Monograph of the Project
POIG.01.03.01-02-002/08

Krzysztof Burnecki

IDENTIFICATION, VALIDATION AND PREDICTION OF FRACTIONAL DYNAMICAL SYSTEMS

INNOVATIVE ECONOMY
NATIONAL COHESION STRATEGY

Wrocław University of Technology

EUROPEAN UNION
EUROPEAN REGIONAL DEVELOPMENT FUND

Krzysztof Burnecki

Identification, validation and prediction of fractional dynamical systems

Oficyna Wydawnicza Politechniki Wrocławskiej
Wrocław 2012

Reviewers:
Aleksander Weron
Wojbor A. Woyczyński

Cover design:
Waldemar Grzebyk
Krzysztof Burnecki

Typesetting:
Krzysztof Burnecki

Printed in the the camera ready form

OFICYNA WYDAWNICZA POLITECHNIKI WROCŁAWSKIEJ
Wybrzeże Wyspiańskiego 27, 50-370 Wrocław
http://www.oficyna.pwr.wroc.pl
e-mail: oficwyd@pwr.wroc.pl

ISBN 978-83-7493-734-4

Drukarnia Oficyny Wydawniczej Politechniki Wrocławskiej. Order No. 1160/2012.

Contents

1

Introduction

In this monograph we propose a systematic methodology how to identify and validate three mechanisms leading to fractional dynamics, namely fractional Brownian motion (FBM), fractional stable motion (FSM) and fractional autoregressive integrated moving average (FARIMA) process. This methodology is based on well-known and novel statistical tools for identification and validation. We claim that the FARIMA process is a universal model for fractional dynamics in engineering and basic sciences. The monograph has been written under the Project: "Detectors and sensors for measuring factors hazardous to environment — modeling and monitoring of threats" (POIG.01.03.01-02-002/08-00). We show that different risk factors affecting the environment can be modelled by fractional dynamical systems.

Fractional dynamical systems are related to the concept of fractional dynamic equations. This is an active field of study in physics, mechanics, mathematics, and economics investigating the behaviour of objects and systems that are described by using differentiation of fractional orders. The celebrated fractional Fokker-Planck equation (FFPE), describing anomalous diffusion in the presence of an external potential $V(x)$, is given by the following formula

$$\frac{\partial w(x,t)}{\partial t} = {}_0D_t^{1-\alpha}\left[\frac{\partial}{\partial x}\frac{V'(x)}{\eta} + K\frac{\partial^2}{\partial x^2}\right] w(x,t). \tag{1.1}$$

It was derived explicitly in [152], where methods of its solution were introduced and for some special cases exact solutions were calculated. Here, the operator ${}_0D_t^{1-\alpha}$, $\alpha \in (0,1)$, is the fractional derivative of the Riemann-Liouville type [185]. It is known that ${}_0D_t^{1-\alpha}$ introduces a convolution integral with a slowly decaying power-law kernel, which is typical for memory effects in complex systems [191]. In equation (1.1), $w(x,t)$ denotes the probability density function (PDF) and the prime (') stands for the derivative with respect to the space coordinates relating the force $F(x)$ and the potential through $F(x) = -V'(x)$. The constant K denotes the anomalous diffusion coefficient, whereas η is the generalized friction constant. For $\alpha \to 1$, equation (1.1) becomes the ordinary Fokker-Planck equation. The FFPE

describes subdiffusion in accordance with the mean-squared displacement in the force-free limit [36] and it obeys some generalized fluctuation-dissipation theorem. Moreover, a generalization of the Einstein-Stokes-Smoluchowski relation $K = k_B T/\eta$ connects the generalized friction and diffusion coefficients [152].

Fractional systems are now recognized in more applied domains. For example the generic term anomalous diffusion (this means described by fractional equations in contrast to the classical Brownian diffusion) receives currently 11 000 citations per year [116]. In the engineering sciences, in which fractional dynamics has a long-standing history in the description of viscoelastic materials or in process control, it is popular as ever [116]. Derivatives and integrals of fractional orders are used to describe objects that can be characterized by long (power-like) memory or self-similarity.

In recent years, there has been a great interest in long-range dependent and self-similar processes, in particular FBM, FSM and FARIMA. This importance can be judged, for example, by a very large number of publications having one of these notions in the title, in areas such as finance and insurance [133, 153, 32, 132], econometrics [178], telecommunication [14, 4, 222, 112, 196], hydrology [166], climate studies [210], linguistics [2], DNA sequencing [114] or medicine [172]. Long-range dependent and self-similar processes also appear widely in other areas like biophysics [201, 138, 48, 115, 33, 42] or astronomy [39, 192]. These publications address a great variety of issues: detection of long memory and self-similarity in the data, statistical estimation of parameters of long-range dependence and self-similarity, limit theorems under long-range dependence and self-similarity, simulation of long memory and self-similar processes, and many others [187].

FARIMA series are fractionally differenced autoregressive moving average (ARMA) series [84]. The model generalizes three broad classes of time series, namely the autoregressive (AR) models, the integrated (I) models, and the moving average (MA) models [13, 30]. In order to illustrate the idea behind the FARIMA time series, let us concentrate on FARIMA$(1, d, 1)$. The basic building blocks of FARIMA$(1, d, 1)$ model are AR(1) process: $X_t = \phi_1 X_{t-1} + Z_t$, and MA(1) process: $X_t = Z_t - \theta_1 Z_{t-1}$, where ϕ_1, θ_1 are constant coefficients and Z_t is the noise. AR(1) is a causal or future-independent function of noise and stands for the regression. The explanatory variable is the observation immediately prior to our current observation. In order to get an idea about the role of MA part let us concentrate on the case when Z_t is a white noise. It appears that if X_t is a stationary 1-correlated time series, i.e. X_s and X_t are independent whenever $|t-s| > 1$, in contrast to an independent and identically distributed (i.i.d.) sequence which is zero-dependent, then it can be represented as the MA(1) process [30]. The dependence (correlation in this case) is only one lag long and it's intensity is fully controlled by the parameter θ_1. Hence, ARMA models introduce short memory of the process.

