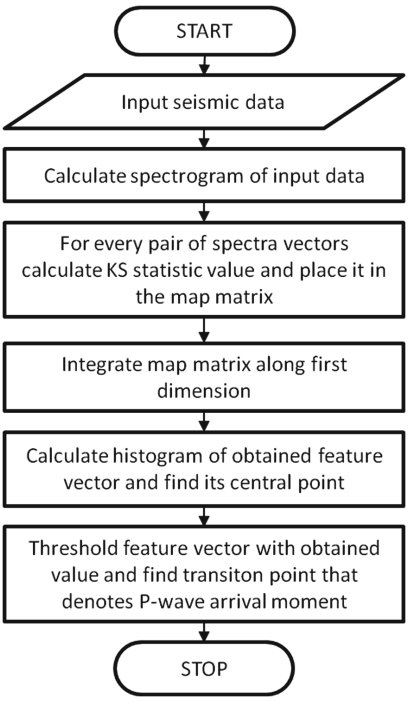


Fig. 1. Flowchart of presented procedure

Presented procedure begins with calculating spectrogram (see Eqs. 1 and 2). Its parameters are selected based on empirical testing and are suited best for the length of signal time series, in relation to the way that prior segmentation algorithm selects and segments the individual seismic events. For analyzed signal the parameters are presented in Table 1.

The short-time Fourier transform (STFT) for the discrete signal $x_0, x_2, \ldots,$ x_{N-1} is defined as follows [8]:

$$\mathrm{STFT}(t,f) = \sum_{m=0}^{L-1} x_{t+m}\omega_m e^{-j2\pi fm/N} \tag{1}$$

for $0 \leq f \leq f_{max}$ and $0 \leq t \leq t_{max}$. In the above equation ω is the shifted window of the length L. Furthermore, the spectrogram is squared absolute value of the STFT:

$$\mathrm{Spec}(t,f) = |\mathrm{STFT}(t,f)|^2. \tag{2}$$

After that, for each non-repeating pair of spectra vectors, the Kolmogorov-Smirnov (KS) statistic is calculated (see Definition 1).

Definition 1 (Kolmogorov-Smirnov statistic). *The two-sample one-sided Kolmogorov-Smirnov (KS) statistic for samples* $y_1^1, y_2^1, \ldots, y_K^1$ *and* $y_1^2, y_2^2, \ldots, y_K^2$ *of lengths* K *is given by following formula:*

$$D_{1,2}^* = \max_v \left(\hat{F}_{y^1}(v) - \hat{F}_{y^2}(v) \right), \tag{3}$$

where $\hat{F}_{y^i}(v)$ *is empirical cumulative distribution functions of vector* y^i, $i = 1, 2$ *in point* v *[6, 14].*

In formula (3) the empirical cumulative distribution function for sample $y_1, y_2, \ldots, y_K$ is defined as follows:

$$\hat{F}(v) = \frac{1}{K} \sum_{k=1}^{K} \mathbb{1}_{\{v_k \leq v\}}, \tag{4}$$

where $\mathbb{1}_A$ is an indicator of a set A.

In our methodology we apply the KS statistic to spectra vectors taken from the spectrogram. More precisely, for all $i, j \in \{0, ..., t_{max}\}$ we calculate the value of KS statistic. All of obtained values $D_{i,j}^*$ are arranged into upper-triangular part of a square matrix KS, such as:

$$KS_{(i,j)} = D_{i,j}^*, \quad i = 0, \ldots, t_{max}, \quad j = i, \ldots, t_{max}, \tag{5}$$

that is reflected with respect to the main diagonal in order to obtain symmetric matrix. Then, such matrix is integrated along one dimension producing a feature vector, which is then thresholded using central point of the histogram. It is performed by dividing spectrogram into two halves according to the item count, then finding two global modes as global maxima in each half of spectrogram, and finally defining central point as an average between modes. Threshold allows to automatically indicate the transition point between the processes (see Sect. 2.1).

2.1 KS Map and Its Interpretation

Based on initial visual evaluation of the input data one can notice two disjoint processes occurring consecutively:

- **Process 1:** No seismic shocks are present, signal can be considered as a noise (area (1,1));
- **Process 2:** Seismic event occurs, signal presents the impact and its dampening (area (2,2)).

In such case one can describe those processes in terms of signal parameters:

- **Process 1:** Low energy, uniform spectrum;
- **Process 2:** High energy, non-uniform and time-varying spectrum;

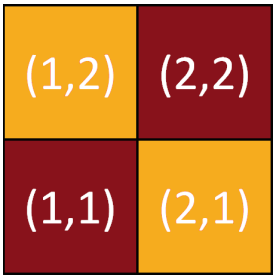


Fig. 2. Conceptual chart of expected KS map

Despite the relatively good visibility of the events in the input data, it is not easy to define one simple, effective and precise evaluation criterion. Thus, proposed methodology is based on the analysis of the map constructed by composing the values of KS statistic.

3 Results

Raw input signal of considered seismic event is presented in Fig. 3. In the first step spectrogram of the signal was calculated according to parameters in Table 1 (see Fig. 4).

After that, individual spectra vectors are compared with each other in a non-repeating manner, and KS statistic produced by each comparison is placed in a upper-triangular half of a square matrix. It is then reflected to form a symmetric square matrix, which is further called a map.

Table 1. Parameters of compared results

Parameter	Value
Sampling frequency	500 Hz
Window	Hamming, 32 samples
Overlap	0.95%
FFT points	256

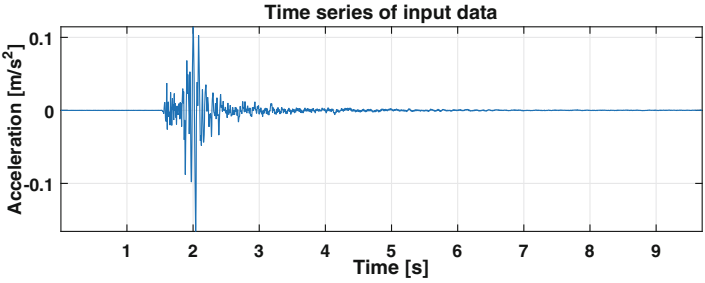


Fig. 3. Raw signal of seismic data

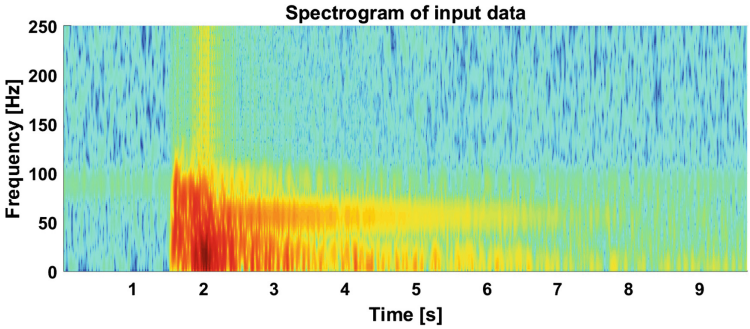


Fig. 4. Spectrogram of seismic data

It is hard to find transition points based on two-dimensional data. Hence, we propose to determine threshold based on one-dimensional statistic. Taking advantage of previously mentioned assumptions (see Sect. 2.1), it is expected that local sum of KS map vectors will vary according to the area (see Figs. 2 and 5).

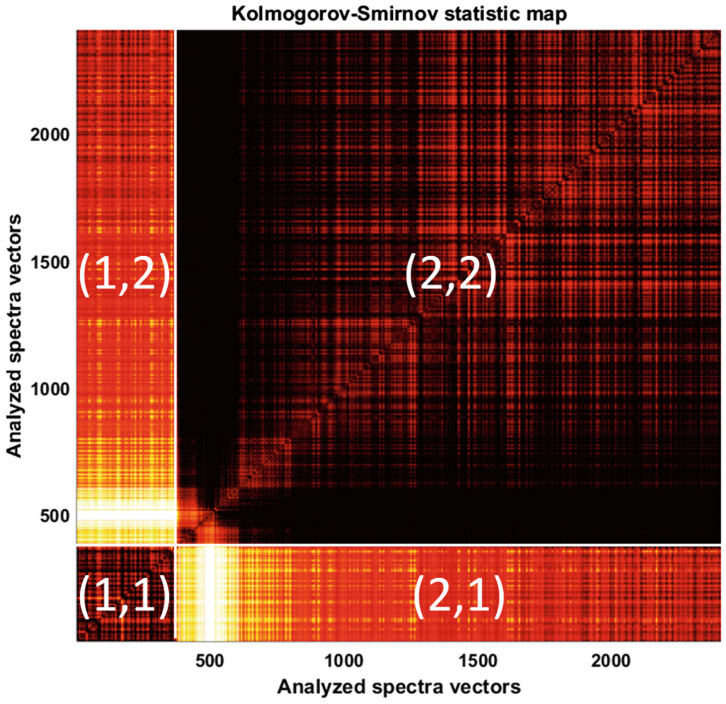


Fig. 5. The map of Kolmogorov-Smirnov statistic.

Hence, in the next step one-dimensional sum of the map is calculated (see Fig. 6). Since the processes share virtually no similarities, one can expect certain behavior of the sum values of particular groups of KS map columns:

- **Group regarding process 1 (columns of areas $(1, x)$):** relatively low energy of the signal. KS statistic values will be low when comparing process 1 to itself, high when the impact occurs and medium and decreasing while the energy damps;
- **Group regarding process 2 (columns of areas $(2, x)$):** relatively high energy. KS statistic values will be low when comparing process 2 to itself, and medium to high when comparing it to process 1.

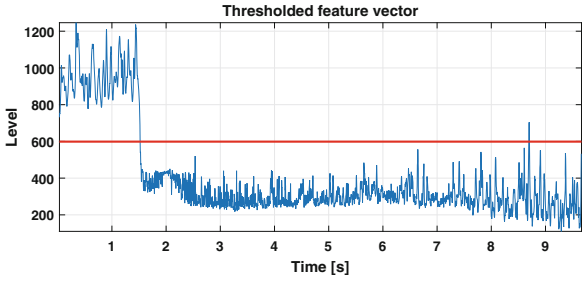


Fig. 6. Feature vector as a result of map integration

To be able to segment out the processes, obtained feature vector is thresholded using central point of histogram as described before (see Fig. 7). It provided the timestamp of 1.524 s, which matches the expert-indicated point exactly, and performs with better precision than STA/LTA method (comparing quotient of a signal under a specific characteristic function in short and long windows [1]) (Fig. 8).

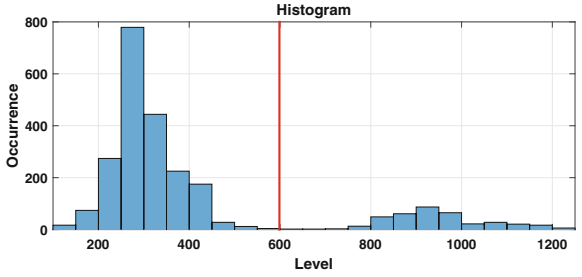


Fig. 7. Histogram of feature vector

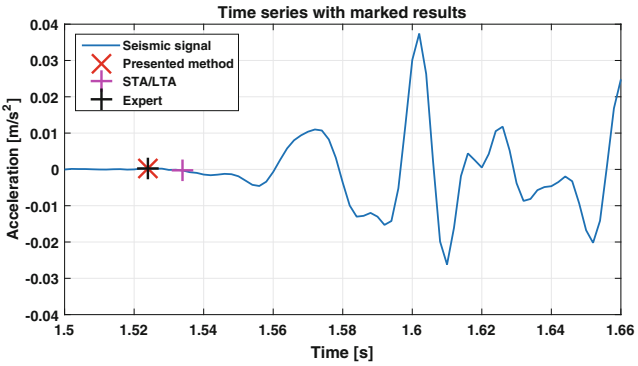


Fig. 8. Raw data with marked results of expert, STA/LTA and presented method. Obtained results: 1.524 s, 1.532 s, 1.524 s respectively.

4 Conclusion

In this article authors present novel approach to P-wave arrival detection in seismic vibration data. Method is based on segmentation of the feature vector constructed from KS statistic map. Entries of the map are KS statistic values that are results of performing KS test on pairs of spectra vectors of signal spectrogram. As a result, algorithm is capable of detection of the P-wave arrival. Presented method produces results consistent with the points indicated manually by seismic experts from the mine and better than commonly used LTA/STA algorithm. Important issue is that the method is automatic and data-driven.

References

1. Allen RV (1978) Automatic earthquake recognition and timing from single traces. Bull Seism Soc Am 68(5):1521–1532
2. Hafez AG, Khan TA, Kohda T (2009) Earthquake onset detection using spectro-ratio on multi-threshold time-frequency sub-band. Digit Signal Process 19(1):118–126
3. Kwiatek G, Ben-Zion Y (2013) Assessment of P and S wave energy radiated from very small shear-tensile seismic events in a deep south African mine. J Geophys Res Solid Earth 118(7):3630–3641
4. Leonard M (2000) Comparison of manual and automatic onset time picking. Bull Seism Soc Am 90(6):1384–1390
5. Leonard M, Kennett B (1999) Multi-component autoregressive techniques for the analysis of seismograms. Phys Earth Planet Inter 113(1):247–263
6. Massey FJ Jr (1951) The Kolmogorov-Smirnov test for goodness of fit. J Am Stat Assoc 46(253):68–78
7. Nurhaida, Subanar, Abdurakhman, Abadi AM (2017) Detecting P and S-wave of Mt. Rinjani seismic based on a locally stationary autoregressive (LSAR) model. In: AIP conference proceedings, vol 1868. AIP Publishing. https://doi.org/10.1063/1.4995121

8. Oppenheim AV (1999) Discrete-time signal processing. Pearson Education India, New Delhi
9. Polak M, Obuchowski J, Wyłomańska A, Zimroz R (2017) Seismic signal enhancement via AR filtering and spatial time-frequency denoising. In: Chaari F, Leskow J, Napolitano A, Zimroz R, Wylomanska A (eds) Cyclostationarity: theory and methods III. Springer, Cham, pp 51–68
10. Sleeman R, van Eck T (1999) Robust automatic P-phase picking: an on-line implementation in the analysis of broadband seismogram recordings. Phys. Earth Planet. Inter. 113(1):265–275
11. Sokolowski J, Obuchowski J, Zimroz R, Wylomanska A (2016) Comparison of recent P-wave arrival picking methods. In: 16th International Multidisciplinary Scientific GeoConference SGEM 2016. SGEM2016 Conference Proceedings, vol 2, pp 133–140
12. Sokolowski J, Obuchowski J, Zimroz R, Wylomanska A, Koziarz E (2016) Algorithm Indicating moment of P-wave arrival based on second-moment characteristic. Shock Vib **2016**. https://doi.org/10.1155/2016/4051701. Article ID 4051701
13. Wang J, Teng TL (1995) Artificial neural network-based seismic detector. Bull Seism Soc Am 85(1):308–319
14. Wang J, Tsang WW, Marsaglia G (2003) Evaluating Kolmogorov's distribution. J Stat Softw 8(18):1–4
15. Withers M, Aster R, Young C, Beiriger J, Harris M, Moore S, Trujillo J (1998) A comparison of select trigger algorithms for automated global seismic phase and event detection. Bull Seism Soc Am 88(1):95–106
16. Wodecki J, Stefaniak P, Michalak A, Wyłomańska A, Zimroz R (2018) Technical condition change detection using Anderson-Darling statistic approach for LHD machines – engine overheating problem. Int J Min Reclam Environ 32(6):392–400. https://doi.org/10.1080/17480930.2017.1388336
17. Xiantai G, Zhimin L, Na Q, Weidong J (2011) Adaptive picking of microseismic event arrival using a power spectrum envelope. Comput Geosci 37(2):158–164
18. Zhang H, Thurber C, Rowe C (2003) Automatic P-wave arrival detection and picking with multiscale wavelet analysis for single-component recordings. Bull Seism Soc Am 93(5):1904–1912
19. Zimroz R, Madziarz M, Żak G, Wyłomańska A, Obuchowski J (2015) Seismic signal segmentation procedure using time-frequency decomposition and statistical modelling. J Vibroengineering 17:3111–3121

Estimation of the Pointwise Hölder Exponent in Time Series Analysis

Adrianna Mastalerz-Kodzis(✉)

Department of Statistics, Econometrics and Mathematics,
University of Economics in Katowice, Katowice, Poland
adrianna.mastalerz-kodzis@ue.katowice.pl

Abstract. Stochastic processes are frequently used, among others, for empirical analyses in technical, medical as well as socio-economic sciences. A time series can be treated as a realization of a stochastic process with stationary or non-stationary increments. Using different quantitative methods we can study the stationarity of increments in the series and analyze the variability of their value. Stochastic processes which use Brownian motion are a frequent topic of theoretical analyses and empirical study. Among others, we model the variability of processes featuring stationary increments using Hurst exponent, whereas those featuring non-stationary increments using Hölder function. A characteristic feature of this type of processes is the analysis of memory present in a series. In hereby article the attention is focused on the estimation of pointwise Hölder exponents. The estimation of pointwise Hölder exponents allows for the analysis of the variability of a process in the surrounding of any argument of the domain. The aim of the article is to select the right surrounding in the process of estimation of the pointwise Hölder exponents for different analytical forms of generating function using the least squares method. The article is composed of two basic parts - theoretical and containing empirical analyses, in particular simulation and optimization analysis.

Keywords: Time series · Stochastic process · Hurst exponent · Hölder function · Estimation of pointwise Hölder exponents · Least squares method

1 Introduction

Multifractional Brown Motion Process can be used in modelling of time series. The process with stationary increments, which are featured by fractional parts of Brownian motion, depends on a constant parameter - Hurst exponent. This exponent belongs to the range (0,1) and divides the time series into: persistent - with a positive correlation between the subsequent implementations ($H \in (0.5, 1)$) and anti-persistent, in which the correlation is negative ($H \in (0, 0.5)$). A general case is considered below - processes dependent on Hölder function. Fractional processes are an exceptional example of multifractional ones, that is, a constant Hölder function is the value of Hurst exponent.

The estimation of pointwise Hölder exponents in Multifractional Brownian Motion Process is present in the papers (Ayache and Lévy Véhel 1999; Peltier and Lévy Véhel

F. Chaari et al. (Eds.): CSTA 2017, ACM 16, pp. 175–184, 2020.
https://doi.org/10.1007/978-3-030-22529-2_10

1995; Barrière 2007; Mastalerz-Kodzis 2018), but does not clearly define what surrounding is to be taken into consideration in the analyses. For the computer generated series we have proposed the methodic of the surrounding length selection. We have used, among others, the least squares method. The first part of the paper includes theoretical basics of the considered points, the second part contains simulation and optimization analyses (Bardet et al. 2003; Istas and Lang 1997; Guyon and León 1989).

2 Stochastic Processes with Memory Effect

Time is the only point in the process of time series modelling with fractional and multifractional processes of Brownian motion (Daoudi et al. 1998; Ayache and Taqqu 2004; Peters 1994; Falconer and Lévy-Véhel 2008; Mastalerz-Kodzis 2018). Processes of stationary kind featured by fractional parts of Brownian motion depend on a constant parameter - Hurst exponent. It is included in the range (0,1) and divides the time series into: persistent - with a positive correlation between the subsequent implementations ($H \in (1/2,\ 1)$) and anti-persistent, in which the correlation is negative ($H \in (0,\ 1/2)$). Below, we consider a general case - processes dependent on Hölder function. Fractional processes make an exceptional example of multi-fractional ones, this means that a constant Hölder function is the value of Hurst exponent.

2.1 Hölder Function Definition

Let (X, d_x) and (Y, d_Y) be metric spaces (Daoudi et al. 1998; Mastalerz-Kodzis 2003). Function $f : X \to Y$ is Hölder function with an exponent $\alpha(\alpha > 0)$, if for each $x, y \in X$ such, that $d_X(x, y) < 1$ the function fulfils an inequality with a positive constant c:

$$d_Y(f(x), f(y)) \leq c \cdot (d_X(x, y))^{\alpha}. \tag{1}$$

By definition, Hölder function is a continuous function in the range. With the function being of class C^1 the function graph's fractional value equals one. If the function were of class C^0, then the graph could feature a fractional measure.

Let there be a function $f : D \to \Re$ $(D \subset \Re)$ and parameter $\alpha \in (0, 1)$. Function $f : D \to \Re$ is Hölder function of class C^{α} $(f \in C^{\alpha})$, if there exist constants $c > 0$ and $h_0 > 0$ such, that for each x as well as all of h such, that $0 < h \leq h_0$ fulfilled is the following inequality:

$$|f(x+h) - f(x)| \leq c\, h^{\alpha}. \tag{2}$$

Let x_0 be any point from the function range f $(x_0 \in D \subset \Re)$. Function $f : D \to \Re$ is at point x_0 Hölder function of class $C^{\alpha}_{x_0}$ $\left(f \in C^{\alpha}_{x_0}\right)$, if constants $\varepsilon,\ c > 0$ exist such, that for each $x \in (x_0 - \varepsilon, x_0 + \varepsilon)$ fulfilled is the following inequality:

$$|f(x) - f(x_0)| \leq c|x - x_0|^a \tag{3}$$

Hölder point exponent of function f at point x_0 is the number $\alpha_f(x_0)$ expressed with the formula $\alpha_f(x_0) = \sup\left\{ \alpha : f \in C^{\alpha}_{x_0} \right\}$. For function f, Hölder function is the function, which to each of the points $x \in D$ assigns the number $\alpha_f(x)$.

2.2 Hölder Function Dependent Process

Let $H_t : [0, \infty) \rightarrow (0, 1)$ be Hölder function with exponent $\alpha > 0$. Multi-fractional Brownian motion process with function parameter H_t is a stochastic process $B_{H_t}(t)$ defined for $t \geq 0$ by the formula (Ayache and Lévy Véhel 1999; Peltier and Lévy Véhel 1995):

$$B_{Ht}(t) = \frac{1}{\Gamma(H_t + \frac{1}{2})} \left\{ \int_{-\infty}^{0} \left[(t-s)^{H_t - 1/2} - (-s)^{H_t - 1/2} + \right] dB(s) + \int_{0}^{t} (t-s)^{H_t - 1/2} dB(s) \right\} \quad (4)$$

where B stands for the standard Brownian motion process.

Hölder point exponents provide information about the characteristics of the process (Ayache and Lévy Véhel 1999; Daoudi et al. 1998). We can note, among others, that the process does not feature stationary increments in the situation where Hölder function is not constant and the variability of the graph increases with values of the function closer to zero, for function values close to one the process is smoother. The local capacitive and Hausdorff value of the process trajectory $B_{H_t}(t)$ for each $t_0 \geq 0$ equals $2 - H(t_0)$ and with probability equal to one Hölder point exponent of the process trajectory $B_{H_t}(t)$ for each $t_0 \geq 0$ equals $H(t_0)$. In the range (0,1) the regularity of the process measured with Hölder point exponents becomes more changeable.

Hölder function is a constant function in the multi-fractional Brownian motion process, which means that the regularity of the process trajectory measured by this function also changes continuously. Further generalization of Brownian motion process is based on the replacement of the continuous Hölder function with an discontinuous one (Peltier and Lévy Véhel 1995). A generalized multi-fractional Brownian motion process with function parameter $H(t)$ and λ - a real number, is process $\left\{ B_{H,\lambda}(t) \right\}_{t \in \Re}$ such, that for each $t \in \Re$:

$$B_{H,\lambda}(t) = \sum_{n=0}^{\infty} \int_{D_n} \frac{e^{it\xi} - 1}{|\xi|^{H_n(t) + 0{,}5}} dB(\xi) \quad (5)$$

where $D_0 = \{\xi : \ |\xi| < 1\}$ and for all $n \geq 1$, $D_n = \left\{ \xi : \ \lambda^{n-1} \leq |\xi| < \lambda^n \right\}$.

2.3 Random Relocation of the Segment Midpoint Method

The multi-fractional Brownian motion process can be generated with the use of random relocation of the segment midpoint method (Mastalerz-Kodzis 2003; Mastalerz-Kodzis 2016; Mastalerz-Kodzis 2018).

Interval [0,1] is divided at half point and we assign to it the following value:

$$B_{H(1/2)}(1/2) = \frac{B(0)+B(1)}{2} + G\frac{\sqrt{1-2^{2*H(1/2)-2}}}{2^{1*H(1/2)}} \tag{6}$$

where *B(0) = 0, B(1)* is equal to Gaussian pseudorandom number with an mean of 0 and variation of 1. *H(1/2)* is a set value, whereas *G* in the successive stages is a series of pseudorandom numbers which form an implementation of a variable of a normal distribution N(0,1).

Also at this stage, generally, the value of the process is set by the formula

$$B_{H(t)}(t) = \frac{B(t-d)+B(t+d)}{2} + G\frac{\sqrt{1-2^{2H(t)-2}}}{2^{i*H(t)}} \tag{7}$$

where $t - d$ and $t + d$ are the former points of time segment interval [0,1] (see Fig. 1).

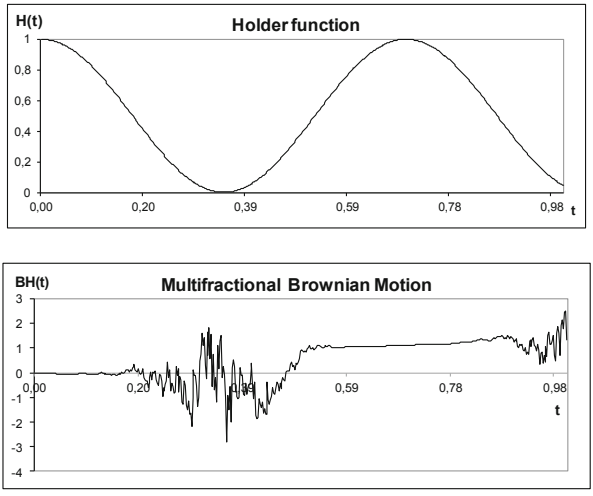


Fig. 1. Multifrational Brownian motion process: Hölder function in form $H(t) = 1 - sin^2(4.5t)$ and process simulation for the set function H(t).

2.4 Estimation Procedure

For series with stationary increments, Hurst exponent, constant in time, is calculated according to the rescaled range analysis (Peters 1994; Mastalerz-Kodzis 2003). With the formula included in (Peltier and Lévy Véhel 1995; Mastalerz-Kodzis 2018) we can also estimate pointwise Hölder exponents.

We will use the symbol $\{B_{i,n} = B_H(\frac{i}{n}), \quad 0 \le i \le n\}$ - to indicate Brownian motion process with Hurst exponent H. Let S_n be given by the formula $S_n = \frac{1}{n-1}\sum_{i=1}^{n-1} |B_{i+1,n} - B_{i,n}|$ and $H_n = -\frac{\log\left(\sqrt{\pi/2}\, S_n\right)}{\log(n-1)}$. Then $\lim_{n\to\infty} H_n = H$.

Let $1 < k < n$ be the length of the neighbourhood (range) used to estimate function $H(t)$. We are going to estimate a function for t from the range $[k/n, 1 - (k/n)]$.

Then the estimator $\hat{H}_{i/(n-1)}$ for $S_{k,n}(i) = \frac{m}{n-1}\sum_{j=i-k/2}^{i+k/2} |B_{j+1,n} - B_{j,n}|$ is as follows

$$\hat{H}_{i/(n-1)} = -\frac{\log\left(\sqrt{\pi/2}\, S_{k,n}(i)\right)}{\log(n-1)} \,. \tag{8}$$

The proper neighbourhood selection is not insignificant. A larger, than required, neighbourhood used in estimation (k is too high) means an excessively smooth Hölder function. For a lower k, on the other hand, the estimated function is very irregular. In the case when n is not a product of the whole numbers k, m, we can estimate the last (closer to present) n′ of the Hölder function value, for n′ with whole factors.

3 Empirical Study

The selection of a proper neighbourhood in the estimation procedure is extremely important. The paper (Peltier and Lévy Véhel 1995) does not clearly specify which criterion is to be followed in the selection of parameters k and m in the pointwise Hölder exponents estimation procedure.

With the use of the least squares method we considered the following function:

$$\sum \left(H(t) - \hat{H}_{i/(n-1)}\right)^2 \to \min \tag{9}$$

in order to define the estimate of Hölder function, the least different, in the mean square sense, to the initial generating function. We considered different divisions of the initial time series characterized by different surrounding lengths (different values of k and m).

For comparison, we also defined the sum: $\sum \left|H(t) - \hat{H}_{i/(n-1)}\right|$.

We considered a computer generated time series using the random relocation of the segment (513 elements long) midpoint method (formulas 6 and 7). It was transformed into a series of 513 feet of return. (Mastalerz-Kodzis 2003). The series was shortened by 1 element (the first, most distant from the present) in order to obtain a series of 512 feet of return. The number 512 can be divided by, inter alia: 128, 64, 32, 16, 8. We, therefore, analyzed five different surrounding lengths in the estimation procedure (the surroundings with the length of 4 and 2 were not taken into consideration as being too short. The questions asked during the analyses were as follows: which length of the surrounding provides a Hölder function closest to the initial one, does a universal method of selecting the surrounding length in the estimation process exist.

Different analytical forms of the generating Hölder function were considered in the empirical analyses (especially constant functions of values from the range (0, 1) – Hurst exponent, continuous functions, (linear, trigonometric) and discontinuous. For a set Hölder function a fractional (for a constant H), multi-fractional (for a continuous H(t) function) as well as a generalized multi-fractional Brownian motion process (for a discontinuous H(t)) was generated, and subsequently, for different surrounding lengths, the pointwise Hölder exponents were estimated. The results of the estimation for the two analyzed cases are presented below. The results of the analyses were presented in Fig. 1.

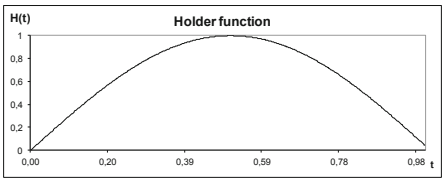
The longer the range used in estimation (the bigger the k), the smoother the estimate, the shorter the range (smaller k), the more 'ragged' the obtained function." The question is, which range should be accepted as optimal one? Whether in each one of the cases, for any form of analytical function $H(t)$ the same range length will, in the most advantageous way, reflect the initial generating function. It turns out that the answer is negative (Fig. 2).

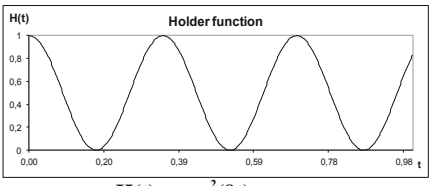
As the selection criterion in choosing the proper neighbourhood we accepted the minimum of the total of the squares of the estimated from the initial Hölder function values pointwise Hölder exponents' deviations (according to formula 9). Table 1 includes the results of the calculations. Considering the deviations' total, or its absolute value, obviously gave different results.

It turned out that, for a Hölder function $H(t) = sin(3,1t)$ the best match is made by the estimate where $k = 32$, whereby for a function of a trigonometric form $H(t) = cos^2(9t)$ the estimation which best reflects the initial function is the one conducted for $k = 8$. During the analyses we also noticed that:

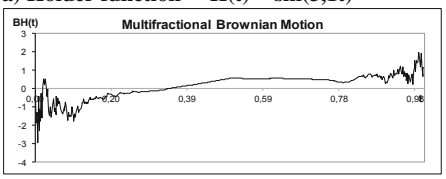
- For a constant Hölder function (Hurst exponent) the estimate that best matches is for the highest k (it smoothens random fluctuations);
- For Hölder function with a low number of extremes (as well as for a function with a low number of largest and smallest values) most often, the best match was made by the estimate for k = 32.
- The more differentiated the values of Hölder function, the more extreme values, the lower k used for estimation brings the generating function closer (captures its shape).

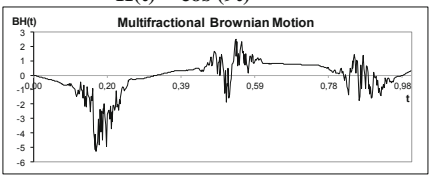
Further, for the function of the following form $H(t) = 0,9cos^2(5t) + 0,05$ the multi-fractional Brownian motion process was generated, the pointwise Hölder exponents were estimated and local standard deviation was added to/subtracted from them (calculated for the surrounding $k = 32$). The results have been presented in Fig. 3. It results from the graph, that the higher the pointwise Hölder exponents, the lower the local standard deviation, lower fractional dimension of the graph, lower variability. According to the pointwise Hölder exponents interpretation, with probability H(t), the series keeps its course, with probability H(t) after an increase in the value, another increase will take place, with probability 1 − H(t) a negative increment of the value will occur. For values close to 1, there is a slight risk of a change in the course, whereas for values close to 0 the change is very probable. (H(t) − s(t) and H(t) + s(t) deflect considerably from the estimated H(t)).



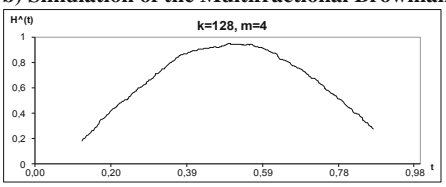


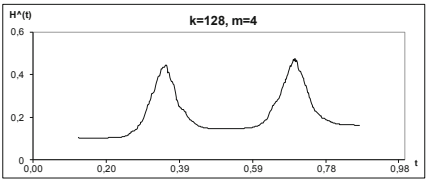
a) Hölder function H(t) = sin(3,1t) **$H(t) = \cos^2(9t)$**



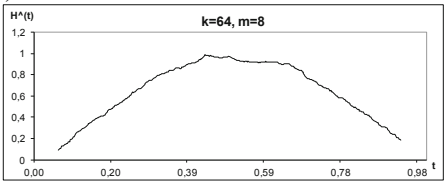


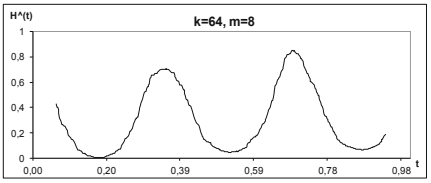
b) Simulation of the Multifractional Brownian Motion Process



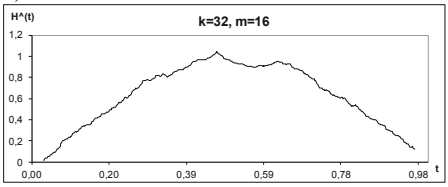


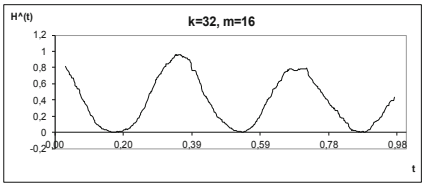
c) estimated Hölder function for k = 128



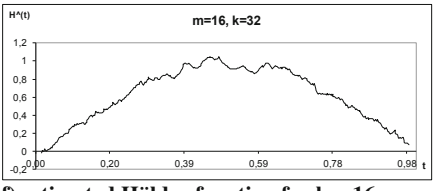


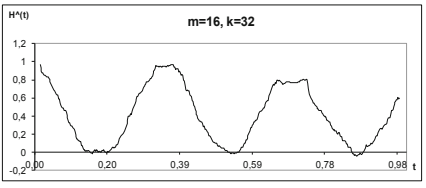
d) estimated Hölder function for k = 64



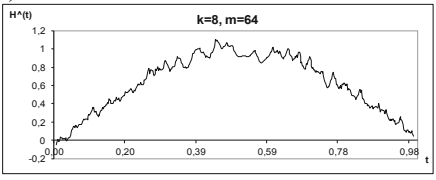


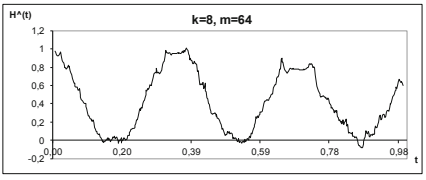
e) estimated Hölder function for k = 32





f) estimated Hölder function for k = 16





g) estimated Hölder function for k = 8

Fig. 2. Multifractional Brownian Motion Process

Table 1. The values of the analyzed sums for the estimated Hölder functions for different values of parameter k

Function ***H(t) = sin(3,1t)***	*K = 128*	*K = 64*	*K = 32*	*K = 16*	*K = 8*
$\sum \left(H(t) - \hat{H}_{i/(n-1)}\right)^2$	6,41	3,41	**2,99**	3,05	3,19
$\sum \left\|H(t) - \hat{H}_{i/(n-1)}\right\|$	48,17	32,83	27,68	27,83	29,71
Function ***H(t) = cos²(9t)***	*K = 128*	*K = 64*	*K = 32*	*K = 16*	*K = 8*
$\sum \left(H(t) - \hat{H}_{i/(n-1)}\right)^2$	54,12	16,64	6,01	3,98	**3,61**
$\sum \left\|H(t) - \hat{H}_{i/(n-1)}\right\|$	118,31	71,72	42,30	32,97	31,52

To conclude, the estimated pointwise Hölder exponents depend on the selection of the surrounding, which is taken into consideration in the estimation procedure. One of the methods of selecting an optimal surrounding is the least squares method. There is no universal or optimal neighbourhood for each form of analytical Hölder function, we must, each time, resolve an optimization problem as in (9).

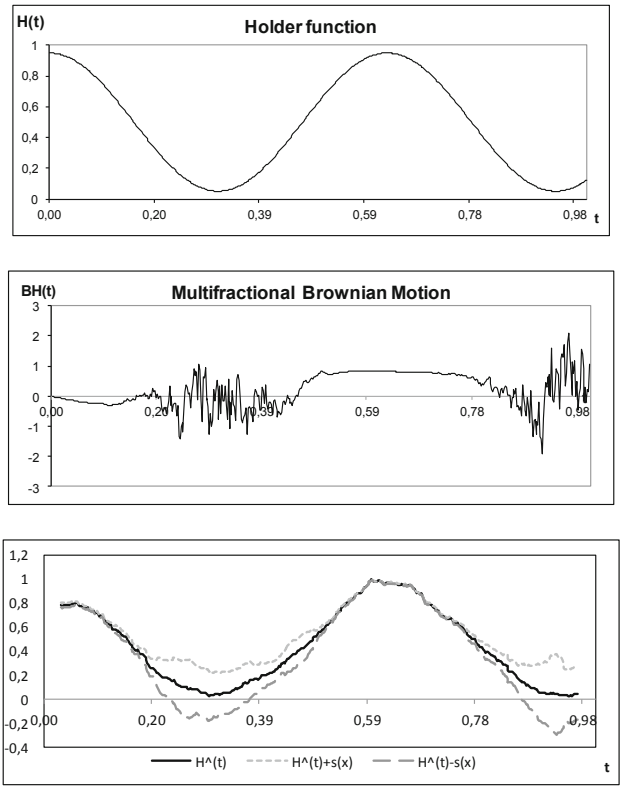


Fig. 3. Estimation of Hölder pointwise exponents with the consideration of local standard deviation for function $H(t) = 0{,}9\cos^2(5t) + 0{,}05$ and for $k = 32$.

The next step will be to use not only the simplest of methods, the least squares method, but also other optimization methods. It would also be important to develop a method for estimating the k parameter in empirical research, not only in computer simulations, when the analytic form of the Hölder function is unknown.

An interesting issue is also the study of cyclostationarity using the pointwise Hölder exponents. If the stochastic process has stationary increments, the value of the pointwise Hölder exponents is constant in time. The non-stationary increments of processes imply visible differences in the point values of the exponents.

4 Conclusion

The application of stochastic processes in empirical analyses has already been widely described in literature. Taking advantage of processes with a so called long memory in technical, medical and socio-economic sciences also is a frequent action. However, there still exist problems that should be continuously attended to, and the reason for that is their resolutions contribute to a more precise and valuable description of the surrounding us reality. And the estimation of pointwise Hölder exponents, according to the author, is one of them.

The article brings up the problem of the estimation of Hölder function, for it is often applied in numerous studies, both of time series, as well as spatial processes with the consideration of the memory effect. Based on the example of computer generated processes (computer simulations), considerations concerning the selection of neighbourhood in the pointwise Hölder exponents estimation processes were carried out. It turned out, that the selection of neighbourhood depends on the analytical form of the generating function, whereas one of the possible selection methods is the least squares method.