The time series X_t is an ARMA$(1,1)$ process if it is stationary and satisfies (for every t) linear difference equation with constant coefficients:

$$X_t - \phi_1 X_{t-1} = Z_t - \theta_1 Z_{t-1},$$

where $t = 0, \pm 1, \ldots$

A stationary solution of ARMA$(1,1)$ equation exists if and only if $\phi_1 \neq \pm 1$. If $|\phi_1| < 1$, then the unique stationary solution exists and is causal, since X_t can be expressed in terms of the current and past values Z_s, $s \leqslant t$. Otherwise, if $|\phi_1| > 1$, then the solution is not causal since X_t is then a function of Z_s, $s \geqslant t$ [30].

Let us observe that ARMA models are short memory stationary processes since their autocorrelation function converges exponentially to zero as the lag goes to infinity. They provide a general framework for studying stationary processes and simple prediction algorithms to describe future behaviour of the dynamics. For example, linear prediction of the causal ARMA$(1,1)$ process looks as follows:

$$\hat{X}_{n+1} = \phi_1 X_n + \theta_{n1}(X_n - \hat{X}_n),$$

where $n \geqslant 1$ and θ_{n1} can be calculated recursively from ϕ_1 and θ_1 [30].

A generalization of this class, which incorporates a wide range of nonstationary series, is provided by ARIMA processes, i.e. processes that reduce to ARMA processes when differenced finitely many times.

Finally, stationary processes with much more slowly decreasing autocorrelation function, i.e. with long (power-like) memory are called FARIMA processes. The FARIMA$(1,d,1)$ process X_t is represented by the following fractional difference equation with constant coefficients:

$$(1-B)^d (X_t - \phi_1 X_{t-1}) = Z_t - \theta_1 Z_{t-1}, \tag{1.2}$$

where $t = 0, \pm 1, \ldots$

The fractional difference operator $(1-B)^d$ is defined in Section 4.3. If the noise Z_t consists of i.i.d. random variables belonging to the domain of attraction of the α-stable law, then $d < 1 - 1/\alpha$ and $0 < \alpha \leqslant 2$ [119]. In this case the noise may be either Gaussian, non-Gaussian with finite variance or it may have infinite variance. For finite variance case, apart from the Gaussian law, the typical choices are exponential, log-normal and normal-inverse Gaussian distributions [42, 35]. For infinite variance variables one may consider, for example, symmetric and skewed stable, and Pareto distributions [192]. For the description of general FARIMA(p,d,q) processes, where p and q are positive integers, see Section 4.3.

First of all, from the physical point of view, it is known that FARIMA is a discrete time analogue of fractional Langevin equation (FLE) that takes into

account the memory parameter d [135]. It also includes other popular models of fractional dynamics as FBM and FLM, which are limiting cases of aggregated FARIMA [195]. FARIMA processes are asymptotically self-similar with the parameter $H = d + 1/\alpha$. Brownian motion (BM) and stable motion (SM) correspond, in the limit sense [195], to FARIMA$(0, 0, 0)$ with Gaussian and α-stable noise, respectively. Similarly, FSM corresponds to FARIMA$(0, d, 0)$ with $d = H - 1/\alpha$, where H is the self-similarity parameter. For $\alpha = 2$ we recognize the FBM model. In contrast to FBM and FSM, it can also describe different light- and heavy-tailed noises and a short (exponential) dependence.

FARIMA process, similarly as increments of FBM, exhibits long-time correlations. In the Gaussian case, i.e. when $\alpha = 2$, the rate of decay of the autocovariance function $\mathrm{Cov}(n) = \mathrm{E}X_nX_0 - \mathrm{E}X_nEX_0$ is n^{2d-1}. Moreover, for $d > 0$ we have $\sum_{n=0}^{\infty} |\mathrm{Cov}(n)| = \infty$. This serves as a classical definition of long memory [13]. However, we call here all FARIMA processes with $d \neq 0$ long-range dependent to distinguish between ARMA models with exponentially decaying autocorrelation function and FARIMA models with the power-law decay. For $\alpha < 2$ the covariance does not exist and one has to replace it, e.g., with the codifference (see [188]). The codifference of the FARIMA process was studied in [119], where it was proved that for $d > 1 - 2/\alpha$ FARIMA possesses long-term dependence in the classical sense.

Summarizing, modelling with FARIMA time series allows one to take into account different light and heavy tails. The FARIMA processes offer also a lot of flexibility in modelling of long (power-like) and short (exponential) dependences by choosing the memory parameter d and appropriate autoregressive and moving average coefficients in equation (1.2). Hence, they can be better tailored to empirical data.