References

Ayache A, Lévy Véhel J (1999) Generalized multifractional brownian motion: definition and preliminary results. In: Dekking M, Lévy Véhel J, Lutton E, Tricot C (eds) Fractals: Theory and Applications in Engineering. Springer, London. https://doi.org/10.1007/978-1-4471-0873-3_2

Ayache A, Taqqu MS (2004) Multifractional processes with random exponent. Stochast Processes Appl 111(1):119–156

Bardet JM, Lang G, Oppenheim G, Philippe A (2003) Generators of long-range dependent processes: a survey. In: Doukhan P, Oppenheim G, Taqqu M (eds) Long-Range Dependence - Theory and Applications. Birkhauser, Basel, pp 579–623

Barrière O (2007) Synthèse et estimation de mouvements browniens multifractionnaires et autres processus à régularité prescrite, Définition du processus autorégulé multifractionnaire et applications. PhD thesis, IRCCyN

Daoudi K, Lévy VJ, Meyer Y (1998) Construction of continuous functions with prescribed local regularity. J Constr Approximations 014(03):349–385

Falconer KJ, Lévy-Véhel J (2008) Multifractional, multistable and other processes with prescribed local form. J Theor Probab 22:375–401. https://doi.org/10.1007/s10959-008-0147-9

Guyon X, León J (1989) Convergence en loi des H-variations d'un processus gaussien stationnaire sur R [Convergence in law of the H-variations of a stationary Gaussian process in R]. Ann Inst H Poincaré Probab Stat 25(3):265–282

Istas J, Lang G (1997) Quadratic variations and estimation of the local Hölder index of a Gaussian process. Ann Inst H Poincaré Probab Stat 33(4):407–436

Mastalerz-Kodzis A (2003) Modelowanie procesów na rynku kapitałowym za pomocą multifraktali, Prace Naukowe, Akademia Ekonomiczna im. Karola Adamieckiego w Katowicach, Katowice

Mastalerz-Kodzis A (2016) Algorytm generowania danych przestrzennych o zadanej lokalnej regularności, Mastalerz-Kodzis A., [w] Metody i modele analiz ilościowych w ekonomii i zarządzaniu. Cz.8; Wydawnictwo Uniwersytetu Ekonomicznego w Katowicach, Katowice 2016; p 11–34

Mastalerz-Kodzis A (2018) Application of the multifractional Brownian motion process in spatial analyses. Argumenta Oeconomica Cracoviensia (18):83–96

Peltier RF, Lévy Véhel J (1995) Multifractional Brownian Motion: Definition and Preliminary Results, INRIA Recquencourt, Rapport de recherche no. 2645

Peters EE (1994) Fractal Market Analysis. Wiley, New York

Application of the CIR Model for Spot Short Interest Rates Modelling on the Polish Market

Katarzyna Brzozowska-Rup(✉)

Kielce University of Technology, Kielce, Poland
krup@tu.kielce.pl

Abstract. The paper examines the estimation of the instantaneous Polish short term interest rate using one of the most popular stochastic differential models for studying the short interest rates, i.e. the Cox, Ingersoll, Ross model (1985) (henceforth CIR). We propose a new approach to estimating an instantaneous short interest rate: our attention is shifted from the whole term structure of the interest rate to the artificial notation of the short rate. In particular, the method focusing on determining a relationship between an observed instantaneous short interest rate and a certain (abstract) unobserved instantaneous rate which is defined as an interest rate demanded over an infinitesimally short period under the risk-neutral measure. To estimate the CIR model, we use a state space model in which estimates of the latent variables and model parameters are obtained by applying an Expectation-Maximisation algorithm combined with particle filters (PF). In practice, the instantaneous rate is identified with an overnight rate, therefore during the research we have adopted daily domestic interbank lending rates which are represented by interest rates on overnight deposits (WIBOR ON). To facilitate the discussion, simulated data are also employed. The obtained results prove the correctness and attractiveness of the method under consideration.

Keywords: Short (term) interest rates · Cox · Ingersoll and Ross model · Particle filter · Maximum likelihood estimation

1 Introduction

1.1 Background and Econometric Challenge

Interest rates are one of the most important categories of the national economy, (Dębski 2010). In the last three decades there has been a tremendous profusion of theories and modelling techniques of interest rate applicable to default-free bonds (commonly referred to risk-free bonds or Treasury bonds) and other interest rate derivatives. An important and useful concept in the modelling of interest rates is an instantaneous (also known as short) interest rate. This idea consists in an inherently discrete-time interest rate demanded over an extremely short period of time. In practice, the instantaneous interest rate does not exist, but this permits us to take an inherently discrete-time interest rate and gives it time-continuity and (additionally) to use the calculus of continuous-parameter stochastic processes in modelling them.

F. Chaari et al. (Eds.): CSTA 2017, ACM 16, pp. 185–203, 2020.
https://doi.org/10.1007/978-3-030-22529-2_11

Beyond doubt the best-known examples of short rates model are the Vasicek and the Cox, Ingersoll and Ross one-factor model. The CIR model is a Markov, equilibrium term structure model with the state variable (as a latent factor) following a square root process, which captures the main important features of real interest rates. Due to its features and capacity, it is one of the most frequently employed interest rate models in literature. The likelihood function for the CIR model is hard to compute: there is no closed-form expression for it, the model does not satisfy all the normality assumptions required for statistical consistency in the ML estimation. There are extensive studies available concerning estimation and implementation of this model, of which several worth mentioning are: Aït-Sahalia (1996), Chatterjee (2005), Kladìvko (2007), De Rossi (2010), Vo (2014).

The practical purpose of this paper is to explore an application of the one-factor CIR model to Polish short interest rate dynamics using an efficient numerical approximation. We propose a new perspective on estimating an instantaneous short interest rate. In our approach, attention is shifted from the whole term structure of interest rate to the artificial notation of the short rate. Rather than using yields-to-maturity computed from the CIR model, we derive a relation between an observed WIBOR ON rate (as a surrogate of a rate bond with an established maturity) and an unobserved short rate.

Therefore, using the yields on WIBOR rate as an input for the estimation process, we propose to estimate the CIR model by an EM algorithm combined with the particle filter.

1.2 Outline of the Paper

The remaining part of the paper is organized as follows: Sect. 2 introduces the definitions and notations used throughout the paper as well as the basic theory of the short-time interest rate. Next, we consider the one-factor interest rate model – CIR. Section 4 provides a brief discussion on the particle filter algorithm coupled with the maximum likelihood estimation method, in particular the Expectation-Maximisation algorithm, to create an iterative process for parameter estimation. Section 5 provides an empirical illustration of the proposed method to estimate the CIR model for a short-term interest rate quoted on the Polish market. The last one is the concluding section.

2 Interest Rates - Definition and Notation

In analysing the level and dynamics of the interest rate, a zero-coupon bond with maturity at time T (also called the T-zero-coupon bond or T-bond for short) is the fundamental notion that allows linking or setting the majority of interest rates.

Let $P(t,T)$ denote a T-maturity zero-coupon bond price at time $t<T$. A zero-coupon bond is a financial instrument defined as follows (Brigo and Mercurion 2001):

Definition. A zero-coupon bond is a contract that guarantees its holder the payment of one unit of currency at a maturity time T, with no intermediate payments. Clearly $P(T,T)=1$, for all T.

Definition. A continuously-compounded zero-coupon (spot) interest rate at time t for maturity T, denoted as $R(t,T)$ is the constant rate defined by the formula

$$P(t,T)\exp(\delta(T-t)R(t,T)) = 1, \tag{1}$$

from which continuously compounded rate R can be expressed as

$$R(t,T) = -\frac{lnP(t,T)}{\delta(T-t)} \tag{2}$$

Accordingly, the function $R : T \to R(t,T)$ is referred to as the (zero-coupon) yield curve. Term structure of interest rate is a functional dependence between the yield and the time to maturity $\delta(T-t) = T-t$ of a bond, where $\delta(T-t)$ is the time difference expressed in years (Brigo and Mercurio 2001). We accept a convention in which months and years are 30 days and 360 days long respectively.

Another important issue in terms of interest rates is instantaneous (short) interest rate $r(t) = r_t$.

Definition. Instantaneous interest rate r_t is the rate of growth of the value of a deposit started today and lasting for any infinitesimal time interval $[t, t+\Delta]$ (Weron and Weron 1999). When the maturity of the interest rate collapses towards its expiry, this leads to the following instantaneous interes rate

$$r_t = \lim_{T\to t^+} R(t,T) = -\frac{\partial}{\partial T} lnP(t,T)\Big|_{T=t} \tag{3}$$

This instantaneous interest rate is usually referred to as an instantaneous spot rate, or briefly, a short rate. The short rate is a mathematical abstract interest rate used for modelling interest rates. In practice, as a surrogate of the instantaneous rate, due to its duration, WIBOR rates are used.

It is worthwhile to explore the essence of creating a model for pricing interest rate derivatives. Quoting from Fabozzi (2002) "when we create a model for pricing interest derivatives, the "underlying" is not the price of a traded security, as it would be in a model for equity options. Instead, we specify a random process for the instantaneous, risk-free spot interest rate, the rate payable on an investment in default-free government bonds for a very short period of time. For convenience, we call this interest rate "the short rate".

The idea of a risk-free interest rate is consistent with the expectations theory of the term structure of interest rates. The theory deals with uncertain future and stresses the role of expectations of future short-term interest rates in the determination of the prices and yields on bonds. Actually, there is a variety of statements of this theory in the literature that differ in terms of the nature of the bond which is priced and the factors that enter into pricing. Broadly speaking, the expectations theory states that investing in

bonds as well as saving on a bank account gives the same (expected) profit. According to the expectations theory (Fobozzi 2002, Weron and Weron 1999):

$$P(t,T) = \mathrm{E}\left[\exp\left(-\int_t^T r_t^* dt\right)\middle|\mathcal{F}_t\right], \tag{4}$$

where r_t^* is the risk adjusted short rate defined as a random process for the short rate plus a function of the premium term. This is accomplished by redefining the model so that the spot rate is equal to expected return from investing over the same short period of time.

The market LIBOR rates L (and its Polish interbank term structure counterpart WIBOR rates)

$$L(t,T) = \frac{1 - P(t,T)}{r(t,T)P(t,T)} \tag{5}$$

are simply-compounded rates linked to the zero-coupon bond. It follows that zero-coupon bond prices can be expressed in terms of L as:

$$P(t,T) = \frac{1}{1 + \delta(t,T)L(t,T)}. \tag{6}$$

Finally, we can remark that the short-term interest rate can be defined as a limit of rates defined below:

$$r_t = \lim_{T \to t^+} R(t,T) = \lim_{T \to t^+} L(t,T). \tag{7}$$

3 The Cox Ingersoll and Ross Model in State Space Form

3.1 CIR Model

The Cox, Ingersoll and Ross model (1985) is one of the better known term structure models. The model generalizes the Vasicek model so that volatility is non constant and the dynamic of the underlying short-term interest rate is a diffusion process. The model links bond yields to one or more latent factors. In particular in the single factor case, the latent variable has an interpretation of the instantaneous spot rate r_t.

The one-factor time-homogeneous, equilibrium CIR model for the risk-free rate of interest is defined as the following stochastic differential equation (Cox et al. 1985)

$$dr_t = \kappa(\mu - r_t)dt + \sigma\sqrt{r_t}dW_t \tag{8}$$

with initial condition $r(0) = r_0$, W_t is a standard Brownian motion under the risk-neutral measure defined on a complete probability space $(\Omega, \mathcal{F}, P)$. The interest rate process $(r_t)_{t \geq 0}$ is usually known as the CIR process or square root process. If all parameters κ, μ, σ (from now on denoted as a vector $\theta = (\kappa, \mu, \sigma)$) are positive, and

$\sigma^2 < 2\kappa\mu$ holds, the CIR process is well defined and has a steady state (marginal) distribution. Gibson et al. (2001) showed that the distribution is a noncentral chi-square distribution of $(2q+2)$ degrees of freedom and the unconditioned parameter $2u$ and is given by

$$p_r(r_t|r_{t-1},\theta) = noncentral\,\chi^2((2q+2),2u) = ce^{-u-v}\left(\frac{v}{u}\right)^{0,5q} I_q\left(2\sqrt{uv}\right) \tag{9}$$

$$\text{where } c = \frac{2\kappa}{\sigma^2\left(1-e^{-\kappa\Delta}\right)}, u = cr_{t-1}e^{-\kappa\Delta}, v = cr_t, q = \frac{2\kappa\mu}{\sigma^2} - 1,$$

and

$$I_q(x) = \sum\nolimits_{n=0}^{\infty} \frac{(0,5x)^{2k+q}}{n!\Gamma(q+n+1)} \tag{10}$$

$I_q(\cdot)$ – is the modified Bessel function of the first kind of order q.[1] They also showed that the density of r_0 can be written as

$$p(r_0) = Gamma\left(c\left(1-e^{-k\Delta}\right), q+1\right) = \frac{\left[c\left(1-e^{-\kappa\Delta}\right)\right]^{q+1}}{\Gamma(q+1)} r_0^q e^{-r_0 c\left(1-e^{-\kappa\Delta}\right)} \tag{11}$$

The CIR process is ergodic and captures three important empirical features of short rate: mean reversion (interest rate tends to fluctuate over long-run trend μ where κ is the reverting speed), conditional heteroscedasticity (this property allows to capture the relationship between volatility (risk) and the level of interest rates) and non-negativity.

3.2 CIR Affine Models in State Space Form

At this point it is worth recalling that the CIR process belongs to the class of processes satisfying the "affine property". Affine term structure models (ATSMs) have become the standard in the literature on the valuation of fixed income securities, such as bonds, interest rate swaps and interest rate derivatives. Affine term structure models are constructed by assuming that the non-arbitrage price $P(r_t,t,T)$ at time t of a discount zero coupon bonds with maturity $T > 0$ and underlying interest rate dynamics described by CIR process, is given by (Cox et al. 1985; Christoffersen et al. 2014)

$$P(r_t,t,T) = P(r_t,\delta_t) = E\left(\exp\left(-\int_t^T r_s ds\right)|\mathcal{F}_t\right) = A(\delta_t)e^{-B(\delta_t)r_t}, \tag{12}$$

where $\delta_t = T-t, A(\cdot), B(\cdot)$ are known deterministic functions of δ_t, and parameters κ, μ, σ.

[1] It is worth noting that Bessel function approaches the plus infinity rapidly, which hinders the optimization routines. However the issue will not be addressed in this paper.

$$A(\delta_t) = \frac{2\kappa\mu}{\delta_t\sigma^2} ln\left(\frac{2\tau\ \exp(0,5(\kappa+\tau)\delta_t)}{(\kappa+\tau)(\exp(\tau\delta_t)-1)+2\tau}\right), B(\delta_t) = \frac{1}{\delta_t}\frac{2(\exp(\tau\delta_t)-1)}{(\kappa+\tau)\exp(\tau\delta_t)-1)+2\tau},$$
$$\tau = \sqrt{\kappa^2+2\sigma^2}$$

In practice, a short rate is seen as a hidden state and bond prices, or yields (e.g. the yield of one-month Treasury bills or money market rates) are used as observations. The yield to maturity at time t of a pure discount bond which expires at time T is defined as:

$$Y(t,T) = \frac{lnP(t,T)}{\delta_t} \tag{13}$$

One solution is to allow for discrepancies between the observed rates and the theoretical rates. These deviations can be explained by actual market features such as the rounding of prices, differences in the timing of observing financial variables and non-synchronous trading. In a modelling context, this can be done by assuming that the observed rates are affected by temporary shocks, which are Gaussian white noise errors. Therefore Eq. (13) which is treated as an exact relationship between factors and yields would now read as a linear function of the state variables r_t

$$y_t(\delta_t) = -A(\delta_t) + B(\delta_t)r_t + \eta_t. \tag{14}$$

It is necessary to note that an estimation of the term structure based on data from the Polish market Treasury bonds is not possible because there is a lack of liquidity of bonds and there are still not enough bonds with sufficient short term to maturity. Therefore we decided to use the interbank rates, i.e. WIBOR rates.

4 Parameter and State Space Estimation

4.1 General State Space Model

Our approach is based on the assumptions that the data (short rates) are not directly available, the observations are contaminated by measurement errors and the transition density does not admit an explicit form. These facts prevent the maximum likelihood estimation and imply that the problem under discussion can be perceived in a category of incomplete observations of the model with a tractable likelihood function.

To facilitate probabilistic reasoning over time, we will be adopting state space representations under a Bayesian framework. There are two main benefits of representing a dynamic system in a state space form. It allows unobserved variables to be incorporated into an observable model. Using a recursive filtering algorithm we can analyse the model and compute sample forecasts for a future value of state variable. Therefore, before starting our discussion of the particle filter we will define a general state-space model.

State space model consists of two equations: the state (transition) equation and the observation equation. Let $\{X_t\}_{t\geq 0}, \{Y_t\}_{t\geq 0}$ be stochastic processes defined on

a (measurable) space $(\Omega, \mathcal{F})$. The discrete-time process $\{X_t\}_{t\geq 0}$ is an unobserved (latent) or hidden Markov process of initial density $\mu(x|\theta)$, and Markov transition density $f(x'|x,\theta)$. The process $\{X_t\}_{t\geq 0}$ is partially observed through the observation process $\{Y_t\}_{t\geq 0}$. The observations $\{Y_t\}_{t\geq 0}$ are assumed to be conditionally independent given $\{X_t\}_{t\geq 0}$ and characterized by the conditional marginal density $g(y|x,\theta)$. To summarize, we have

$$X_0 \sim \mu(x|\theta) \tag{15}$$

$$X_t|X_{1:t-1} = x_{1:t-1}, Y_{1:t-1} = y_{1:t-1} \sim f(x_t|x_{t-1},\theta) \tag{16}$$

$$Y_t|(X_{1:t} = x_{1:t}, Y_{1:t-1} = y_{1:t-1}) \sim g(y|x_{t-1},\theta) \tag{17}$$

The system (15–17) defines a Bayesian model (also known as hidden Markov models, HHM), in which inference about $X_{1:n}$ given the realization $Y_{1:n} = y_{1:n}$ relies upon the posterior distribution (known as the posterior or target distribution in literature)

$$p(x_{0:n}|y_{1:n},\theta) = \frac{p(x_{0:n}, y_{1:n}|\theta)}{p(y_{1:n}|\theta)} = \frac{p(x_{0:n-1}|y_{1:n-1},\theta)g(y_n|x_n,\theta)f(x_n|x_{n-1},\theta)}{p(y_n|y_{1:n},\theta)}. \tag{18}$$

This class includes many non-linear and non-Gaussian time series models.

It is worth noting that in the literature the filtering problem is defined as the recursion satisfied by the marginal distribution $p(x_t|y_{1:t},\theta)$

$$p(x_t|y_{1:t},\theta) = \frac{g(y_t|x_t,\theta)f(x_t|y_{1:t-1},\theta)}{p(y_t|y_{1:t-1},\theta)}, \tag{19}$$

where

$$p(x_t|y_{1:t-1},\theta) = \int p(x_{t-1}|y_{1:t-1},\theta)p(x_t|x_{t-1},\theta)dx_{t-1}, \tag{20}$$

$$p(y_t|y_{1:t-1},\theta) = \int p(x_{t-1}|y_{1:t-1},\theta)p(y_t|x_t,\theta)p(x_t|x_{t-1},\theta)dx_{t-1:t}, \tag{21}$$

From the Bayesian perspective the posterior probability density function (PDF) contains all the state information both from the system model and from the observations. It means that knowing posterior PDF gives an opportunity to derive many well-known estimators, such as minimum-mean-square-error estimators, maximum likelihood estimators and maximum posterior estimators.

In linear Gaussian models, the problem of estimation is solved using the Kalman filter. Unfortunately, when the model is nonlinear and/or non-Gaussian, the Kalman filter, as well as its modifications (e.g. the Extended Kalman Filter, Unscented Kalman Filter) can only be used for initial approximation.

4.2 The Idea of Particle Filter

Particle filters have become a popular class of numerical methods for inference in non-linear non-Gaussian scenarios. In particular particle filtering can be reinterpreted as some special instances of Sequential Monte Carlo (SMC) method.