We would like to underline that this monograph is aimed at parametric system identification only. Nonparametric system identification for nonlinear systems (like Hammerstein or Wiener) was widely studied in the literature. We refer here to [87], where theoretical foundations of nonparametric systems identification for nonlinear block-oriented systems were presented. For the theory of nonparametric regression, see [89, 91]. The nonparametric regression approach to identifying the Hammerstein system was proposed in [86]. The distribution-free properties of the kernel regression estimate were studied in [85]. The use of wavelets in the statistical inference is examined in [92]. Wavelet expansions have been widely applied in system identification [169, 93, 94, 95, 98, 193, 97, 203, 204]. Nonlinear system identification by Haar wavelets was systematically studied in [202]. The multidimensional nonlinear Schur parametrization problem for higher-order and non-Gaussian stochastic signals was studied in the series of papers [227, 229, 228], where a generalization of the linear prediction/innovations fil-

ter problem was solved. Nonparametric identification algorithms based on ordered input observations were studied [81, 180, 181, 182]. Identification of discrete Hammerstein systems was done in [123, 124]. Comprehensive studies of semiparametric Hammerstein-Wiener models were given in [96, 170].

Nonparametric methods are beyond the scope of this monograph; nevertheless, we mention here that nonparametric and parametric methods are supposed to be applied in different situations. The two approaches do not compete with each other since they are designed to be applied in quite different situations. The first is used when a priori information is nonparametric, that is, when we wish to recover an infinite-dimensional object with underlying assumptions as weak as possible. Clearly, in such a case, parametric methods can only approximate, but not estimate, the unknown characteristics [87]. However, when the information is parametric, parametric methods provide the natural choice. This is the situation studied here for FARIMA models.

Guide to the chapters

The monograph is structured as follows. In Chapter 2 we present the main highlights of the Project. Organization of the Project and covered risk areas are described in detail. Three problems are introduced which are tied to the environmental protection risk area: electromagnetic field intensity, solar flare variability and molecular biology. Chapter 3 provides basic characteristics of both light- and heavy-tailed distributions that are most commonly used as underlying probability laws in fractional dynamical systems. Parameter estimators and methods of simulation are discussed. The stable distribution which is a limiting law for a large amount of distributions is studied in detail. A new algorithm of distinguishing between Gaussian and non-Gaussian stable distributions with the stability parameter α close to 2 is presented. In Chapter 4 three models for fractional dynamical systems are considered, namely FBM, FSM and FARIMA time series. The former two models are self-similar and their increments form long-range dependent processes. The latter discrete-time FARIMA process is stationary, and when aggregated, in the limit, it converges to either FBM or FSM. In this sense it generalizes both models. In contrast to them, it allows for different light- and heavy-tailed distributions, and, as a stationary process, provides tools for calculating predictions. We choose this flexible model for further considerations. We present a new estimation algorithm for the FARIMA parameters in the case of noise belonging to the attraction of the stable law for both positive and negative memory parameters. A new algorithm for fractional differencing is also provided. Chapter 5 is devoted to the description of different identification and validation tools. They are introduced for testing of stationarity and ergodicity, distribution type, self-similarity and long memory. In Sections 5.3.2 and 5.4.3 we present

two new methods of estimation self-similarity index and memory parameter with the help of p-variation and sample mean-squared displacement respectively. In Chapter 6 we discuss three case studies related to the problems introduced in Chapter 2. They are successfully resolved by means of the universal FARIMA time series model with various noises. In Chapter 7 we present a novel algorithm for identification and validation of fractional dynamical systems. It was implemented in the ASP module which was written in Matlab. We present a specification of the module and a case study on sample data collected by a telecommunication group within the Project.

2

Detectors and sensors for measuring factors hazardous to environment

Project: "Detectors and sensors for measuring factors hazardous to environment – modeling and monitoring of threats" is financed by the European Union via the European Regional Development Fund and the Polish state budget, within the framework of the Operational Programme Innovative Economy 2007–2013 (POIG.01.03.01-02-002/08-00) [1]. The project is carried out by Wrocław University of Technology as the beneficiary at the Faculty of Electronics, along with four other science and research institutions: Krakow AGH University of Science and Technology, Wroclaw Medical University, Institute of Immunology and Experimental Therapy, and Institute of Low Temperature and Structure Research of the Polish Academy of Sciences in Wroclaw as consortium members.

The general objective of the Project is to increase the role of science in general economy, through research and development activities being performed in priority areas for social and economical development of Poland. The Project carried out by a multidisciplinary scientific research team, allows for improved application of modern solutions for a balanced development of knowledge-based economy. In the scientific and technical sense, the direct objective of the project is to develop innovative technical and implementation solutions for environment hazard measurement and control systems.

This is a multidisciplinary project consisting of: design and development of prototype detectors and sensors, development of measurement methods and construction of a demonstration platform of an environment hazard monitoring system for a selected area, using detectors and sensors developed under this project, as well as other commercially available or adapted solutions. The research teams under this project were divided into two theme groups: „Detectors and sensors" and „Teleinformatics". The first team carries out research tasks relevant to the following five themes:

- electromagnetic field sensors,
- acoustic sensors,

- gas sensors,
- industrial waste sensors,
- biological sensors.

The second team carries out research tasks relevant to the following five teleinformatics-related themes:

- interfaces,
- teletransmission, network, protocols,
- databases,
- statistics and forecasts,
- on-line analysis.

The project defines selected risk areas. These can relate to different subjects like: environmental protection, medicine, food industry, chemical industry, farming, transport, mining, intelligent buildings, airports or other public facilities. It is assumed that each such area will use certain detectors and sensors, which will monitor the subset of all factors important to that given subject. The teletransmission systems are subject to special requirements as the accuracy of automatic or expert system decisions heavily rely on their performance and the methods used for internal and external communications. The research objectives include development of new detectors and sensors as well as measurement methodologies, and they will be conducted in parallel, independent of all of the other tasks.