The method has been an active area of research since the original work of Gordon et al. (1993). There is a large body of literature on the theory as well as a number of successful very extensive and varied applications described in literature. However, there has been no comprehensive review so far. Taking into account the direction and topic of the book, it is worth paying attention to the implementation of PF algorithms in the following areas, such as signal processing: positioning, navigation, target tracking (see, for instance Gustafsson et al. 2002, Nordlund and Gustafsson 2008); in the field of robotics and automatics (especially particle filter algorithm known as FastSLAM, Chen 2012), statistics (Doucet 2008) and communication (Djurić et al. 2003, Ghirmai 2016). The interested reader can also find comprehensive and in-depth expositions of the different algorithms and extensive lists of references on the subject in Doucet et al. 2001, Cappé et al. 2007 and Kantas et al. 2014.

The basic idea of the PF consists in estimating target probability distributions by empirical distributions focused on a set of samples (called particles) determined by the sequential importance sampling (SIS) extended by the technique of resampling (re-weight the particles using importance weights so that the correct distribution is targeted). The specificity of PF approximation relies on sampling from a known importance density $q_\theta(x_t|x_{t-1}, y_{1:t}, \theta)$ (also known as proposal density) instead of sampling from the transition density $f(x_t|x_{t-1}, \theta)$ which may frequently be daunting. A more detailed discussion of the choice of importance density can be found in (Doucet et al. 2001; Cappe 2007; Brzozowska-Rup and Dawidowicz 2009).

In a standard PF (also known as the sequential importance resampling algorithm), the posterior PDF $p(x_t|y_{1:t}, \theta)$ is approximated by a set of N particles as

$$\hat{p}_N(x_t|y_{1:t}) = \sum_{i=1}^{N} w_t^{(i)} \delta(x_t - x_t^{(i)}) \tag{22}$$

Where δ denotes the Dirac delta function, $x_t^{(i)}$ denotes the i-th particle with importance weight $w_t^{(i)} = W\left(x_t^{(i)}\right) / \sum_{j=1}^{N} W\left(x_t^{(j)}\right)$, $\sum_{i=1}^{N} w_t^{(i)} = 1, w_t^{(i)} \geq 0$ for $i = 1, \ldots, N$. We define the importance weights

$$w_t^{(i)} = \frac{p(x_{1:t}^{(i)}|y_{1:t}, \theta)}{q\left(x_{1:t}^{(i)}\right)} = w_{t-1}^{(i)}\left(x_{1:t-1}^{(i)}\right) \frac{g\left(y_t \middle| x_t^{(i)}, \theta\right) f\left(x_t^{(i)} \middle| x_{t-1}^{(i)}, \theta\right)}{q\left(x_t^{(i)} \middle| x_{t-1}^{(i)}, y_t\right)} \tag{23}$$

Given current particles set $\{x_t^{(i)}, w_t^{(i)}\}_{i=1}^{N}$ from $\hat{p}_N(x_t|y_{1:t})$ the particles are generated from the algorithm consists of two steps:

1. For $i = 1, \ldots, N$ simulate $x_{t+1}^{(i)} \sim q_t(x_{t+1}|x_t, y_{t+1})$,

2. For $i = 1, \ldots, N$ using (16) compute weights $w_{t+1}^{(i)}$ compute Effective Sample Size N_{ESS}, $N_{ESS} = \left(\sum_{i=1}^{N} \left(w_t^{(i)}\right)^2\right)^{-1}$

   If $\widehat{N}_{ESS} < N_T$, then *RESAMPLE* :

$$\text{Drawz } (i) \sim Mult\left(N, w_t^{(1)}, \ldots, w_t^{(N)}\right)$$

$$\text{update } \left\{x_{0:t+1}^{(i)}, w_{t+1}^{(i)}\right\}_{i=1}^{N} = \left\{x_{0:t+1}^{(z(i))}, \frac{1}{N}\right\}_{i=1}^{N}$$

   Else keep $\left\{x_{0:t+1}^{(i)}, w_{t+1}^{(i)}\right\}_{i=1}^{N}$

   $\left\{x_{0:t+1}^{(i)}, w_{t+1}^{(i)}\right\}_{i=1}^{N}$ approximates the filtering density at time $t+1$.

Specialist literature features different implementation schemes of resampling, for example: systematic, residual, residual-systematic resampling. Review and comparison of these techniques can be found in Hol et al. (2006). During the last twenty years, in parallel with algorithmic developments, the theoretical properties of the method have been studied extensively, and there is currently a number of available results describing the convergence and difficulties of PF estimation, e.g. the monographs (Del Moral 2004; 2013; Chopin 2004; Douc and Moulines 2008).

4.3 Forward Filtering-Backward Smoothing

It is known that sampling from a joint distribution $p(x_{1:T}|y_{1:T})$ for approximation marginals $\{p(x_t|y_{1:T}, \theta)\}_{t=1}^{T}$, can be inefficient when T is large and $t \ll T$ (this is the result of the successive resampling step which causes particle degeneracy). In order to overcome these difficulties, particle smoothing algorithm is recommended (Doucet and Johansen 2008).

In this article we will use the concept of the forward-backward smoothing formula, in which particle smoothing is performed recursively backward in time according to the formula:

$$\begin{aligned} p(x_t|y_{1:T}, \theta) &= \int p(x_t, x_{t+1}|y_{1:T}, \theta) dx_{t+1} \\ &= p(x_t|y_{1:t}, \theta) \int \frac{f(x_t|x_{t+1}, \theta)}{p(x_{t+1}|y_{1:t}, \theta)} p(x_{t+1}|y_{1:T}, \theta) dx_{t+1} \end{aligned} \tag{24}$$

The preceding expression reveals that the marginal smoothing posterior density can be approximated by the following expression:

$$\hat{p}(x_t|y_{1:T}, \theta) = \sum_{i=1}^{N} w_{t|T}^{(i)} \delta\left(x_t - x_t^{(i)}\right) \tag{25}$$

where smoothing weights are defined by

$$w_{t|T}^{(i)} = w_t^{(i)} \sum_{j=1}^{N} w_{t+1|T}^{(j)} \frac{f\left(x_{t+1}^{(j)} \middle| x_t^{(i)}, \theta\right)}{\sum_{k=1}^{N} w_t^{(k)} p\left(x_{t+1|T}^{(j)} \middle| x_t^{(k)}, \theta\right)}, \tag{26}$$

For a detailed treatment on the FFBS algorithm we suggest readers refer (Briers et al. 2010 and Zhou 2013)

4.4 CIR as a State Space Model Representation

Estimation of the unobservable factors r_t within an exact state space model is not simple because of the highly non-normal transition density $p(r_t|r_{t-1}, \theta)$. Therefore it is necessary to determine its approximation form.

In our approach we use the fact that yields data are only seen at an equally spaced discrete time, rather than diffusion form (8) so it would be more convenient to consider observations for state variables at discrete time intervals with the size Δ. Therefore using Euler discretization[2] for CIR model, the transition dynamic can be written as follows:

$$\begin{aligned} f(r_t|r_{t-1}, \theta) &= \frac{1}{\sqrt{2\pi\Delta\sigma^2|r_{t-1}|}} \exp\left(-\frac{(r_t - r_{t-1} - \kappa(\mu - r_{t-1})\Delta)^2}{2\sigma^2\Delta|r_{t-1}|}\right) \\ &= \frac{1}{\sqrt{2\pi\Delta\sigma^2|r_{t-1}|}} \exp\left(-\frac{(r_t - \kappa\mu\Delta + (\kappa\Delta - 1)r_{t-1})^2}{2\sigma^2\Delta|r_{t-1}|}\right), \end{aligned} \tag{27}$$

We assume that the measurement error $\in_t$ is IID Gaussian, and there is no serial correlation then the observations conditional density, given by the pricing function (14) can be derived as:

$$g(y_t(\delta_i)|r_t, \theta) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp\left(-\frac{(y_t(\delta_i) + A(\delta_i) - B(\delta_i)r_t)^2}{2\sigma_\eta^2}\right) \tag{28}$$

From now on Eqs. (27), (28) constitute the state space model under consideration. We can use the particle filter algorithm to obtain information about r_t conditional on the observation of simply-compounded WIBOR rates y_t.

[2] We use the Euler discretization with correction, known as "the full truncation scheme" studied in (Higham et al. 2002).

4.5 Expectation-Maximisation Algorithm

The Expectation-Maximisation (EM) algorithm (Dempster et al. 1977) is useful in situations where the observations can be viewed as incomplete data and/or the direct maximisation of the likelihood is more complex than the maximisation of the complete likelihood:

$$\ell_c(\theta|x_{0:T}, y_{1:T}) = logp(x_{0:T}, y_{1:T}|\theta), \tag{29}$$

The method consists of two steps recursively executed so long, until the criterion of convergence is reached (Cappé 2011, Olsson 2008). The states of model are not observed so the first (E-step) approximates the joint log-likelihood function by taking the expected value over the unobserved states based upon some current estimation of the parameters θ_k

$$\begin{aligned} Q(\theta, \theta_k) &= E[\log p(x_{0:T}, y_{1:T}|\theta)|y_{1:T}, \theta_k] \\ &= \int \log p(x_{0:T}, y_{1:T}|\theta) p(x_{0:T}|y_{1:T}, \theta_k) dx_{0:T} \end{aligned} \tag{30}$$

The (M-step) then maximises the above expectation with respect to the parameter θ

$$\theta_{k+1} := \text{argmax}_\theta \, Q(\theta, \theta_k) \tag{31}$$

In the offline version of EM, based on the Bayesian rule, and assuming that the importance weights are known, the desired approximation of the function Q takes the form of:

$$\begin{aligned} \widehat{Q}(\theta, \theta_k) = &\sum_{i=1}^{N} logp\left(x_{0|T}^{(i)}|y_{1:T}, \theta\right) w_{0|T}^{(i)} + \sum_{t=1}^{T} \sum_{i=1}^{N} logp\left(x_{t|T}^{(i)} \middle| x_{t-1|T}^{(i)}, \theta\right) w_{t|T}^{(i)} \\ &+ \sum_{t=1}^{T} \sum_{i=1}^{N} logp\left(y_t \middle| x_{t|T}^{(i)}, \theta\right) w_{t|T}^{(i)}, \end{aligned} \tag{32}$$

where the weights $w_t^{(i)}, w_{t|T}^{(i)}$ are determined respectively during the filtration and smoothing procedures of formulas (23) and (26) (Cappé 2011, Olsson et al. 2008, Schön et al. 2011).

4.6 Particle Filtering Backward Smoothing PFBS-EM Algorithm

The discussed EM and particle smoothing estimation method can be summarised in the following algorithm.

PFBS-EM algorithm

I. Assume a started parameter value θ_0
For $k = 0, \ldots K$
For $t = 0, \ldots T$,
For $i = 1, \ldots N$
run the PF (algorithm 1) obtain $\left\{x_{1:t|t}^{(i)}, w_t^{(i)}\right\}$,
compute $SW_i = \sum_{j=1}^{N} w_t^{(j)} p(x_{t+1}^{(i)} \left| x_t^{(j)}, \theta \right.)$
End for
End for
II. For $i = 1, \ldots, N$ $w_{T|T}^{(i)} = w_T^{(i)}$ end for
For $t = T-1, \ldots, 1$, for $i = 1, \ldots, N$ using (27) compute weights $w_{t|T}^{(i)}$
If necessary $Resample\left\{x_{t|T}^{(i)}, w_{t|T}^{(i)}\right\} \rightarrow \left\{\tilde{x}_{t|T}^{(i)}, \frac{1}{N}\right\}$
Else set $\tilde{x}_{t|T}^{(i)} = x_{t|T}^{(i)}$.
End for
IIII Apply EM and update the estimation of θ.

Set k=k+1 End for
Set $\hat{\theta} = \hat{\theta}_k$

It is worth noting that in practice it is proposed to perform only step E for the first 10–20 observations. A close analysis of the approximation of EM algorithm using the FFBS technique was presented in (Olsson et al. 2008; Douc et al. 2011; Zhou 2013)

5 Empirical Analysis and Application

5.1 Estimation of the CIR

In this section, through simulation and real data application, we present performance of the method that has been discussed so far.

We proposed to consider the conditional optimal proposal distribution, which for a fixed δ_i can be approximated (up to constant) by

$$\begin{aligned} q(r_t|r_{t-1},y_t,\theta) &\propto f(x_t|x_{t-1},\theta)g(y_t|x_t,\theta) \\ &\propto \frac{1}{\sigma\sigma_\eta\sqrt{\Delta|r_{t-1}|}}\exp\left(-\frac{(r_t-(\kappa\mu\Delta-(\kappa\Delta-1)r_{t-1}))^2}{2\sigma^2\Delta|r_{t-1}|}\right) \\ &\quad\times\exp\left(-\frac{(B(\delta_i)r_t-(y_t(\delta_i)+A(\delta_i)))^2}{2\sigma_\eta^2}\right) \\ &\quad\propto N(r_t;\Sigma m,\Sigma) \end{aligned} \tag{33}$$

$$m=\frac{B(\delta_i)(y_t(\delta_i)+A(\delta_i))}{\sigma_\eta^2}+\frac{\kappa\mu\Delta-(\kappa\Delta-1)r_{t-1}}{\sigma^2\Delta|r_{t-1}|} \tag{34}$$

$$\Sigma^{-1}=\frac{B^2(\delta_i)}{\sigma_\eta^2}+\frac{1}{\sigma^2\Delta|r_{t-1}|} \tag{35}$$

and incremental importance weight then becomes

$$w_t=\frac{f(r_t|r_{t-1},\theta)g(y_t|r_t)}{q(r_t|r_{t-1},y_t,\theta)} \tag{36}$$

The joint likelihood function with the accuracy to multiplicative constant has the following formula

$$\begin{aligned} \ell_c(\theta)=-\Bigg\{&\left(\frac{2\kappa_0\mu_0}{\sigma_0^2}\right)log\left(\frac{\left[c_0\left(1-e^{-\kappa_0\Delta}\right)\right]}{\Gamma\left(\frac{2\kappa_0\mu_0}{\sigma_0^2}\right)}\right)+\frac{2\kappa_0\mu_0}{\sigma_0^2}log(r_0)-r_0c_0\left(1-e^{-\kappa_0\Delta}\right) \\ &+\frac{1}{2}\left(\sum_{t=1}^{T}\log\left(2\pi\Delta\sigma^2|r_{t-1}|\right)+\frac{(r_t-\kappa\mu\Delta+(\kappa\Delta-1)r_{t-1})^2}{\sigma^2|r_{t-1}|\Delta}+\left(\frac{y_t(\delta_t)+A-Br_t}{\sigma_\eta}\right)^2\right. \\ &\left.+Tlog\left(2\pi\sigma_y\right)\right)\Bigg\} \end{aligned} \tag{37}$$

For further calculations we will use the modified form of the above function, where the change consists in omitting the distribution of the initial state, the normalization constant and the last component (they do not depend on the parameters sought).

$$\begin{aligned} -2Q(\theta,\theta_k)=E\Bigg[&\sum_{t=1}^{T}\left(log\left(2\pi\Delta\sigma^2|r_{t-1}|\right)+\frac{(r_t-\kappa\mu\Delta+(\kappa\Delta-1)r_{t-1})^2}{\sigma^2|r_{t-1}|\Delta}\right. \\ &\left.+\left(\frac{y_t(\delta_t)+A-Br_t}{\sigma_y}\right)^2\right)+Tlog\left(2\pi\sigma_\eta^2\right)|y_{1:T},\theta_k\Bigg] \end{aligned} \tag{38}$$

The estimates are given by the following formulas:

$$\hat{\sigma}^2 = \frac{1}{\Delta \mathrm{T}} \mathrm{E}\left[\frac{\mathrm{S}_0(\mathrm{S}_3^2 - \mathrm{S}_4^2) + \mathrm{S}_0\mathrm{S}_5(\mathrm{S}_4 - \mathrm{S}_5) + \mathrm{S}_1(\mathrm{S}_1\mathrm{S}_5 - 2\mathrm{S}_2\mathrm{S}_3) + S_2^2 S_4}{\mathrm{S}_0\mathrm{S}_4 - \mathrm{S}_1^2} | y_{1:T}, \theta_k\right]$$

$$\hat{\sigma}_\eta^2 = \frac{1}{T} E[S_6 | y_{1:T}, \theta_k],$$

$$\hat{\kappa} = E\left[-\frac{S_0(S_3 - S_5) + S_1(S_1 - S_2)}{(S_0 S_4 - \mathrm{S}_1^2)\Delta} | y_{1:T}, \theta_k\right],$$

$$\hat{\mu} = \mathrm{E}\left[\frac{\mathrm{S}_1(\mathrm{S}_3 + \mathrm{S}_4) - \mathrm{S}_1\mathrm{S}_5 - \mathrm{S}_2\mathrm{S}_4}{\mathrm{S}_0(\mathrm{S}_3 - \mathrm{S}_5) + \mathrm{S}_1(\mathrm{S}_1 - \mathrm{S}_2)} | \mathrm{y}_{1:\mathrm{T}}, \theta_\mathrm{k}\right].$$

Where $S = (S_0, S_1, \ldots, S_6)$ admits the following sufficient statistics:

$$S_0 = \sum_{t=1}^{T} \frac{1}{|r_{t-1}|}, S_1 = \sum_{t=1}^{T} \frac{r_{t-1}}{r_{t-1}}, S_2 = \sum_{t=1}^{T} \frac{r_t}{|r_{t-1}|}, S_3 = \sum_{t=1}^{T} \frac{r_t r_{t-1}}{|r_{t-1}|},$$
$$S_3 = \sum_{t=1}^{T} \frac{r_t r_{t-1}}{|r_{t-1}|}, S_4 = \sum_{t=1}^{T} \frac{r_{t-1}^2}{|r_{t-1}|}, S_5 = \sum_{t=1}^{T} \frac{r_t^2}{|r_{t-1}|}$$
$$S_6 = \sum_{t=1}^{T} \left(y(\delta_t) + \hat{A} - \hat{B} r_t\right)^2,$$

5.2 Data Description

To illustrate the estimation method, we rely on both simulated and real data, in particular we apply the daily observations of WIBOR ON (from January 2015 to October 2017). The data are collected from the most popular finance and business website https://www.money.pl. The total number of observations T = 719. In the estimation, the number of particles $N = 500$, the EM iteration number is 200, while the first 20 iterations are not taken into account.

To evaluate the estimation accuracy, we use the RMSE which is defined as the differences between the estimated observation yH_t and the real observation $y0_t$ as follows:

$$RMSE = \left(\frac{1}{T}\sum_{t=1}^{T} (y0_t - yH_t)^2\right)^{\frac{1}{2}}, \tag{39}$$

where $yH_t(\delta_t) = -A(\delta_t) + B(\delta_t)\widehat{r}_t, \quad \delta_t = 1$ and $\widehat{r}_t$ is the estimate of the state r_t which is the mean state of all the particles in time t.

To confirm the effectiveness of the presented method we simulate data from the discretized form of CIR model with parameters: $\kappa = 0,6$, $\mu = 0,7$, $\sigma = 0,05$, $A(\delta_t) = 0,1$, $B(\delta_t) = 1$, $\sigma_\eta = 0,05$, $\Delta = 1$. As the initial parameters for EM method we use

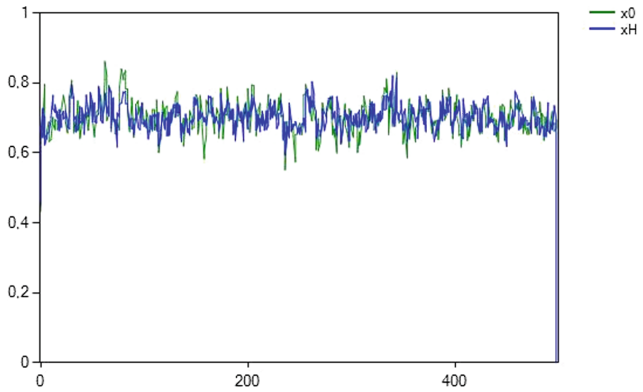


Fig. 1. Time series of state variable generated from CIR model with initial parameter (green solid line) and its filtered value obtained by PF (blue line).

parameters $\kappa^{(0)} = 0,06$, $\mu^{(0)} = 0,1$, $\sigma^{(0)} = 0,5$, $A(\delta_t) = 0,1$, $B(\delta_t) = 1, \delta_t = 1$, $\sigma_\eta^{(0)} = 0,5$, the T = 500, N = 200 number of particles and 100 iterations of PFFBS-EM algorithm. The results of particle filtering are given in Fig. 1. We can see that despite the limited number of particles, the estimated trajectory ideally coincides with the simulated data.

Figure 2 shows the convergence through iteration of estimates for each parameter. The 100^{th} estimates and their standard deviation are $\widehat{\kappa} = 1,657$, $\widehat{\mu} = 0,685$, $\widehat{\sigma} = 0,085$, $\widehat{\sigma}_\eta = 0,051$.

The estimators very quickly stabilize at the right level except for the parameter κ. This fact will be the goal of the future research.

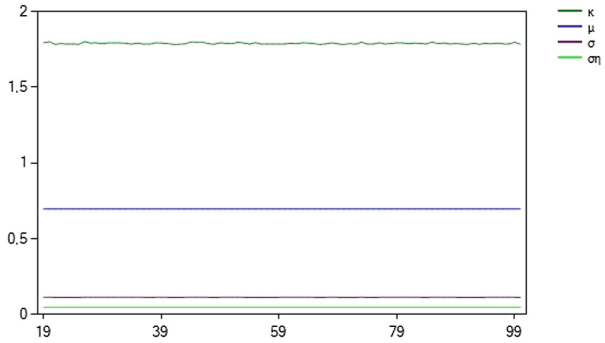


Fig. 2. Convergence of the PFBS-EM estimates of the CIR model parameter for simulated data

5.3 Estimating CIR Model for Polish Short Interest Rate

We have carried out a number of simulations assuming various initial conditions, however, due to the limitations of the article, we chose to present one particular case

result, which we selected as optimal on the basis of RMSE criterion. The data presented in the Table 1 are averaged results of the estimation from 20 iterations of the complete PFBS-EM algorithm.

Table 1. Estimation of CIR model from daily WIBOR ON rate, 2015-2017.

Parameters	Initial value of parameters for EM	Mean value of parameter estimators	Standard deviation
κ	0,501	0,483	0,023
μ	0,800	2,521	0,240
σ	0,250	0,201	0,011
σ_η	0,250	0,141	0,021

In Figs. 3 and 4 we compare the WIBOR ON rate and the CIR-estimated rate, where Fig. 4 is an excerpt from the whole trajectory in order to illustrate the resemblance of both rates. This proves the efficiency of the applied method (Fig. 5).

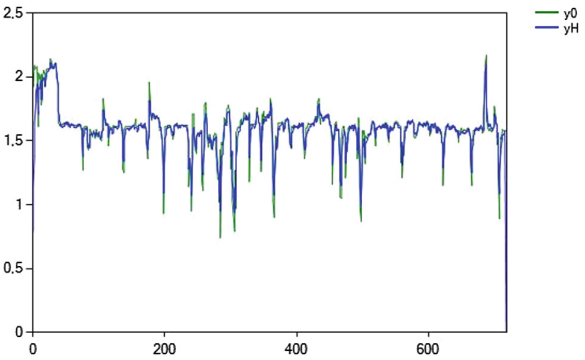


Fig. 3. The daily WIBOR ON rate from January 2015 to October 2017 vs estimated rate obtained by the presented method with initial parameters from the table

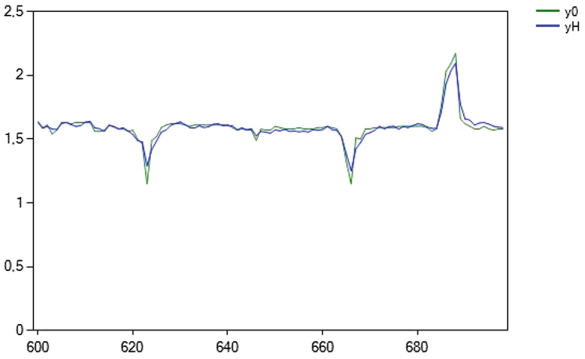


Fig. 4. The graph is a selected excerpt from Fig. 3

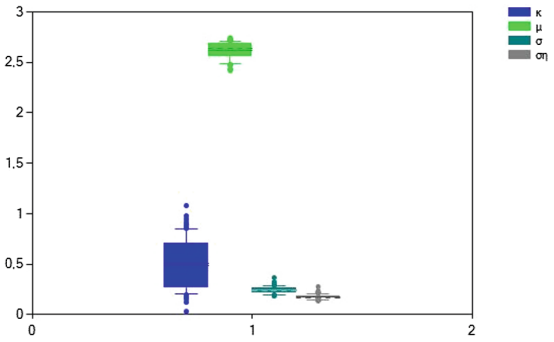


Fig. 5. Boxplots of estimates of the CIR parameters produced by PFBS-EM algorithm for Polish short interest rate

All the numerical results obtained in this section have been produced in the author's original program using C Sharp.

6 Concluding Remarks

The models of term structure of interest rates are probably the most computationally difficult part of the modern finance due to the relative complexity of application techniques. The knowledge of the CIR model estimators, the current interest rate y_t and the short-term interest rate r_t gives an opportunity to estimate the expected future value of the short-term interest rate, which are very important for the whole economy of the country. This study aims to test the feasibility of using WIBOR ON rates to model short interest rates for the Polish market. To achieve this goal, we have used a state space model to estimate the CIR of term structure of interest rates. The paper demonstrates the attractiveness of PFBS-EM estimation in the state variable combined with EM estimation of unknown parameter estimation for the analysed model. The advantage of the presented method is that it is easy to implement and the fact that, in contrast to the popular techniques of estimating CIR model, reduces the complexity of calculation effort and simultaneously returns reliable estimators of unknown parameters. However, it has been pointed out that this method tends to be sensitive to the selection of the initial values. An indication of methods allowing to determine the proper initial parameter values will be the direction of the author's further research.

Acknowledgements. This work is supported by the National Centre of Science granted on the basis of decision number DEC-2013/11/D/HS4/04014.

References

Journal article

Aït-Sahalia Y (1996) Nonparametric pricing of interest rate derivative securities. Econometrica 64:527–560

Briers M, Doucet A, Maskell S (2010) Smoothing algorithms for state-space models. Ann Inst Stat Math 62:61–89

Brzozowska-Rup K, Dawidowicz A (2009) Metoda filtru cząsteczkowego. Matematyka Stosowana 10:69–107

Cappé O, Godsill SJ, Moulines E (2007) An overview of existing methods and recent advances in sequential Monte Carlo. Proc IEEE 95(5):899–924

Cappé O (2011) Expectation-Maximisation. In: Mengersen K, Titterington M, Robert CP (eds) Mixtures: estimation and applications. Wiley, pp 1–53

Cox JC, Ingersoll JE, Ross S (1985) A theory of term structure of interest rates. Econometrica 53:385–407

Chatterjee S (2005) Application of the Kalman filter for estimating Continuous time term structure models: the case of UK and Germany. Working paper, University of Glasgow

Chen SY (2012) Kalman filter for robot vision: a survey. IEEE Trans Ind Electron 59(11): 4409–4420

Christoffersen P, Dorion Ch, Jacobs K, Karoui L (2014) Nonlinear Kalman filtering in affine term structure models. CIRPEE, working paper 14-04

Chopin N (2004) Central limit theorem for sequential Monte Carlo methods and its application to Bayesian inference. Ann Stat 32(6):2385–2411

Dempster A, Larid N, Rubin D (1977) Maximum likelihood from incomplete data via the EM algorithm. J Roy Stat Soc B 39(1):1–38

Djurić PM, Kotecha JH, Zhang J, Huang Y, Ghirmai T, Bugallo MF, Míguez J (2003) Particle filtering. IEEE Signal Process. Mag 20(5):19–38

Douc R, Garivier A, Moulines E, Olsson J (2011) Sequential Monte Carlo smoothing for general state space hidden Markov models. Ann Appl Probab 21(6):2109–2145

Douc R, Moulines E (2008) Limit theorems for weighted samples with applications to sequential Monte Carlo methods. Ann Stat 36(5):2344–2376

De Rossi G (2010) Maximum likelihood estimation of the Cox-Ingersoll-Ross model using particle filters. Comput Econ 36(1):1–16

Doucet A, Johansen AM (2008) A tutorial on particle filtering and smoothing: fifteen years later. Technical report, Department of Statistics, University of British Columbia

Ghirmai T (2016) Distributed particle filter for target tracking: With reduced sensor communications. Sensors 16(9):1454

Gordon N, Salmond D, Smith AF (1993) Novel approach to nonlinear/non-Gaussian Bayesian state estimation. IEE Proc F Radar Signal Process 140:107–113

Gustafsson F, Gunnarsson F, Bergman N, Forssell U, Jansson J, Karlsson R, Nordlund PJ (2002) Particle filters for positioning, navigation and tracking. IEEE Trans Signal Process 50(2): 425–437

Higham DJ, Mao X, Stuart AM (2002) Strong convergence of Euler-type methods for nonlinear stochastic differential equations. SIAM J 40:1041–1063

Kantas N, Doucet A, Singh SS, Maciejowski J, Chopin N (2014) On particle methods for parameter estimation in state-space models. Stat Sci 30(3):328–351

Kladìvko K (2007) Maximum likelihood estimation of Cox-Ingersoll-Ross process: the Matlab implementation. Tech Comput Prague. http://www2.humusoft.cz/www/papers/tcp07/kladivko.pdf
Langrock R (2011) Some applications of nonlinear and non-Gaussian state-space modelling by means of hidden Markov models. J Appl Stat 38(12):2955–2970
Nordlund PJ, Gustafsson F (2008) Research on an improved terrain aided positioning model. Technical report LiTH-ISY-R2870, Linköpings Universitet
Olsson J, Douc R, Cappé O, Moulines E (2008) Sequential Monte Carlo smoothing with application to parameter estimation in nonlinear state-space models. Bernouli 14(1):155–179
Schön TB, Wills A, Ninness B (2011) System identification of nonlinear state-space models. Automatica 47(1):39–49
Vo LH (2014) Application of Kalman Filter on Modelling Interest Rates. J Manag Sci 1:1–15
Zhou HS (2013) Modified backward sampling smoothing with EM algorithm - application to economics and finance. Pre-print, Unpublished report

Book

Brigo D, Mercurio F (2001) Interest rate models. Theory and practice. Springer, Heidelberg
Cairns AJG (2004) Interest rate models: an introduction. Princeton University Press, Princeton
Del Moral P (2004) Feynman-Kac formulae. Genealogical and interacting particle systems with applications. Springer, Heidelberg
Del Moral P (2013) Mean field simulation for Monte Carlo integration. Chapman and Hall, Boca Raton
Dębski W (2010) Rynek finansowy i jego mechanizmy. Podstawy teorii i praktyki. Wydawnictwo Naukowe PWN, Warszawa
Doucet A, de Freitas N, Gordon N (2001) Sequential Monte Carlo methods in practice. Springer, Heidelberg
Gibson R, Lhabitant FS, Talay D (2001) Modeling the term structure of interest rates: a review of the literature. http://www.risklab.ch/ftp/papers/TermStructureSurvey.pdf
Hol JD, Schon TB, Gustafsson F (2006) On resampling algorithms for particle filters. In: Proceedings IEEE Nonlinear Statistical Signal Processing Workshop, pp 79–82
Hull JC (2015) Risk management and financial institutions, 4th edn. Wiley, New York
Weron A, Weron R (1999) Inżynieria finansowa. Wycena instrumentów pochodnych. Symulacje komputerowe. Statystyka rynku. Wydawnictwa Naukowo-Techniczne, Warszawa
Fabozzi FJ (2002) Interest rate, therm structure and valuation modeling. Wiley, New Jersey

An Overview of Robust Spectral Estimators

Valdério Anselmo Reisen[1,3(✉)], Céline Lévy-Leduc[2], Higor Henrique Aranda Cotta[1,3], Pascal Bondon[3], Marton Ispany[4], and Paulo Roberto Prezotti Filho[3,5]

[1] DEST and PPGEA-Universidade Federal do Espirito Santo-UFES, Vitória, Brazil
valderioanselmoreisen@gmail.com, valderio.reisen@ufes.br
[2] UMR MIA-Paris, AgroParisTech, INRA, Université Paris-Saclay, Paris, France
[3] Laboratoire des Signaux et Systémes (L2S), CNRS-CentraleSupélec-Université Paris-Sud, Gif sur Yvette, France
[4] University of Debrecen, Debrecen, Hungary
[5] PPGEA-UFES and IFES, Vitória, Brazil

Abstract. The periodogram function is widely used to estimate the spectral density of time series processes and it is well-known that this function is also very sensitive to outliers. In this context, this paper deals with robust estimation functions to estimate the spectral density of univariate and periodic time series with short and long-memory properties. The two robust periodogram functions discussed and compared here were previously explicitly and analytically derived in Fajardo et al. (2018), Reisen et al. (2017) and Fajardo et al. (2009) in the case of long-memory processes. The first two references introduce the robust periodogram based on M-regression estimator. The third reference is based on the robust autocovariance function introduced in Ma and Genton (2000) and studied theoretically and empirically in Lévy-Leduc et al. (2011). Here, the theoretical results of these estimators are discussed in the case of short and long-memory univariate time series and periodic processes. A special attention is given to the M-periodogram for short-memory processes. In this case, Theorem 1 and Corollary 1 derive the asymptotic distribution of this spectral estimator. As the application of the methodologies, robust estimators for the parameters of AR, ARFIMA and PARMA processes are discussed. Their finite sample size properties are addressed and compared in the context of absence and presence of atypical observations. Therefore, the contributions of this paper come to fill some gaps in the literature of modeling univariate and periodic time series to handle additive outliers.

Keywords: Time series · M-estimation · Q_N-estimation · Long-memory · Periodic processes · Robustness

1 Introduction

It is well known that outlying observations may completely destroy most of the standard estimators and several authors developed robust approaches in order to

F. Chaari et al. (Eds.): CSTA 2017, ACM 16, pp. 204–224, 2020.
https://doi.org/10.1007/978-3-030-22529-2_12

mitigate the impact of additive outliers, specially in time series models which is the process considered in this paper. However, most of the work is devoted to the robust estimation of the location, scale and other statistical tools. In this direction, the classical periodogram is the natural estimator of the spectral density of a time series and recent studies indicate that the periodogram is highly sensitive to the presence of outliers, and, thus, it becomes useless in any sub-sequential analysis. As a viable approach to attenuate this issue, the M-regression method applied to build alternative spectral estimators given in Fajardo et al. (2018) and Reisen et al. (2017) and the Q_N-periodogram introduced in Fajardo et al. (2009) are some methodologies proposed recently in the literature of time series to handle additive outliers.

The M-periodogram is discussed in Fajardo et al. (2018) and Reisen et al. (2017) for the long-memory time series. The short-range process was still an open problem and is one main contribution of this paper. The asymptotic property of the M-periodogram is derived for the process which is identified to have short-memory property such as an ARMA model (Theorem 1). As a second contribution of this paper, the recent results given Fajardo et al. (2018) and Reisen et al. (2017), for long-memory model, are summarized and these methods are compared empirically with Q_N-periodogram and the classical periodogram which is widely used in modelling time series data. Here, these methods are empirically studied and compared in time series with and without additive outliers with the aim to verify their finite sample size robustness properties, that is, to verify their capacity to accommodate the additive outlier's effect.

The use of M- and Q_N-periodograms in periodic ARMA (PARMA) models is also discussed here in the context of handling atypical or aberrant observations (additive outliers). This becomes the third contribution of this paper.

This paper is organized as follows: Sect. 2 discusses robust periodograms based on M-regression method and Q_N function for short and long-memory time series. Section 3 presents some simulation results for the methods discussed in Sect. 2. Section 4 gives some applications of the alternative periodograms in short and long-memory and periodic processes.

2 Robust Periodograms

Let $\{Y_t\}_{t\in\mathbb{Z}}$ be a second order stationary process. Since this paper deals with short and long-memory processes, additional assumptions on the process $\{Y_t\}_{t\in\mathbb{Z}}$ will be given in the sequel of the paper. For a sample $\{Y_1, Y_2, ..., Y_N\}$, the classical periodogram function, at the Fourier frequency $\lambda_j = 2\pi j/N, j = 1, \ldots, [N/2]$, is defined as

$$I_N(\lambda_j) = \frac{1}{2\pi N}\left|\sum_{k=1}^{N} Y_k \exp(\mathrm{i}k\lambda_j)\right|^2 . \tag{1}$$

Next subsections deal with alternative periodogram functions which present similar performance (from theoretical and empirical meaning) to $I_N(\lambda)$, $\lambda \in (-\pi, \pi)$, but with robustness property against additive outliers and asymmetric and heavy-tail distributions.

2.1 *M*-periodogram

One alternative way to derive the periodogram function $I_N(\lambda_j)$ is based on the Least Square (LS) estimates of a two-dimensional vector $\boldsymbol{\beta}' = (\beta^{(1)}, \beta^{(2)})$ in the linear regression model

$$Y_i = c'_{Ni}\boldsymbol{\beta} + \varepsilon_i = \beta^{(1)}\cos(i\lambda_j) + \beta^{(2)}\sin(i\lambda_j) + \varepsilon_i\ ,\ 1 \le i \le N,\ \boldsymbol{\beta} \in \mathbb{R}^2\ , \quad (2)$$

where ε_i denotes the deviation of Y_i from $c'_{Ni}\boldsymbol{\beta}$ and $\mathbb{E}(\varepsilon_i) = 0$ and $\mathbb{E}(\varepsilon_i^2) < \infty$. In the sequel (ε_i) is assumed to be a function of a stationary Gaussian process, see (10) for a precise definition. Then,

$$\hat{\boldsymbol{\beta}}_N^{\mathrm{LS}}(\lambda_j) = \underset{\boldsymbol{\beta}\in\mathbb{R}^2}{\mathrm{Arg\ min}} \sum_{i=1}^{N} (Y_i - c'_{Ni}(\lambda_j)\boldsymbol{\beta})^2\ , \quad (3)$$

where

$$c'_{Ni}(\lambda_j) = (\cos(i\lambda_j)\ \ \sin(i\lambda_j))\ . \quad (4)$$

The solution of (3) is

$$\hat{\boldsymbol{\beta}}_N^{\mathrm{LS}}(\lambda_j) = (C'C)^{-1}C'\mathbf{Y}\ , \quad (5)$$

where $\mathbf{Y} = (Y_1, \ldots, Y_N)'$, C and $C'C$ are defined by

$$C = \begin{pmatrix} \cos(\lambda_j) & \sin(\lambda_j) \\ \cos(2\lambda_j) & \sin(2\lambda_j) \\ \vdots & \vdots \\ \cos(N\lambda_j) & \sin(N\lambda_j) \end{pmatrix} \quad (6)$$

and

$$C'C = \begin{pmatrix} \sum_{k=1}^{N}\cos(k\lambda_j)^2 & \sum_{k=1}^{N}\cos(k\lambda_j)\sin(k\lambda_j) \\ \sum_{k=1}^{N}\cos(k\lambda_j)\sin(k\lambda_j) & \sum_{k=1}^{N}\sin(k\lambda_j)^2 \end{pmatrix} = \frac{N}{2}\,\mathrm{Id}_2 \quad (7)$$

where Id_2 is the identity matrix 2 by 2. Hence,

$$\begin{aligned} \hat{\boldsymbol{\beta}}_N^{\mathrm{LS}}(\lambda_j) &= \frac{2}{N}C'\mathbf{Y} = \frac{2}{N}\left(\sum_{k=1}^{N} Y_k\cos(k\lambda_j) \quad \sum_{k=1}^{N} Y_k\sin(k\lambda_j)\right)' \\ &= (\hat{\beta}_N^{\mathrm{LS},(1)}(\lambda_j), \hat{\beta}_N^{\mathrm{LS},(2)}(\lambda_j))'\ . \end{aligned} \quad (8)$$

In view of (1),

$$I_N(\lambda_j) = \frac{N}{8\pi}\|\hat{\boldsymbol{\beta}}_N^{\mathrm{LS}}(\lambda_j)\|^2 = \frac{N}{8\pi}\left((\hat{\beta}_N^{\mathrm{LS},(1)}(\lambda_j))^2 + (\hat{\beta}_N^{\mathrm{LS},(2)}(\lambda_j))^2\right) =: I_N^{\mathrm{LS}}(\lambda_j)\ , \quad (9)$$

where $\|\cdot\|$ denotes the classical Euclidean norm and $\hat{\boldsymbol{\beta}}_N^{\mathrm{LS}}(\lambda_j) = (\hat{\beta}_N^{\mathrm{LS},(1)}(\lambda_j), \hat{\beta}_N^{\mathrm{LS},(2)}(\lambda_j))'$ is the least square estimates of $\boldsymbol{\beta}' = (\beta^{(1)}, \beta^{(2)})$ see, for example, Fajardo et al. (2018) and Reisen et al. (2017) and references therein. Note that $I_N(\lambda_j)$ (9) can be derived for different choices of ε_i, $i = 1, \ldots, N$.