This is an applied project, having real-life practical application. The project will result in a pilot measurement and teletransmission system, which will provide a practical source of environmental hazard information data for a selected area.

Author's work in this project is related to the project's theme "Statistics and forecasts". In Sections 2.1 – 2.3 we introduce three problems which are tied to the environmental protection risk area. In Section 2.1 we present a problem of constant and preferably automated monitoring of the electromagnetic field with the ultimate objective of predicting alarming trends that could threaten the society. In Section 2.2 we address the issue of high solar activity. Prediction of solar flare events is an actual problem as these disturbances can pose serious threats to man-made spacecrafts, disrupt electronic communication channels and can even set up huge electrical currents in power grids which can lead to blackouts. In Section 2.3 we present a problem of identifying basic mathematical mechanisms behind recent nanoscale single-molecule biophysics experiments. Two types of data are analysed: the data describing the motion of individual fluorescently labelled mRNA molecules inside live *E. coli* cells and the data describing transient anomalous diffusion of telomeres in the nucleus of living human cells (U2OS cancer). We note that *E. coli* cells were intensively studied in the Project in the context of biological sensors. Our results described here answer few open problems raised

by experimenters from molecular biology labs in Bar-Ilan University, Heidelberg University and Princeton University. Solutions to the three problems by means of fractional dynamics are presented in Chapter 6.

2.1 Case study: Electromagnetic field measurements

The progress of civilization is marked by intensive growth of emissions to the environment. The intensification of the electromagnetic field (EMF) is a result of increasing energy demand (more and more devices are powered by electric energy) and the popularization of radio communication systems, especially data transmission [17, 18, 34, 19, 16]. In the past, the dominant EMF sources of broadcasting systems used to be radio and TV transmitters, usually located at dedicated sites. In modern duplex systems, however, EMF sources can be found both in dense networks of base stations and within mobile terminals. The latter are usually low power sources, but exist in large quantity. Despite the fact that in most cases the measured EMF intensity is well below hazardous levels, there is a need for constant and preferably automated monitoring of this exposure with the ultimate objective of predicting alarming trends that could threaten the society [52].

The electromagnetic field is characterized by a number of parameters. However, when considering the influence of EMF on the biosphere, the most important information is the intensity of particular EMF elements in defined frequency ranges. By analyzing frequency bands, it is possible to identify the EMF source class, for example, in radio communication: radio diffusion and mobile radio communication, etc. This approach prefers broadband measurements or single band measurements using meters with passive broadband probes, which allow measurements in direct surroundings of primary and secondary EMF sources. Broadband meters are widely used in measurements related to work safety and environment protection.

In Section 6.1 we analyse a set of Universal Mobile Telecommunications System (UMTS) data. The electromagnetic field intensity was measured in Wroclaw in an urban area in 2011. We show that a long memory FARIMA model describes the data well. We believe that the proposed FARIMA methodology can be also applied for other telecommunication data. This would allow to better predict hazardous levels of the field intensity.

2.2 Case study: High solar flare activity

Understanding the long-term solar variability is one of the most important problems in solar physics. Solar activity has clearly a periodic character, but solar

cycles (taken separately) are different in form, amplitude and length. Solar flare events are random in time and strength. Their prediction is an actual problem as these disturbances can pose serious threats to man-made spacecrafts, disrupt electronic communication channels and can even set up huge electrical currents in power grids [60]. It is enough to remind about serious problems with the spacecrafts GOES, Deutsche Telecom, Telstars 401, etc. Satellite operators would be glad to escape the unhappy surprise, and mission planners are compelled to take into account the future space weather forecast. The cost to the airline industry arose as planes were re-routed to lower altitudes, burning more fuel in force of atmospheric drag. Now the space weather forecast has the considerable importance as well as the weather forecast on the Earth.

Within the last few years theoretical models of solar dynamo have been developed to explain various aspects of the solar activity. In [68] the authors have made the first attempt of using a theoretical dynamo model to predict the strength of the upcoming cycle 24. They have shown that this cycle will be the strongest in 50 years. But later it was pointed out in [59] that some assumptions in the Dikpati–Gilman model are unjustified. On the contrary, their model, based on the earlier work of [160], predicts that the cycle 24 will be weaker than the 23-rd. The key problem here is the following: the dominant processes like the magnetic field advection and toroidal field generation by differential rotation are fairly regular during the rising phase of a cycle from a minimum to a maximum, and hence a good knowledge of magnetic configurations during a minimum would enable a good theoretical model to predict the next maximum reliably. However, the dominant process in the declining phase of a cycle contains the poloidal field generation by the Babcock-Leighton mechanism which involves randomness (primary cause of solar cycle fluctuations) and cannot be predicted in advance by any deterministic model. Thus the prediction methods of flare activity should be based on statistical analysis of temporal and spatial features of the experimental evidence. For example, the distribution of the X-ray flares from 1987 to 1992 with respect to helio longitude was studied in [129]. They have shown that the flares were not uniformly distributed in longitude. The analysis of temporal features of the X-flare statistics basically has been related with the waiting-time distributions (see, for example, [21, 128, 220, 221].

Our aim is to present a proper statistical model for analyzing and predicting soft X-ray solar emission activity in the period of solar maximum. As a candidate we consider in Section 6.2 the FARIMA model with Pareto noise.

2.3 Case study: Single particle tracking dynamics in molecular biology

A phenomenon observed in recent nanoscale single-molecule biophysics experiments is subdiffusion, which largely departs from the classical Brownian diffusion theory since the mean-squared displacement (MSD) is sublinear. In the case of classical diffusion, the second moment is linear in time, whereas anomalous diffusion processes exhibit distinct deviations from this fundamental property: $\mathrm{E}X^2(t) \sim t^a$, where for $0 < a < 1$ is subdiffusive and for $a > 1$ is superdiffusive [152]. Determining anomalous diffusion in crowded fluids, e.g., in the cytoplasm of living cells [212, 208, 83, 88], or in more controlled in vitro experiments [167, 201, 108], is a challenging problem. Subdiffusion can have three different physical origins: obstruction, binding and viscoelasticity [107].

The most popular theoretical models that are commonly employed are: continuous-time random walk (CTRW) [83, 107], obstructed diffusion (OD) [88, 99], FBM [201, 115, 38], FSM [48], and fractional Langevin equation (FLE) [108, 31, 115]. These models can be divided into two categories: with short memory (CTRW, OD) and fractional with long (power-like) memory (FBM, FSM and FLE).

The FARIMA process unifies the latter category [45]. First of all, from the physical point of view, it is known that FARIMA is a discrete time analogue of FLE that takes into account the memory parameter d [135]. It also includes other popular models of subdiffusive dynamics as FBM and FSM, which are limiting cases of aggregated FARIMA. In contrast to FBM and FSM, it can also describe different light- and heavy-tailed noises and a short (exponential) dependence. FARIMA with finite variance noises, similarly as increments of FBM, exhibits long-time (power-like) correlations. This results in an anomalous diffusion exponent MSD. Therefore, FARIMA can be associated with the third origin of anomalous diffusion, i.e. viscoelasticity [88].

In Section 6.3 we look for a unified approach to the fractional subdiffusive dynamics in biological cells by means of FARIMA processes. To this end we investigate three sets of data: the data describing the motion of individual fluorescently labelled mRNA molecules inside live *E. coli* cells presented by Golding and Cox [83], the data extracted from the video movie of Golding and Cox and the data describing transient anomalous diffusion of telomeres in the nucleus of living human cells (U2OS cancer) presented in [115]. In the latter context let us mention that the 2009 Nobel Prize in Physiology or Medicine was awarded to Elizabeth H. Blackburn, Carol W. Greider and Jack W. Szostak for the discovery of how chromosomes are protected by telomeres and the enzyme telomerase [198, 199, 200].

3

Light and heavy-tailed distributions

In this chapter we discuss basic characteristics of both light- and heavy-tailed distributions that are most commonly used as underlying probability laws in fractional dynamical systems, namely the exponential, Gaussian, log-normal, normal-inverse Gaussian, Pareto and stable. We present their parameter estimators and methods of simulation. The stable distribution which is a limiting law for heavy-tailed distributions is studied in detail in Section 3.6. A new algorithm of distinguishing between Gaussian and non-Gaussian stable distributions with the stability parameter close to 2 is presented.

3.1 Exponential distribution

Definition 3.1 *A random variable X is said to have an exponential distribution if it has density:*

$$f(x) = \beta e^{-\beta x}, \qquad x > 0. \tag{3.1}$$

The parameter $\beta > 0$ is called the intensity. The Laplace transform of (3.1) is

$$L(t) := \int_0^\infty e^{-tx} f(x)dx = \frac{\beta}{\beta + t}, \qquad t > -\beta,$$

yielding the general formula for the k-th raw moment

$$m_k := (-1)^k \frac{\partial^k L(t)}{\partial t^k}\Big|_{t=0} = \frac{k!}{\beta^k}.$$

The mean and variance are thus β^{-1} and β^{-2}, respectively. The maximum likelihood estimator (equal to the method of moments estimator) for β is given by:

$$\hat{\beta} = \frac{1}{\hat{m}_1},$$

where

$$\hat{m}_k = \frac{1}{n}\sum_{i=1}^{n} x_i^k, \tag{3.2}$$

is the sample k-th raw moment.

Exponential random variable can be very easily simulated by using the standard inverse transform method [179].

3.2 Gaussian distribution

Definition 3.2 *A random variable X is said to have a Gaussian (normal) distribution if it has density*

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right\}, \qquad -\infty < x < \infty.$$

The parameter μ is the mean or expectation (location of the peak) and σ^2 is the variance. σ is known as the standard deviation.

The maximum likelihood estimator (equal to the method of moments estimator) for μ is given by:

$$\hat{\mu} = \bar{x} = \hat{m}_1.$$

The maximum likelihood estimator (equal to the method of moments estimator) for σ^2 is given by:

$$\hat{\sigma}^2 = \sum_{i=1}^{n} (x_i - \bar{x})^2 .$$

The problem of simulating sequences of normal random variables results from the fact that there are no analytic expressions for the inverse F^{-1} of the cumulative distribution function (CDF). The most commonly approaches used in the literature are: Box–Muller method [26], Marsaglia polar method which is a modification of the Box–Muller method algorithm, which does not require computation of trigonometric functions [144], and the ziggurat algorithm [145]. The latter has been a default method of generating Gaussian random variables in Matlab since version 7.4.

3.3 Log-normal distribution

Consider a random variable Y which has the normal distribution. Let $X = e^Y$ so that $Y = \log X$. The distribution of X is called log-normal.

Definition 3.3 *A random variable X is said to have a log-normal distribution if its density is given by*

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma x} \exp\left\{-\frac{1}{2}\frac{(\log x - \mu)^2}{\sigma^2}\right\}, \qquad x > 0,$$

where $\sigma > 0$ is the scale and $-\infty < \mu < \infty$ is the location parameter.