It is supposed here that

$$\varepsilon_i = G(\eta_i)\ . \tag{10}$$

In (10), G is a non null real-valued and skew symmetric measurable function (*i.e.* $G(-x) = -G(x)$, for all x) and $(\eta_i)_{i\geq 1}$ is a stationary Gaussian process with zero mean and unit variance. Additional assumptions of $(\eta_i)_{i\geq 1}$ will be given in the sequel of the paper.

Let $\psi(.)$ be a function satisfying the following assumptions.

(**A1**) $0 < \mathbb{E}[\psi^2(\varepsilon_1)] < \infty$.
(**A2**) The function ψ is absolutely continuous with its almost everywhere derivative ψ' satisfying $\mathbb{E}[|\psi'(\varepsilon_1)|] < \infty$ and such that the function $z \mapsto \mathbb{E}[|\psi'(\varepsilon_1 - z) - \psi'(\varepsilon_1)|]$ is continuous at zero.
(**A3**) ψ is nondecreasing, $\mathbb{E}[\psi'(\varepsilon_1)] > 0$ and $\mathbb{E}[\psi'(\varepsilon_1)^2] < \infty$.
(**A4**) ψ is skew symmetric, *i.e.* $\psi(-x) = -\psi(x)$, for all x.

It is now introduced the M-periodogram based on the M-estimator $\hat{\boldsymbol{\beta}}_N^{\mathrm{M}}$ of the parameter $\boldsymbol{\beta}$ defined in Eq. (2). The M-estimator $\hat{\boldsymbol{\beta}}_N^{\mathrm{M}} = (\hat{\beta}_N^{(1)}, \hat{\beta}_N^{(2)})'$ is defined as the solution (t_1, t_2) of

$$\sum_{i=1}^{N} \cos(i\lambda_j)\, \psi(Y_i - \cos(i\lambda_j)t_1) = 0 \text{ and } \sum_{i=1}^{N} \sin(i\lambda_j)\, \psi(Y_i - \sin(i\lambda_j)t_2) = 0. \tag{11}$$

$\hat{\beta}_N^{(1)}$ and $\hat{\beta}_N^{(2)}$ can be also seen as the minimizers with respect to t_1 and t_2, respectively, of

$$\left|\sum_{i=1}^{N} \cos(i\lambda_j)\, \psi(Y_i - \cos(i\lambda_j)t_1)\right| \text{ and } \left|\sum_{i=1}^{N} \sin(i\lambda_j)\, \psi(Y_i - \sin(i\lambda_j)t_2)\right|, \tag{12}$$

where ψ satisfies the same assumptions as in Koul and Surgailis (2000). By analogy to (9), the robust periodogram $I_N^M(\lambda_j)$ at $\lambda_j = 2\pi j/N, j = 1, \ldots, [N/2]$, is defined by

$$I_N^M(\lambda_j) = \frac{N}{8\pi} \|\hat{\boldsymbol{\beta}}_N^{\mathrm{M}}(\lambda_j)\|^2 = \frac{N}{8\pi} \left((\hat{\beta}_N^{(1)}(\lambda_j))^2 + (\hat{\beta}_N^{(2)}(\lambda_j))^2 \right). \tag{13}$$

2.1.1 *M*-periodogram in Short-Memory Processes

In this subsection the asymptotic properties of $\hat{\boldsymbol{\beta}}_N^{\mathrm{M}}$ are established in the short-range dependence framework. For this, the following assumptions are introduced. This result helps to establish the theoretical properties of the robust periodogram I_N^M given in Corollary 1.

(**A5**) Let η_t, $t \in \mathbb{Z}$, be i.i.d. standard Gaussian random variables and let a_j be real numbers such that $\sum_{j\geq 0} |a_j| < \infty$ and $a_0 = 1$. Then,

$$\varepsilon_i = \sum_{j\geq 0} a_j \eta_{i-j}.$$

(**A6**) ψ is the Huber function that is $\psi(x) = \max[\min(x, c), -c]$, for all x in $\mathbb{R}$, where c is a positive constant.

Theorem 1. *Assume that (A5) and (A6) hold and that* $\boldsymbol{\beta} = \mathbf{0}$ *in (2) so that* $Y_i = \varepsilon_i$. *Then, for any fixed* j, $\hat{\boldsymbol{\beta}}_N^M$ *defined by (12) satisfies*

$$\sqrt{\frac{N}{2}}(F(c) - F(-c))\hat{\boldsymbol{\beta}}_N^M(\lambda_j) \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, \boldsymbol{\Delta}^{(j)}\right), \ N \to \infty ,$$

where F *is the c.d.f. of* ε_1 *and*

$$\boldsymbol{\Delta}^{(j)} = \sum_{k\in\mathbb{Z}} \mathbb{E}\{\psi(\varepsilon_0)\psi(\varepsilon_k)\} \begin{pmatrix} \cos(k\lambda_j) & \sin(k\lambda_j) \\ -\sin(k\lambda_j) & \cos(k\lambda_j) \end{pmatrix}.$$

Theorem 1 is proved in Sect. 5.

Corollary 1. *Under the assumptions of Theorem 1,* $I_N^M(\lambda_j)$ *defined in (13) satisfies for any fixed* j,

$$I_N^M(\lambda_j) \xrightarrow{d} \frac{X^2 + Y^2}{4\pi(F(c) - F(-c))^2}, \quad \text{as } N \to \infty ,$$

where

$$X \sim \mathcal{N}\left(0, \sum_{k\in\mathbb{Z}} \mathbb{E}\{\psi(\varepsilon_0)\psi(\varepsilon_k)\}\cos(k\lambda_j)\right), \ Y \sim \mathcal{N}\left(0, \sum_{k\in\mathbb{Z}} \mathbb{E}\{\psi(\varepsilon_0)\psi(\varepsilon_k)\}\cos(k\lambda_j)\right)$$

and

$$\mathrm{Cov}(X, Y) = \sum_{k\in\mathbb{Z}} \mathbb{E}\{\psi(\varepsilon_0)\psi(\varepsilon_k)\}\sin(k\lambda_j) .$$

The proof of Corollary 1 is a straightforward consequence of Theorems 1 and (13).

2.1.2 *M*-periodogram for Long-Memory Processes

Now, consider the following assumption for $(\eta_i)_{i\geq 1}$ in the case of long-memory process. The results in this subsection are well detailed in Fajardo et al. (2018).

(**A7**) $(\eta_i)_{i\geq 1}$ is a stationary zero-mean Gaussian process with covariances $\rho(k) = \mathbb{E}(\eta_1\eta_{k+1})$ satisfying:

$$\rho(0) = 1 \text{ and } \rho(k) = k^{-D}L(k), \ 0 < D < 1 ,$$

where the function L is slowly varying at infinity and is positive for large k. Recall that a slowly varying function $L(x), x > 0$ is such that $L(xt)/L(x) \to 1$, as $x \to \infty$ for any $t > 0$. Constants and logarithms are example of slowly varying functions.
Moreover, the spectral density f of $(\eta_i)_{i\geq 1}$ can be expressed as:

$$f(\lambda) = |1 - \exp(-\mathrm{i}\lambda)|^{-2d} f^*(\lambda) \,, \tag{14}$$

where $d \in (0, 1/2)$ and f^* is an even, positive, continuous function on $(-\pi, \pi]$, bounded above and bounded away from zero.

Note that

$$D = 1 - 2d \,, \tag{15}$$

where D is defined in Assumption (A7) and d is the standard long-memory parameter notation given in the literature of long-memory models. The fact that $(\eta_i)_{i\geq 1}$ is required to satisfy (A7) essentially means that both $L(x)$, $x \geq 1$ and $f^*(\lambda)$, λ in $(-\pi, \pi]$ satisfy some smoothness properties.

Theorem 2. *Assume that (A7), (A1), (A2), (A3) and (A4) hold and that $\boldsymbol{\beta} = 0$ in (2) so that $Y_i = \varepsilon_i$. Then, for any fixed j, $\hat{\boldsymbol{\beta}}_N^M(\lambda_j)$ defined by (12) satisfies*

$$\sqrt{\frac{N}{2}}\hat{\boldsymbol{\beta}}_N^M(\lambda_j) = \frac{J_1}{\mathbb{E}[\psi'(\varepsilon_1)]} \left\{ \sqrt{\frac{2}{N}} \sum_{i=1}^{N} \begin{pmatrix} \cos(i\lambda_j) \\ \sin(i\lambda_j) \end{pmatrix} \eta_i \right\} + o_p(N^{(1-D)/2}) \,, \text{ as } N \to \infty \,, \tag{16}$$

where $J_1 = \mathbb{E}[\psi(G(\eta))\eta] \neq 0$, η being a standard Gaussian random variable and $D = 1 - 2d$. Moreover,

$$N^{D/2}\hat{\boldsymbol{\beta}}_N^M(\lambda_j) \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, \frac{J_1^2}{(\mathbb{E}[\psi'(\varepsilon_1)])^2}\widetilde{\Gamma}\right) \,, \; N \to \infty \,, \tag{17}$$

where

$$\widetilde{\Gamma} = \lim_{N\to\infty} \frac{4}{N^{2-D}} \sum_{1\leq k,\ell\leq N} c_{Nk}(\lambda_j) c_{N\ell}^T(\lambda_j) \rho(k-\ell) \tag{18}$$

$$= 8\pi \times (2\pi j)^{-2d} f^*(0) \begin{pmatrix} \mathcal{L}_1 & 0 \\ 0 & \mathcal{L}_2 \end{pmatrix}. \tag{19}$$

In Relation (18), the vector $c_{Nk}(\lambda_j)$ is defined in (4),

$$\mathcal{L}_1 = \frac{1}{\pi}\int_{\mathbb{R}} \frac{\sin^2(\lambda/2)}{(2\pi j - \lambda)^2} \left|\frac{\lambda}{2\pi j}\right|^{-2d} \mathrm{d}\lambda - \frac{1}{\pi}\int_{\mathbb{R}} \frac{\sin^2(\lambda/2)}{(2\pi j - \lambda)(2\pi j + \lambda)} \left|\frac{\lambda}{2\pi j}\right|^{-2d} \mathrm{d}\lambda \,, \tag{20}$$

and

$$\mathcal{L}_2 = \frac{1}{\pi}\int_{\mathbb{R}} \frac{\sin^2(\lambda/2)}{(2\pi j - \lambda)^2} \left|\frac{\lambda}{2\pi j}\right|^{-2d} \mathrm{d}\lambda + \frac{1}{\pi}\int_{\mathbb{R}} \frac{\sin^2(\lambda/2)}{(2\pi j - \lambda)(2\pi j + \lambda)} \left|\frac{\lambda}{2\pi j}\right|^{-2d} \mathrm{d}\lambda \,. \tag{21}$$

Corollary 2. *Under the assumptions of Theorem 2, the periodogram* I_N^M *defined in (13) satisfies*

$$N^{D-1} I_N^M(\lambda_j) \xrightarrow{d} (Z_1^2 + Z_2^2)\ ,\ \text{as } N \to \infty\ , \tag{22}$$

where (Z_1, Z_2) *is a zero-mean uncorrelated Gaussian vector with covariance matrix equal to*

$$\frac{J_1^2}{8\pi(\mathbb{E}[\psi'(\varepsilon_1)])^2}\widetilde{\Gamma}\ , \tag{23}$$

with $\widetilde{\Gamma}$ *defined in (18).*

Theorem 2 and Corollary 2 are proved in Fajardo et al. (2018).

2.2 Q_N-periodogram

Another possible approach to obtain the classical periodogram (1) is to write it in terms of the sample autocovariance function

$$I_N(\lambda_j) = \frac{1}{2\pi} \sum_{h=-(N-1)}^{N-1} \widehat{\gamma}(h)\cos(h\lambda_j), \tag{24}$$

where $\lambda_j = 2\pi j/N, j = 1, \ldots, [N/2]$ and $\widehat{\gamma}(h)$ is the classical sample autocovariance function for a sample $\{Y_1, ..., Y_N\}$.

A straightforward approach to robustify (24) is to plug in a robust autocovariance function replacing the classical one. This methodology is now addressed.

For a sample $x_1, ..., x_N$ Rousseeuw and Croux (1993) proposed a robust scale estimator function $Q_N(\cdot)$ which is based on the τth order statistic of $\binom{N}{2}$ distances $\{|x_j - x_k|, j < k\}$, and can be written as

$$Q_N(x) = \kappa \times \{|x_j - x_k|; j < k\}_{(\tau)}, \tag{25}$$

where κ is a constant used to guarantee consistency ($\kappa = 2.2191$ for the Gaussian distribution) and $\tau = \lfloor(\binom{N}{2} + 2)/4\rfloor + 1$. The above function can be evaluated using the algorithm proposed by Croux and Rousseeuw (1992), which is computationally efficient.

Based on $Q_N(\cdot)$, Ma and Genton (2000) proposed a highly robust estimator for the autocovariance function:

$$\widehat{\gamma}_{Q_N}(h) = \frac{1}{4}\left[Q_{N-h}^2(\mathbf{u}+\mathbf{v}) - Q_{N-h}^2(\mathbf{u}-\mathbf{v})\right], \tag{26}$$

where $\mathbf{u}$ and $\mathbf{v}$ are vectors containing the initial $N-h$ and the final $N-h$ observations of $x_1, ..., x_N$, respectively. The robust estimator for the autocorrelation function is

$$\widehat{\rho}_{Q_N}(h) = \frac{Q_{N-h}^2(\mathbf{u}+\mathbf{v}) - Q_{N-h}^2(\mathbf{u}-\mathbf{v})}{Q_{N-h}^2(\mathbf{u}+\mathbf{v}) + Q_{N-h}^2(\mathbf{u}-\mathbf{v})}\ . \tag{27}$$

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

Now, returning to (24), the robust Q_N-periodogram for a sample $\{Y_1, ..., Y_N\}$ is defined by

$$I_N^{Q_N}(\lambda_j) = \frac{1}{2\pi} \sum_{h=-(N-1)}^{N-1} \widehat{\gamma}_{Q_N}(h) \cos(h\lambda_j), \tag{28}$$

where $\lambda_j = 2\pi j/N, j = 1, \ldots, [N/2]$.

The theoretical properties of $I_N^{Q_N}$ are still under study. Therefore, in the sequel, the asymptotic properties of $\hat{\gamma}_{Q_N}$ are summarized for short and long memory processes. These are well detailed in Lévy-Leduc et al. (2011).

2.2.1 Main Asymptotic Results for Short Memory Process

In the short-memory scenario, the process under study $(Y_i)_{i\geq 1}$ satisfies the following assumption (see, also, Lévy-Leduc et al. 2011):

(**A8**) $(Y_i)_{i\geq 1}$ is a stationary zero-mean Gaussian process with autocovariance sequence $\gamma(h) = \mathbb{E}(Y_1 Y_{h+1})$ satisfying:

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

Theorem 3. *Assume that (A8) holds and let h be a non negative integer. Then, the autocovariance estimator $\widehat{\gamma}_{Q_N}(h)$ satisfies the following Central Limit Theorem:*

$$\sqrt{N}\left(\widehat{\gamma}_{Q_N}(h) - \gamma(h)\right) \xrightarrow{d} \mathcal{N}(0, \check{\sigma}_h^2), N \to \infty\ ,$$

where

$$\check{\sigma}^2(h) = \mathbb{E}[\zeta^2(Y_1, Y_{1+h})] + 2\sum_{k\geq 1} \mathbb{E}[\zeta(Y_1, X_{1+h})\zeta(Y_{k+1}, Y_{k+1+h})]\ , \tag{29}$$

and the function ζ is defined by

$$\zeta : (x, y) \mapsto \left\{(\gamma(0)+\gamma(h))\ \mathrm{IF}\left(\frac{x+y}{\sqrt{2(\gamma(0)+\gamma(h))}}, Q, \Phi\right) - (\gamma(0)-\gamma(h))\ \mathrm{IF}\left(\frac{x-y}{\sqrt{2(\gamma(0)-\gamma(h))}}, Q, \Phi\right)\right\}. \tag{30}$$

where IF *is defined by*

$$\mathrm{IF}(x, Q, \Phi) = \kappa\left(\frac{1/4 - \Phi(x+1/\kappa) + \Phi(x-1/\kappa)}{\int_{\mathbb{R}} \phi(y)\phi(y+1/\kappa)\mathrm{d}y}\right)\ , \tag{31}$$

where Φ and ϕ denote the c.d.f. and p.d.f. of a standard Gaussian random variable, respectively with κ defined in (25).

Theorem 3 is proved in Lévy-Leduc et al. (2011).

2.2.2 Main Asymptotic Results for Long-Memory Process

The following results concern the robust autocovariance function for long-memory process see, also, Lévy-Leduc et al. (2011).

(**A9**) $(Y_i)_{i\geq 1}$ is a stationary zero-mean Gaussian process with autocovariance $\gamma(h) = \mathbb{E}(Y_1 Y_{h+1})$ satisfying:

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

where L is slowly varying at infinity and is positive for large h. Note that, as previously stated, $D = 1 - 2d$.

Theorem 4. *Assume that (A9) holds and that L has three continuous derivatives. Assume also that $L_i(x) = x^i L^{(i)}(x)$ satisfy: $L_i(x)/x^\epsilon = O(1)$, for some ϵ in $(0, D)$, as x tends to infinity, for all $i = 0, 1, 2, 3$, where $L^{(i)}$ denotes the ith derivative of L. Let h be a non negative integer. Then, $\widehat{\gamma}_{Q_N}(h)$ satisfies the following limit theorems as N tends to infinity.*

(i) If $D > 1/2$,

$$\sqrt{N}\left(\widehat{\gamma}_{Q_N}(h) - \gamma(h)\right) \xrightarrow{d} \mathcal{N}(0, \check{\sigma}^2(h)),$$

where

$$\check{\sigma}^2(h) = \mathbb{E}[\zeta^2(Y_1, Y_{1+h})] + 2\sum_{k\geq 1} \mathbb{E}[\zeta(Y_1, Y_{1+h})\zeta(Y_{k+1}, Y_{k+1+h})],$$

ζ being defined in (30).

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

$$\beta(D)\frac{N^D}{\widetilde{L}(N)}\left(\widehat{\gamma}_{Q_N}(h) - \gamma(h)\right) \xrightarrow{d} \frac{\gamma(0) + \gamma(h)}{2}\left(Z_{2,D}(1) - Z_{1,D}(1)^2\right)$$

where $\beta(D) = \mathrm{B}((1-D)/2, D)$, B denotes the Beta function, the processes $Z_{1,D}(\cdot)$ and $Z_{2,D}(\cdot)$ are defined as follows:

$$Z_{1,D}(t) = \int_{\mathbb{R}} \left[\int_0^t (u - x)_+^{-(D+1)/2} \mathrm{d}u\right] \mathrm{d}B(x), \quad 0 < D < 1, \tag{32}$$

$$Z_{2,D}(t) = \int_{\mathbb{R}^2}' \left[\int_0^t (u - x)_+^{-(D+1)/2}(u - y)_+^{-(D+1)/2} du\right] \mathrm{d}B(x)\mathrm{d}B(y),\ 0 < D < 1/2\ , \tag{33}$$

and

$$\widetilde{L}(N) = 2L(N) + L(N + h)(1 + h/N)^{-D} + L(N - h)(1 - h/N)^{-D}, \tag{34}$$

where B is the standard Brownian motion. The symbol $\int'$ means that the domain of integration excludes the diagonal.