For large σ its tail is (semi-)heavy – heavier than the exponential but lighter than power-law. For small σ the log-normal resembles a normal distribution. The log-normal CDF is given by:

$$F(x) = \Phi\left(\frac{\log x - \mu}{\sigma}\right), \qquad x > 0,$$

where $\Phi(\cdot)$ is the standard normal (with mean 0 and variance l) distribution function. The k-th raw moment m_k of the log-normal variate can be easily derived using results for normal random variables:

$$m_k = \mathrm{E}\left(X^k\right) = \mathrm{E}\left(e^{kY}\right) = M_Y(k) = \exp\left(\mu k + \frac{\sigma^2 k^2}{2}\right),$$

where $M_Y(z)$ is the moment generating function of the normal distribution. In particular, the mean and variance are

$$\begin{aligned} \mathrm{E}(X) &= \exp\left(\mu + \frac{\sigma^2}{2}\right), \\ \mathrm{Var}(X) &= \left\{\exp\left(\sigma^2\right) - 1\right\}\exp\left(2\mu + \sigma^2\right), \end{aligned}$$

respectively. For both standard parameter estimation techniques the estimators are known in closed form. The method of moments estimators are given by:

$$\begin{aligned} \hat{\mu} &= 2\log\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) - \frac{1}{2}\log\left(\frac{1}{n}\sum_{i=1}^{n} x_i^2\right), \\ \hat{\sigma}^2 &= \log\left(\frac{1}{n}\sum_{i=1}^{n} x_i^2\right) - 2\log\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right), \end{aligned}$$

while the maximum likelihood estimators by:

$$\begin{aligned} \hat{\mu} &= \frac{1}{n}\sum_{i=1}^{n} \log(x_i), \\ \hat{\sigma}^2 &= \frac{1}{n}\sum_{i=1}^{n} \left\{\log(x_i) - \hat{\mu}\right\}^2. \end{aligned}$$

Finally, the generation of a log-normal variate is straightforward. We simply have to take the exponent of a normal variate.

3.4 Normal-inverse Gaussian distribution

Definition 3.4 *A random variable X is said to have a normal-inverse Gaussian (NIG) distribution if it has density*

$$f(x) = \frac{\alpha\delta}{\pi} e^{\delta\sqrt{\alpha^2-\beta^2}+\beta(x-\mu)} \frac{K_1(\alpha\sqrt{\delta^2+(x-\mu)^2})}{\sqrt{\delta^2+(x-\mu)^2}}, \quad -\infty < x < \infty. \tag{3.3}$$

The NIG distribution was introduced in [6]. It is described by four parameters $(\alpha, \beta, \delta, \mu)$, where α stands for tail heavyness, β for asymmetry, δ is the scale parameter, and μ is the location. The normalizing constant $\mathrm{K}_\lambda(t)$ in formula (3.3) is the modified Bessel function of the third kind with index λ, also known as the MacDonald function. It is defined as:

$$\mathrm{K}_\lambda(t) = \frac{1}{2}\int_0^\infty x^{\lambda-1} e^{-\frac{1}{2}t\left(x+x^{-1}\right)} dx, \qquad t > 0.$$

The normal-inverse Gaussian distributions form a subclass of the generalised hyperbolic distributions. The fact that there is a simple expression for the moment generating function implies that simple expressions for all moments are available. Like for the hyperbolic law the calculation of the density function is straightforward, but the distribution function has to be numerically integrated from (3.3).

At the expense of four parameters, the NIG distribution is able to model symmetric and asymmetric distributions with possibly long tails in both directions. Its tail behaviour is often classified as "semi-heavy", i.e. the tails are lighter than those of stable laws, but much heavier than Gaussian [24]. However, if we let $\alpha \to 0$ the NIG distribution converges to the Cauchy distribution (with location parameter μ and scale parameter δ), which exhibits extremely heavy tails. Obviously, the NIG distribution may not be adequate to deal with cases of extremely heavy tails such as those of Pareto or non-Gaussian stable laws. However, empirical experience suggests excellent fits of the NIG law to financial data [113, 211, 216, 24] and it has been recently fitted to biological data [42].

Parameters of the NIG distribution can be obtained via maximum likelihood method by numerical maximization of the likelihood function (for the usage of expectation-maximization algorithm in this this context see [113]). The most

natural way of simulating NIG variables is derived from the property that they can be represented as normal variance-mean mixtures [24].

3.5 Pareto distribution

Suppose that a variate X has (conditional on β) an exponential distribution with intensity β (i.e. with mean β^{-1}). Further, suppose that β itself has a gamma distribution. The unconditional distribution of X is a mixture and is called the Pareto distribution. Moreover, it can be shown that if X is an exponential random variable and Y is a gamma random variable, then X/Y is a Pareto random variable.

Definition 3.5 *A random variable X is said to have a Pareto distribution if it has density*

$$f(x) = \frac{\alpha\lambda^\alpha}{(\lambda + x)^{\alpha+1}}, \qquad x > 0.$$

Clearly, the shape parameter α and the scale parameter λ are both positive. The k-th raw moment:

$$m_k = \lambda^k k! \frac{\Gamma(\alpha - k)}{\Gamma(\alpha)},$$

exists only for $k < \alpha$. The mean and variance are thus:

$$\begin{aligned} \mathrm{E}(X) &= \frac{\lambda}{\alpha - 1}, \\ \mathrm{Var}(X) &= \frac{\alpha\lambda^2}{(\alpha-1)^2(\alpha-2)}, \end{aligned}$$

respectively. Note, that the mean exists only for $\alpha > 1$ and the variance only for $\alpha > 2$. Hence, the Pareto distribution has very thick (or heavy) tails. The method of moments estimators are given by:

$$\begin{aligned} \hat{\alpha} &= 2\frac{\hat{m}_2 - \hat{m}_1^2}{\hat{m}_2 - 2\hat{m}_1^2}, \\ \hat{\lambda} &= \frac{\hat{m}_1\hat{m}_2}{\hat{m}_2 - 2\hat{m}_1^2}, \end{aligned}$$

where, as before, $\hat{m}_k$ is the sample k-th raw moment (3.2). Note, that the estimators are well defined only when $\hat{m}_2 - 2\hat{m}_1^2 > 0$. Unfortunately, there are no closed form expressions for the maximum likelihood estimators and they can only be evaluated numerically.

Pareto random variable can be very easily simulated by using the standard inverse transform method [179].

3.6 Stable distribution

Stable distributions, also called α-stable are ubiquitous in nature due to generalized central limit theorem. It says that the stable distributions, like the Gaussian one, attract distributions of sums of independent identically distributed random variables [105, 188, 3]. Due to this reason, stable distributions naturally appear when evolution of a system or result of an experiment are determined by a sum of many random factors.

Definition 3.6 *A random variable X is said to have a stable distribution if it has a* domain of attraction*, i.e. if there is a sequence of i.i.d. random variables $Y_1, Y_2, \ldots$ and sequences of positive numbers (d_n) and real numbers (a_n), such that*

$$\frac{Y_1 + Y_2 + \cdots Y_n}{d_n} + a_n \xrightarrow{D} X,$$

where $\xrightarrow{D}$ denotes convergence in distribution.

It appears that $d_n{=}n^{1/\alpha}L(n)$, for some $\alpha \in (0, 2]$ and L being a slowly varying function in infinity [188]. A stable random variable is called symmetric stable if its distribution is symmetric, that is, if X and $-X$ have the same distribution.

Hence, the parameter α is one of the main characteristics of X and is known as the index of stability. A stable random variable X with index α is called α-stable. When $\alpha = 2$, the distribution reduces to Gaussian.

The second equivalent definition of the stable law provides the form of the characteristic function [105, 188].

Definition 3.7 *A random variable X is stable if for some $\alpha \in (0; 2]$, $\sigma \in (0; \infty)$, $\beta \in [-1; 1]$, $\mu \in \mathbb{R}$ its characteristic function has the following form:*

$$\phi_X(t) = \mathrm{E}\exp(itX) = \begin{cases} \exp(-\sigma^\alpha |t|^\alpha (1 - i\beta \operatorname{sgn}(t) \tan \frac{\pi\alpha}{2}) + it\mu) & \text{if} \quad \alpha \neq 1, \\ \exp(-\sigma|t|(1 + i\beta \frac{2}{\pi} \operatorname{sgn}(t) \ln|t|) + it\mu) & \text{if} \quad \alpha = 1. \end{cases}$$

Therefore, each stable distribution is characterized by four parameters. The parameter α is called the index of stability, σ the scale parameter, β the skewness

parameter and μ is the location parameter. Stable variable X is denoted by $X \sim S_\alpha(\sigma, \beta, \mu)$, whereas symmetric α-stable random variable is denoted by $X \sim S\alpha S$. In the latter case $\beta = \mu = 0$. For the Gaussian distribution β is irrelevant, and it is characterized by standard deviation and mean.

It is known that for general stable distributions, although they are continuous, there is no elementary form of probability density function. Computational formulas for stable densities and distribution functions are given in [163]. The Fox function representation for the stable distributions is exhibited in [189] and stable densities in terms of the incomplete hypergeometric function are expressed in [100].

Remark 3.1 *The only stable distributions with elementary probability density functions are:*

- *the Gaussian distribution, namely* $X \sim S_2(\sigma, 0, \mu) = N(\mu, 2\sigma^2)$,
- *the Cauchy distribution, namely* $X \sim S_1(\sigma, 0, \mu)$,
- *the Lèvy distribution, namely* $X \sim S_{\frac{1}{2}}(\sigma, 1, \mu)$.

Some measurable quantities obey stable statistics exactly, such as, e.g., gravitational field of masses distributed randomly in space, hitting times for one-dimensional Brownian motion, and points of arrival in two-dimensional one [75]; interesting examples can be also found in [105, 209]. Stable statistics may also appear asymptotically due to generalized central limit theorem, like, e.g., in non-Brownian random walks with jumps and/or waiting times obeying heavy-tailed distributions, see the reviews [152, 151]; we mention here such illuminating examples as circulation of dollar bills [28], Lévy flights for light in fractal medium called Lévy glass [7], and Lévy-like behaviour of the marine vertebrates in response to patchy distribution of food resources [190]. Stably distributed random noises are observed in such diverse applications as plasma physics (density and electric field fluctuations [157, 109]), stochastic climate dynamics [69], physiology (heartbeats [172]), electrical engineering [161], telecommunication [154], biology [48, 33], chemistry [174], and economics [105, 148, 143, 216, 217, 35]. These and a lot of other observations of stably (or stable-like) distributed quantities require reliable methods of random data analysis which allow to detect stable distributions (or distributions which belong to their domains of attraction) and estimate their parameters.