Theorem 4 is proved in Lévy-Leduc et al. (2011).

3 Monte Carlo Simulation

In this section, small sample size experiments are conducted with the aim to clarify the empirical performance of the spectral estimates discussed previously in a different context such as time series with additive outliers. Based on this, some standard questions, such as (1) what is the best method to be used in a real application? (2) which method (if any) should be considered when dealing with outliers? (3) Does the large observation (if any) make similar outlier's effect on the statistical time series modelling functions, that is, on the ACF and periodogram functions? Among others, are expected to be answered or, at least, clarified.

Let $\{X_t\}_{t=1,\ldots,N}$ be a sample from a Gaussian second order stationary process and let $\{Y_t\}_{t=1,\ldots,N}$ be a sample of the process defined by

$$Y_t = X_t + \omega W_t \tag{35}$$

where the parameter ω represents the magnitude of the outlier, and W_t is a random variable with probability distribution

$$\mathbb{P}(W_t = -1) = \mathbb{P}(W_t = 1) = \delta/2 \text{ and } \mathbb{P}(W_t = 0) = 1 - \delta,$$

where $\mathbb{E}[W_t] = 0$ and $\mathbb{E}[W_t^2] = \text{Var}(W_t) = \delta$. Note that (35) is based on the parametric models proposed by Fox (1972). W_t is the product of $Bernoulli(\delta)$ and *Rademacher* random variables; the latter equals 1 or -1, both with probability 1/2. X_t and W_t are independent random variables. Note that, if $\omega = 0.0$ $\{Y_t\}$ is an outlier free time series.

In order to compare the performance of M- and Q_N-periodogram, a Monte Carlo investigation was carried out under different contamination scenarios. For the simulations, the number of replications was 5000, the samples $\{X_t\}$ of size $N = 500$ were generated according to a model autocorrelation structure, which is given in what follows, and the contaminated data Y_t were generated from (35) with $\delta = 0.01$ for magnitudes $\omega = 0$ (no outliers) and 10.

The comparison between the methods is performed by estimating α in the linear regression $\log(I(\lambda_j)) \simeq const + \alpha \log(\lambda_j) + E_j$, $j = 1, \ldots, N^{0.7}$, where $I(.)$ is either $I_N(.)$, $I_N^M(.)$ or $I_N^{Q_N}(.)$. The data were generated based on

$$X_t = (1 - B)^{-d} Z_t = \sum_{j \geq 0} \frac{\Gamma(j + d)}{\Gamma(j + 1)\Gamma(d)} \epsilon_{t-j}, \tag{36}$$

where ϵ_t is an AR(1) model, that is, $\epsilon_t = \phi \epsilon_{t-1} + \eta_t$, where η_t, $t = 1, ..., N$, are i.i.d. standard Gaussian random variables.

In the finite sample size investigation, the model correlation structures are divided in two cases:

1. An AR(1) model with $\phi = 0.6$ and $d = 0$.
2. An ARFIMA$(0, d, 0)$ model with $d = 0.3$.

Figure 1 displays the plots of the empirical densities of $\hat{\alpha}_{I_N}$, $\hat{\alpha}_{I_N^M}$ and $\hat{\alpha}_{I_N^{Q_N}}$ for the case of AR(1) models without contamination ($\omega = 0$). Although, $\hat{\alpha}_{I_N^M}$ has a slight better performance than $\hat{\alpha}_{I_N^{Q_N}}$, that is, the first method and the classical periodogram presented very close densities, all the methods provided similar results showing that, even for small sample sizes, the empirical density is very close which corroborate the theoretical results discussed previously. Based on the asymptotic theory and the empirical results all three methods can be used to estimate the spectral density of a time series when there is no contamination of additive outliers. This opens an important contribution in the context that alternative spectral estimators such as I_N^M and $I_N^{Q_N}$ can be used instead of the classical periodogram I_N in the step procedure for modelling time series data. For example, these estimators can be an alternative tools to be used in the Whittle function to obtain the parameter estimates. This will be also discussed in what follows. Note that, the disadvantage of $I_N^{Q_N}$ over I_N^M and I_N is that the ACF using $Q_N(.)$ does not have the positive definite property.

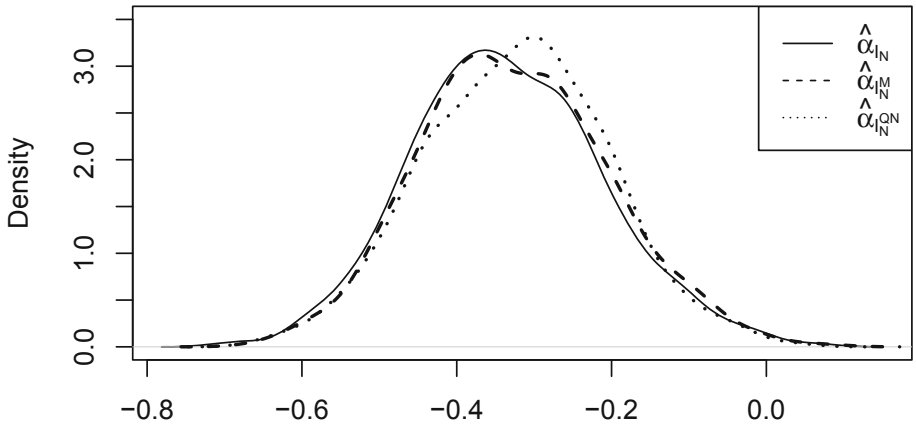


Fig. 1. Densities of $\hat{\alpha}_{I_N}$, $\hat{\alpha}_{I_N^M}$ and $\hat{\alpha}_{I_N^{Q_N}}$ for AR(1) models with $\phi = 0.6$ and $\omega = 0$.

When the data is contaminated with additive outliers the scenario changes significantly. As well known, the periodogram, which depends on the classical autocovariance, is corrupted by the outliers. Therefore, the alternative methods are almost unaffected. This is displayed in Fig. 2 in which $\omega = 10$ and $\delta = 0.01$. The empirical density of $\hat{\alpha}_{I_N}$ is shifted to the right side which is an expected result since the variance increases with outliers. The empirical densities of $\hat{\alpha}_{I_N^M}$ and $\hat{\alpha}_{I_N^{Q_N}}$ remain almost unchangeable.

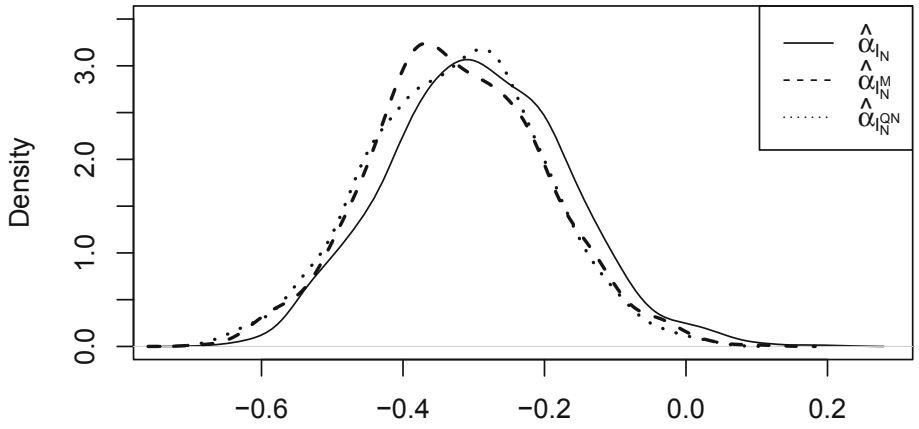


Fig. 2. Densities of $\hat{\alpha}_{I_N}$, $\hat{\alpha}_{I_N^M}$ and $\hat{\alpha}_{I_N^{Q_N}}$ for AR(1) models with $\phi = 0.6$, $\delta = 0.01$ and $\omega = 10$.

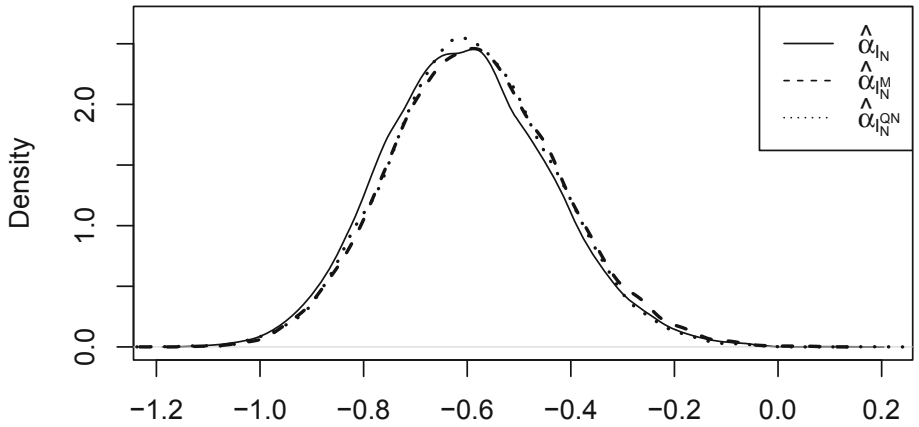


Fig. 3. Densities of $\hat{\alpha}_{I_N}$, $\hat{\alpha}_{I_N^M}$ and $\hat{\alpha}_{I_N^{Q_N}}$ when $d = 0.3$, $N = 500$ and $\omega = 0$.

In the case of long-memory process, the empirical density plots are given in Figs. 3 and 4 for non-contaminated and contaminated time series, respectively. Similar conclusions of the AR case are drawn. That is, in the uncontaminated scenarios, all three methods displayed similar densities although the method M and the classical one (periodogram) are very close. In the contaminated case, the classical one is totally affected by the additive outliers. Reinforcing that the ACF using Q_N does not have the positiveness property.

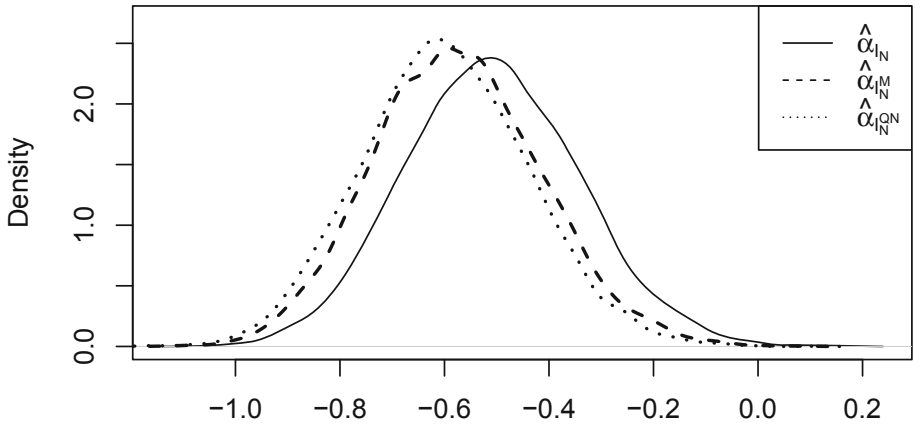


Fig. 4. Densities of $\hat{\alpha}_{I_N}$, $\hat{\alpha}_{I_N^M}$ and $\hat{\alpha}_{I_N^{Q_N}}$ when $d = 0.3$, $N = 500$, $\delta = 0.01$ and $\omega = 10$.

4 Applications of M and Q_N-periodograms

4.1 Robust Estimation of the Fractional Parameter

Based on the theoretical results discussed previously, this section introduces some applications related to the use of M-regression and Q_N estimation functions. The application is divided in two cases: (a) Estimation of the fractional parameter d in long-memory processes; (b) Estimation in periodic AR (PAR) processes. Some finite sample size investigation is also addressed in the context of time series with and without outliers.

(a) *Estimation of the fractional parameter in long-memory process*

The estimation methods of the fractional parameter d discussed here are derived from the well-known semi-parametric regression method (GPH) originally proposed by Geweke and Porter- Hudak (1983). The regression estimation methods based on I_N^M and $I_N^{Q_N}$ were previously introduced in Reisen et al. (2017) and Fajardo et al. (2009), respectively, papers where the reader will find more details related to theoretical and empirical results of these estimation methodologies.

(**A10**) $(\varepsilon_i)_{i\geq 1}$ is a stationary mean-zero Gaussian process with spectral density given in Assumption (A7).

For estimating the fractional parameter d of long-memory processes having their spectral density satisfying (14), it is usual to use the standard GPH (Geweke and Porter-Hudak 1983) estimator defined in the following. This estimator is motivated heuristically by starting from

$$\log(f(\lambda_j)) = -2d\log(|2\sin(\lambda_j/2)|) + \log(f^*(\lambda_j)) = -2dX_j + \log(f^*(\lambda_j))$$
$$= \log(f_0^\star) - 2dX_j + \log(f_j^\star/f_0^\star), \tag{37}$$

where $X_j = \log|2\sin(\lambda_j/2)|$ and $f_j^\star = f^\star(\lambda_j)$. If

$$\varepsilon_j^R = \log\left(\frac{I_N(\lambda_j)}{f(\lambda_j)}\right), \tag{38}$$

then

$$\log(I_N(\lambda_j)) = \varepsilon_j^R + \log(f(\lambda_j)),$$

and, by (37),

$$\log(I_N(\lambda_j)) = \log(f_0^\star) - 2dX_j + \log(f_j^\star/f_0^\star) + \varepsilon_j^R. \tag{39}$$

The GPH estimator is given by

$$\hat{d}^{\mathrm{GPH}} = \frac{-0.5\sum_{j=1}^{m_N}(X_j - \bar{X})\log(I_N^{\mathrm{LS}}(\lambda_j))}{\sum_{k=1}^{m_N}(X_k - \bar{X})^2}, \tag{40}$$

where $X_j = \log|2\sin(\lambda_j/2)|$, $\bar{X} = \sum_{j=1}^{m_N} X_j/m_N$, $I_N^{\mathrm{LS}}(\lambda_j)$ is defined in (9) and m_N is a function of N.

Based on the above discussion, one way to define a M-regression estimator of d consists in replacing I_N^{LS} in (40) by I_N^M defined in (13):

$$\hat{d}^M = \frac{-0.5\sum_{j=1}^{m_N}(X_j - \bar{X})\log(I_N^M(\lambda_j))}{\sum_{k=1}^{m_N}(X_k - \bar{X})^2}, \tag{41}$$

where $X_j = \log|2\sin(\lambda_j/2)|$, $\bar{X} = \sum_{j=1}^{m_N} X_j/m_N$ and m_N is a function of N which is specified in Theorem 5.

The theoretical properties of $\hat{d}^{\mathrm{M}}$ are established under the following assumptions. The random process (ε_j) is obtained through a moving average process:

$$\varepsilon_j = \sum_{k\le j} a_{j-k}\zeta_k, \quad a_j = L(j)j^{-(1+D)/2}, \ j \ge 1, \tag{42}$$

for some D in $(0,1)$, where $L(\cdot)$ is a positive slowly varying function at infinity and where the random variables ζ_k are i.i.d. with zero mean and variance 1. It is assumed that the distribution of ζ_0 satisfies

$$\left|\mathbb{E}(\mathrm{e}^{\mathrm{i}u\zeta_0})\right| \le C(1+|u|)^{-\delta}, \ u \in \mathbb{R}. \tag{43}$$

where $C < \infty$ and $\delta > 0$ are constants. Note that, Conditions (42) and (43) imply that the cumulative distribution function F_{ε_0} of ε_0 is infinitely boundedly differentiable, see Koul and Surgailis (2000).

Theorem 5. *Let $Y_i = \varepsilon_i$, for all i in $\{1, \ldots, N\}$, where ε_i satisfy (42) and (A10). Assume that $1/D$ is not an integer and that $\boldsymbol{\beta} = \mathbf{0}$ in (2). Assume moreover that $\mathbb{E}(\zeta_0^{4 \vee 2k^\star}) < \infty$, where $k^\star = [1/D]$, ζ_0 is defined in (42) and satisfies (43), $\nu_1 \neq 0$, $\nu_2 = 0$ and $\nu_3 \neq 0$, where the ν_k are defined by*

$$\nu_k = \int_0^\infty \psi(y) \left[1 - (-1)^k\right] f^{(k)}(y) \mathrm{d}y, \quad \text{for all integer } k \geq 0 \,, \tag{44}$$

where ψ is the Huber function. Then, if $1/3 < D < 1$,

$$\sqrt{m_N}(\hat{d}^M - d) \xrightarrow{d} \mathcal{N}(0, \pi^2/24), \ \text{as } N \to \infty \,, \tag{45}$$

where $\hat{d}^M$ is defined in (41) and $m_N = N^\beta$ with $0 < \beta < (1 - D)/3$.

This result is proved in Reisen et al. (2017).

Another way of defining a robust estimator of d is to consider:

$$\hat{d}^{Q_N} = \frac{-0.5 \sum_{j=1}^{m_N} (X_j - \bar{X}) \log(I_N^{Q_N}(\lambda_j))}{\sum_{k=1}^{m_N} (X_k - \bar{X})^2} \,, \tag{46}$$

where $X_j = \log|2\sin(\lambda_j/2)|$, $\bar{X} = \sum_{j=1}^{m_N} X_j/m_N$, $I_N^{Q_N}(\lambda_j)$ is defined in (28) and m_N is a function of N. For further information, see Fajardo et al. (2009). The asymptotic property of $\hat{d}^{Q_N}$ is still an open problem, however, the empirical results given in Fajardo et al. (2009) support the use of this method under time series with and without outliers. The performance of fractional estimators $\hat{d}^{GPH}$, $\hat{d}^M$ and $\hat{d}^{Q_N}$ is the motivation of the next subsection for long-memory time series with and without additive outliers.

4.1.1 Finite Sample Size Investigation

In this subsection, the numerical experiments were carried out in accordance with the model of Sect. 3. For the simulations, $N = 500$, $\omega = 10$ and $\delta = 0.01$ for 5000 replications. The results are displayed in Figs. 5, 6 and Table 1. Since there is not short-memory component in the model m_N was fixed at $N^{0.7}$ for all tree methods.

Figure 5 presents the boxplots with the results of $\hat{d}_{GPH}$, $\hat{d}_M$ and $\hat{d}_{Q_N}$ estimators for the uncontaminated scenario. $\hat{d}_M$ and $\hat{d}_{Q_N}$ seem to present positive

Table 1. Empirical Mean, Bias and RMSE of $\hat{d}_{GPH}$, $\hat{d}_M$ and $\hat{d}_{Q_N}$ when $\omega = 10$ and $\delta = 0, 0.01, 0.05$.

d	δ	MEAN			BIAS			RMSE		
		$\hat{d}_{GPH}$	$\hat{d}_M$	$\hat{d}_{Q_N}$	$\hat{d}_{GPH}$	$\hat{d}_M$	$\hat{d}_{Q_N}$	$\hat{d}_{GPH}$	$\hat{d}_M$	$\hat{d}_{Q_N}$
0.3	0.0	0.3029	0.2950	0.2933	0.0029	−0.0049	−0.0066	0.0601	0.0596	0.0558
	0.01	0.2226	0.2899	0.3052	−0.0773	−0.0101	0.0052	0.0972	0.0581	0.0584
	0.05	0.1225	0.2681	0.3236	−0.1775	−0.0318	0.0236	0.1873	0.0689	0.0682

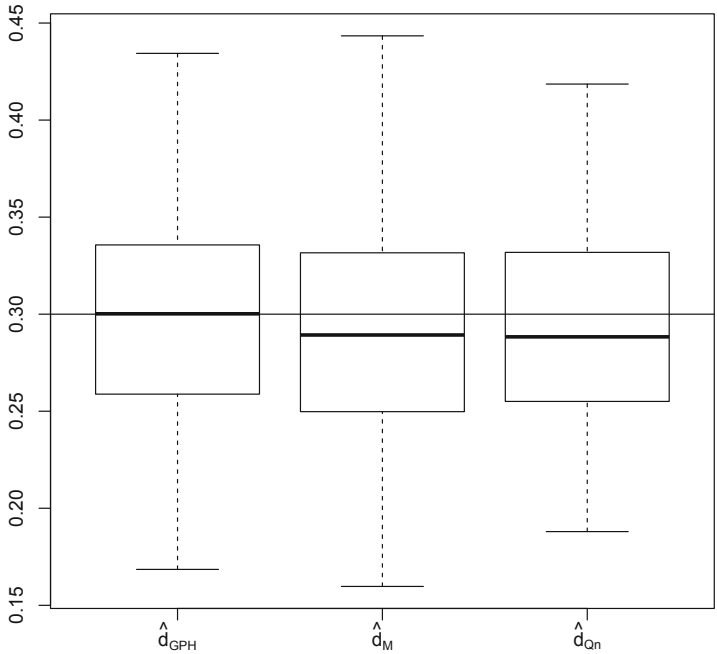


Fig. 5. Boxplots of $\hat{d}_{GPH}$, $\hat{d}_M$ and $\hat{d}_{Q_N}$ when $\delta = 0$.

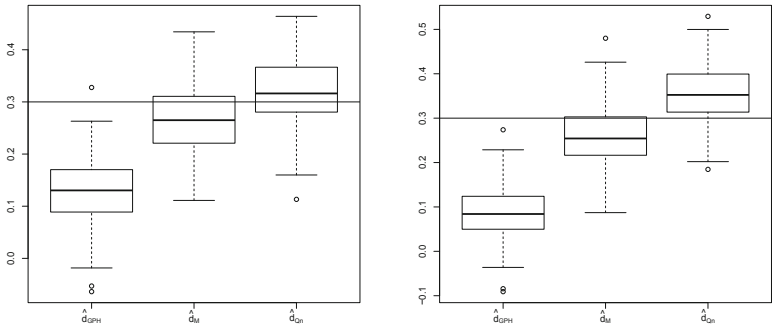


Fig. 6. Boxplots of $\hat{d}_{GPH}$, $\hat{d}_M$ and $\hat{d}_{Q_N}$ when $\delta = 0.05$ and $\delta = 0.1$, respectively.

bias and, surprisingly, $\hat{d}_{Q_N}$ displays smaller deviation. However, in general, all methods perform similarly, i.e., all estimation methods leaded to comparable estimates close to the real values of d.

Figure 6 displays the boxplots of $\hat{d}_{GPH}$, $\hat{d}_M$ and $\hat{d}_{Q_N}$ when the series has outliers. As can be perceived from the boxplots, the GPH estimator is clearly affected by additive outliers while the robust ones keep almost the same picture as the one of the non-contaminated scenario, except that the bias of $\hat{d}_{Q_N}$ becomes negative, that is, this estimator tends to overestimate the true parameter.

The empirical mean, bias and mean square root are displayed in Table 1. This numerically corroborates the results discussed based on Figs. 5, 6, that is, the estimators have similar performance in the absence of outliers in the data. While the performance of $\hat{d}_{GPH}$ changes dramatically in the presence of outliers, the estimates from $\hat{d}_{Q_N}$ and $\hat{d}_M$ keep almost unchangeable. As a general conclusion, the empirical result suggests that all the methods can be used to estimate the parameter d when there is not a suspicion of additive or abrupt observation. However, in the existence of a single atypical observation, the methods $\hat{d}_{Q_N}$ and $\hat{d}_M$ should be preferred. Similar conclusions are given in Fajardo et al. (2009) and Reisen et al. (2017) for $\hat{d}_{Q_N}$ and $\hat{d}_M$, respectively.

4.2 Q_n and M-estimators in PARMA Models

One of the most popular periodic causal process is the PARMA model which generalizes the ARMA model. $\{Z_t\}_{t\in\mathbb{Z}}$ is said to be a PARMA model if it satisfies the difference equation

$$\sum_{j=0}^{p_\nu} \phi_{\nu,j} Z_{r\mathcal{S}+\nu-j} = \sum_{k=0}^{q_\nu} \theta_{\nu,k} \varepsilon_{r\mathcal{S}+\nu-k}, r \in \mathbb{Z} \tag{47}$$

where for each season ν ($1 \leq \nu \leq \mathcal{S}$) where $\mathcal{S}$ is the period, p_ν and q_ν are the AR and MA orders, respectively, $\phi_{\nu,1}, \ldots, \phi_{\nu,p_\nu}$ and $\theta_{\nu,1}, \ldots, \theta_{\nu,q_\nu}$ are the AR and MA coefficients, respectively, and $\phi_{\nu,0} = \theta_{\nu,0} = 1$. The sequence $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ is zero-mean and uncorrelated, and has periodic variances with period $\mathcal{S}$, i.e. $\mathrm{E}(\varepsilon^2_{r\mathcal{S}+\nu}) = \sigma^2_\nu$ for $\nu = 1, \ldots, \mathcal{S}$. In the following, $p = \max_\nu p_\nu$, $q = \max_\nu q_\nu$, $\phi_{\nu,j} = 0$ for $j > p_\nu$, $\theta_{\nu,k} = 0$ for $k > q_\nu$, and (47) is referred as the PARMA$(p, q)_\mathcal{S}$ model (see, for example, Basawa and Lund 2001 and Sarnaglia et al. 2015).

To deal with outliers effect in the estimation of PAR model, Sarnaglia et al. (2010) proposed the use of the $Q_N(.)$ function in this model. Following the same lines of the linear time series model described previously, the $Q_N(.)$ function is used to compute an estimator of the periodic autocovariance function $\gamma^{(\nu)}(h)$ at lag h and this sample ACF based on $Q_N(.)$ estimator, denoted here as $\gamma_Q^{(\nu)}(h)$, replaces the classical periodic ACF $\gamma^{(\nu)}(h)$ in the Yule-Walker periodic equations (see, for example, McLeod 1994 and Sarnaglia et al. 2010) to derive an alternative parameter estimator method for a periodic AR model. The authors derived some asymptotic and empirical properties of the proposed estimator. They showed that the method well accommodate the effect of additive outliers, that is, it presented robustness against these type of observations in the finite sample size series as well as in a real data set.

Let now $Z_1, \ldots Z_N$, where $N = n\mathcal{S}$, be a sample from PAR process which is a particular case of the model definition in (47) with $q_\nu = 0$ and let now $Q_N(.)$ for PAR process be defined as

$$Q_N^{(\nu)}(Z) = Q_N(\{Z_{r\mathcal{S}+\nu}\}_{0\leq r\leq N}). \tag{48}$$

Based on $Q_N^{(\nu)}(Z)$, the authors derived the sample ACF for periodic stationary processes $\hat{\gamma}_Q^{(\nu)}(h)$. Under some model assumptions, they proved the following main results.

1. For a fixed lag h, $\hat{\gamma}_Q^{(\nu)}(h)$ satisfies the following central limit theorem: As $N \longrightarrow \infty$,
$$\sqrt{N}\left(\hat{\gamma}_Q^{(\nu)}(h) - \gamma^{(\nu)}(h)\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \check{\sigma}_h^2),$$
where $\gamma^{(\nu)}(h)$ is the periodic ACF function and $\check{\sigma}_h^2$ is the variance, more details are given in Sarnaglia et al. (2010).
2. The $Q_N^{(\nu)}$ Yule-Walker estimators $(\tilde{\phi}_{\nu,i})_{1\leq i\leq p_\nu,\nu=1,\ldots,\mathcal{S}}$ satisfy $\tilde{\phi}_{\nu,i} - \phi_{\nu,i} = O_P(N^{-1/2})$ for all $i = 1, \ldots, p_\nu$ and ν in $\{1, \ldots, \mathcal{S}\}$.

Recently, Solci et al. (2018) compared the Yule-Walker estimator (YWE), the robust least squares estimator (Shao 2008) and the ACF Q_n estimator $(\hat{\gamma}_Q^{(\nu)}(h)$, denoted here RYWE, in the context of estimating the parameters in PAR models with and without outliers. Their main conclusion is similar to the cases discussed previously, that is, for the case of ARFIMA model $\hat{\gamma}_Q^{(\nu)}(h)$ displayed good performance in estimating the parameters in PAR models, periodic samples with and without outliers. As expected, the YWE estimator performed very poorly with the presence of outliers in the data. One of their simulation results is reproduced in the table below (Table 2) in which $n = 100, 400$ (cycles), $\mathcal{S} = 4$, ϵ_t is a Gaussian white noise process and $\delta = 0.01$ (outlier's probability) and magnitude $\omega = 10$. The results correspond to the mean of 5000 replications.

Table 2. Bias and RMSE for Model 1 and outliers with probability $\delta = 0.01$.

ω	ϵ_t	n	$\phi_{\nu,1}$	YWE		RYWE	
				Bias	RMSE	Bias	RMSE
0	$\mathcal{N}(0,1)$	100	0.9	−0.007	0.077	−0.003	0.103
			0.8	−0.002	0.065	0.004	0.084
			0.7	0.000	0.063	−0.001	0.083
			0.6	−0.005	0.066	−0.003	0.083
		400	0.9	−0.001	0.037	−0.001	0.047
			0.8	−0.001	0.031	0.000	0.038
			0.7	−0.001	0.032	0.001	0.038
			0.6	0.000	0.032	0.000	0.039
7	$\mathcal{N}(0,1)$	100	0.9	−0.181	0.247	0.014	0.120
			0.8	−0.118	0.176	0.012	0.096
			0.7	−0.105	0.157	0.015	0.091
			0.6	−0.097	0.151	0.012	0.091
		400	0.9	−0.183	0.203	0.017	0.055
			0.8	−0.129	0.144	0.012	0.046
			0.7	−0.108	0.124	0.013	0.044
			0.6	−0.103	0.119	0.014	0.043

As an alternative estimator of $\tilde{\phi}_{\nu,i}$, Sarnaglia et al. (2016) proposed the use of M-periodogram function to obtain estimates of the parameters in PARMA models. The estimator is based on the approximated Whittle function suggested in Sarnaglia et al. (2015). Basically, the Whittle M-estimator of PARMA parameters is derived by the ordinary Fourier transform with the non-linear M-regression estimator for periodic processes in the harmonic regression equation that leads to the classical periodogram. The empirical simulation investigation in Sarnaglia et al. (2016) considered the scenarios of periodic time series with presence and absence of additive outliers. Their small sample size investigation leaded to a very promising estimation method under the context of modelling periodic time series with additive outliers and heavy-tailed distributions. The theoretical justification of the proposed estimator is still an open problem and it is now a current research theme of the authors.

Table 3 displays results of a simple simulation example to show the empirical performance of the Whittle M-estimator with the Huber function $\psi(x)$ (Huber 1964) compared to the maximum Gaussian and Whittle likelihood estimators to estimate a PAR(2) model with parameters $\phi_{1,1} = -0.2$, $\phi_{2,1} = -0.5$, $\sigma^2_{1,1} = 1.0$ and $\sigma^2_{2,1} = 1.0$. The sample sizes are $N = n\mathcal{S} = 300, 800$ ($n = 150, 400$, respectively) and the Huber function was used with constant equal to 1.345, which ensure that the M-estimator is 95% as efficient as the least squares estimator for univariate multiple linear models with independent and identically distributed Gaussian white noise. The sample root mean square error (RMSE) was computed over 5000 replications. The PAR(2) model with additive outliers was generated with outlier's probability $\delta = 0.01$ and magnitude $\omega = 10$. The values with "$*$" refer to the RMSE for the contaminated series.

Table 3. Empirical RMSE results for estimating an PAR(2) model.

Method	N	$\phi_{1,1}$	$\sigma^2_{1,1}$	$\phi_{2,1}$	$\sigma^2_{2,1}$
MLE	300	0.067; 0.121*	0.117; 1.366*	0.079; 0.252*	0.111; 1.363*
	800	0.048; 0.101*	0.079; 1.122*	0.046; 0.239*	0.074; 1.253*
WLE	300	0.068; 0.121*	0.117; 1.368*	0.079; 0.252*	0.111; 1.364*
	800	0.048; 0.101*	0.079; 1.122*	0.046; 0.239*	0.074; 1.253*
RWLE	300	0.067; 0.067*	0.147; 0.179*	0.083; 0.089*	0.147; 0.189*
	800	0.051; 0.054*	0.118; 0.149*	0.051; 0.058*	0.108; 0.152*

In the absence of outliers, in general, all estimators present similar behaviour. Relating to the estimation of the variance of the innovations, the MLE and WLE seem to be more precise which is an expected result since the data is Gaussian with zero-mean and these two methods are asymptotically equivalents. The RMSE of the estimators decreases as the sample size increases. When the simulated data has outliers, as an expected result the MLE and WLE estimates are totally corrupted by the atypical observations while the RWLE estimator presents generally accurate estimates. This simple example of simulation leads to the same conclusions of the models discussed previously in which M-regression method was also considered.

The methods discussed above give strong motivation to use the methodology in practical situations in which periodically correlated time series contain additive outliers. For example, Sarnaglia et al. (2010) applied the robust ACF estimator $\hat{\gamma}_Q^{(\nu)}(h)$ to fit a model for the quarterly Fraser River data. Sarnaglia et al. (2016) and Solci et al. (2018) analysed air pollution variables using the robust methodologies discussed in these papers. In the first paper, the authors considered the daily average SO_2 concentrations and, in the second one, it was analysed the daily average PM_{10} concentrations. Both data set were collected at Automatic Air Quality Monitoring Network (RAMQAr) in the Great Vitória Region GVR-ES, Brazil, which is composed by nine monitoring stations placed in strategic locations and accounts for the measuring of several atmospheric pollutants and meteorological variables in the area. In general, the models well fitted the series and all these applied examples revealed outliers effects on the estimates.

5 Proof of Theorem 1

By Propositions 1 and 4 and Example 1 of Wu (2007) the assumptions of Theorem 1 of Wu (2007) hold. Thus,

$$\sqrt{\frac{N}{2}}(F(c)-F(-c))\hat{\boldsymbol{\beta}}_N^{\mathrm{M}}(\lambda_j) \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, \boldsymbol{\Delta}^{(j)}\right), \ N \to \infty ,$$

with

$$\boldsymbol{\Delta}^{(j)} = \sum_{k\in\mathbb{Z}} \mathbb{E}\{\psi(\varepsilon_0)\psi(\varepsilon_k)\}\boldsymbol{\Delta}_k^{(j)} ,$$

where

$$\boldsymbol{\Delta}_k^{(j)} = \lim_{N\to\infty} \frac{2}{N} \sum_{\ell=1}^{N-|k|} \begin{pmatrix} \cos(\ell\lambda_j) \\ \sin(\ell\lambda_j) \end{pmatrix} (\cos((\ell+k)\lambda_j) \ \sin((\ell+k)\lambda_j)) .$$

Observe that

$$\begin{aligned}\boldsymbol{\Delta}_k^{(j)} &= \lim_{N\to\infty} \frac{2}{N} \sum_{\ell=1}^{N-|k|} \begin{pmatrix} \frac{\cos(k\lambda_j)+\cos((2\ell+k)\lambda_j)}{2} & \frac{\sin(k\lambda_j)+\sin((2\ell+k)\lambda_j)}{2} \\ \frac{-\sin(k\lambda_j)+\sin((2\ell+k)\lambda_j)}{2} & \frac{\cos(k\lambda_j)-\cos((2\ell+k)\lambda_j)}{2} \end{pmatrix} \\ &= \begin{pmatrix} \cos(k\lambda_j) & \sin(k\lambda_j) \\ -\sin(k\lambda_j) & \cos(k\lambda_j) \end{pmatrix} + \lim_{N\to\infty} \frac{2}{N} \sum_{\ell=1}^{N-|k|} \begin{pmatrix} \frac{\cos((2\ell+k)\lambda_j)}{2} & \frac{\sin((2\ell+k)\lambda_j)}{2} \\ \frac{\sin((2\ell+k)\lambda_j)}{2} & \frac{-\cos((2\ell+k)\lambda_j)}{2} \end{pmatrix} .\end{aligned}$$

By observing that

$$\begin{aligned}&\frac{1}{N}\sum_{\ell=1}^{N-|k|} \cos((2\ell+k)\lambda_j) = \frac{\cos(k\lambda_j)}{N}\sum_{\ell=1}^{N-|k|}\cos(2\ell\lambda_j) + \frac{\sin(k\lambda_j)}{N}\sum_{\ell=1}^{N-|k|}\sin(2\ell\lambda_j) \\ &= \frac{\cos(k\lambda_j)}{N}\cos(\lambda_j(N-|k|-1))\frac{\sin(\lambda_j(N-|k|))}{\sin(\lambda_j)} + \frac{\sin(k\lambda_j)}{N}\sin(\lambda_j(N-|k|-1))\frac{\sin(\lambda_j(N-|k|))}{\sin(\lambda_j)}\end{aligned}$$

tends to zero as N tends to infinity and that the same holds for $N^{-1}\sum_{\ell=1}^{N-|k|}\sin(2\ell+k)$, this concludes the proof.

Acknowledgements. V. A. Reisen gratefully acknowledges partial financial support from FAPES/ES, CAPES/Brazil and CNPq/Brazil and CentraleSupélec. Màrton Ispàny was supported by the EFOP-3.6.1-16-2016-00022 project. The project is cofinanced by the European Union and the European Social Fund. Paulo Roberto Prezotti Filho and Higor Cotta are Ph.D students under supervision of V. A. Reisen and P. Bondon. The authors would like to thank the referee for the valuable suggestions.

References

Basawa I, Lund R (2001) Large sample properties of parameter estimates for periodic ARMA models. J Time Ser Anal 22(6):651–663

Croux C, Rousseeuw PJ (1992) Time-efficient algorithms for two highly robust estimators of scale. Comput Stat 1:1–18

Fajardo F, Reisen VA, Cribari-Neto F (2009) Robust estimation in long-memory processes under additive outliers. J Stat Plan Infer 139(8):2511–2525

Fajardo FA, Reisen VA, Lévy-Leduc C, Taqqu M (2018) M-periodogram for the analysis of long-range-dependent time series. Statistics 52(3):665–683

Fox AJ (1972) Outliers in time series. J R Stat Soc 34(3):350–363

Geweke J, Porter-Hudak S (1983) The estimation and application of long memory time series model. J Time Ser Anal 4(4):221–238

Huber P (1964) Robust estimation of a location parameter. Ann Math Stat 35:73–101

Koul HL, Surgailis D (2000) Second order behavior of M-estimators in linear regression with long-memory errors. J Stat Plan Infer 91(2):399–412

Lévy-Leduc C, Boistard H, Moulines E, Taqqu M, Reisen VA (2011) Robust estimation of the scale and the autocovariance function of Gaussian short and long-range dependent processes. J Time Ser Anal 32(2):135–156

Ma Y, Genton M (2000) Highly robust estimation of the autocovariance function. J Time Ser Anal 21(6):663–684

McLeod AI (1994) Diagnostic checking periodic autoregression models with application. J Time Ser Anal 15(2):221–33

Reisen VA, Lévy-Leduc C, Taqqu MS (2017) An M-estimator for the long-memory parameter. J Stat Plan Infer 187:44–55

Rousseeuw PJ, Croux C (1993) Alternatives to the median absolute deviation. J Am Stat Assoc 88(424):1273–1283

Sarnaglia AJQ, Reisen VA, Bondon P (2015) Periodic ARMA models: application to particulate matter concentrations. In: 23rd European Signal Processing Conference, pp 2181–2185

Sarnaglia AJQ, Reisen VA, Bondon P, Lévy-Leduc C (2016) A robust estimation approach for fitting a PARMA model to real data. In: IEEE Statistical Signal Processing Workshop, pp 1–5

Sarnaglia AJQ, Reisen VA, Lévy-Leduc C (2010) Robust estimation of periodic autoregressive processes in the presence of additive outliers. J Multivar Anal 101(9):2168–2183

Shao Q (2008) Robust estimation for periodic autoregressive time series. J Time Ser Anal 29(2):251–263

Solci CC, Reisen VA, Sarnaglia AJQ, Pascal B (2018) Empirical study of robust estimation methods for PAR models with application to PM_{10} data. Commun Stat 15 (in press)

Wu WB (2007) M-estimation of linear models with dependent errors. Ann Stat 35(2):495–521

Author Index

F. Chaari et al. (Eds.): CSTA 2017, ACM 16, p. 225, 2020.
https://doi.org/10.1007/978-3-030-22529-2