The tails of non-Gaussian stable distributions decrease like a power function: $|x|^{-\alpha}$. Due to this reason, they appear naturally in the description of random processes with large outliers. The rate of decay mainly depends on the parameter α. The smaller the α, the slower the decay and the heavier the tails. Consequently,

for a stable random variable X with index $\alpha < 2$ one has $\mathrm{E}|X|^\delta = \infty$ for any $\delta \geqslant \alpha$ and $\mathrm{E}|X|^\delta < \infty$ for $0 < \delta < \alpha$.

The comprehensive theory of α-stable distributions is presented in monographs [105, 188, 209].

3.6.1 Estimation

There are at least three procedures for estimating stable law parameters: (i) the maximum likelihood method based on numerical approximation of the stable likelihood function [71, 162, 216]; (ii) the quantile method using tabulated quantiles of stable laws [147]; and (iii) the method using regression on the sample characteristic function [122]. The regression method is considered to be both fast and accurate [216].

3.6.2 Simulation

The complexity of the problem of simulating sequences of α-stable random variables results from the fact that there are no analytic expressions for the inverse F^{-1} of the cumulative distribution function. The first breakthrough was made by Kanter [111], who gave a direct method for simulating $S_\alpha(1, 1, 0)$ random variables, for $\alpha < 1$. It turned out that this method could be easily adapted to the general case. Chambers, Mallows, and Stuck [53] were the first to give the formulas.

It is interesting to note that for $\alpha = 2$ the Chambers–Mallows–Stuck method reduces to the well-known Box–Muller algorithm for generating Gaussian random variables [105]. Although many other approaches have been proposed in the literature, this method is regarded as the fastest and the most accurate [216].

3.6.3 Testing

To test whether a data follow a stable distribution one can apply strict testing procedure described in Section 5.2 for arbitrary distributions or to employ specific properties of the distribution. We concentrate here on the latter case. One way to test for stability is to check the distinctive property of stable random variables, namely summability, see, e.g., [27]. The second array of tests checks the power-law behaviour of the underlying data, see, e.g., [168]. In this context an issue arises of recognizing stable distributions with the stability index close to 2. In this case, if the sample is not long enough, the shape of empirical PDF is close to a Gaussian [49], and both log-log scale analysis and standard estimators of the power-law exponent estimators like Hill give overestimated value of α for the number of observation less than 10^6 [215, 25]. In applications, the number of observations is often less.

The problem of recognizing α-stable distribution with α close to 2 from experimental data when the sample size of available data is not large, was addressed in [49]. Following [49] we introduce a testing procedure combining a simple visual test based on empirical fourth moment, and the Anderson–Darling (AD) and Jarque–Bera (JB) statistical tests (tests are explained in detail in Section 5.2). Namely, we calculate the empirical cumulative fourth moment (ECFM) of the simulated data sets, which for a sample of observations $\{x_1, \ldots, x_n\}$ is defined as follows:

$$C(k) = \frac{1}{k}\sum_{i=1}^{k}(x_i - \bar{x})^4, \ k = 1, 2, ..., n, \tag{3.4}$$

where $\bar{x}$ is the mean of the random sample.

Formula (3.4) can be calculated for any sample obtained from an arbitrary probability distribution. For a fixed k it forms a random variable. For distributions with finite fourth moment (e.g., Gaussian), the ECFM, as a function of k, converges to a constant, whereas for distributions with infinite fourth moment (e.g., stable with $\alpha < 2$) it diverges to infinity. The latter, for a finite sample, can be observed as an irregular chaotic behaviour. Such reasoning stands for the idea of the first step of our algorithm.

This simple test, which is presented in Figure 3.1 (inset), clearly indicates that there is a noticeable difference between non-Gaussian stable and Gaussian distributions. For the former the ECFM does not tend to a constant value with number k of observations increasing, and behaves chaotically, while for the latter one for large numbers of observations it goes to the theoretical fourth moment that in this case is equal to 12. Now, if we take stably distributed variables with smaller stability index, say, $\alpha = 1.9$, and the same total number of observations n, then the difference between the empirical stable PDF and Gaussian PDF becomes more visible; this is shown in Figure 3.2. Of course, in this case we also observe the chaotic behaviour of the ECFM for stably distributed variables, see the inset in Figure 3.2.

In Figure 3.3 we present an asymmetric case with $\beta = 1$ and $\alpha = 1.98$. Similarly to the symmetric case $\beta = 0$, we plot the PDF of the Gaussian distribution with mean equal zero and variance equal to 2, and the slope of the asymptotics of the stable PDF. The difference between the considered PDFs is almost not visible, similar to Figure 3.1. However, in contrast to Figure 3.1, the visual test based on ECFM does not indicate clearly a difference between the analyzed distributions, since for both empirical PDFs the ECFM tends to a constant, see the inset in Figure 3.3. This example demonstrates that obviously there is a need to use a more advanced technique for recognition of the stable distribution with α close to 2.

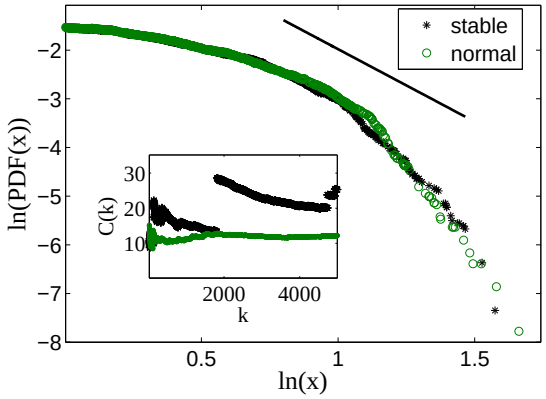

Figure 3.1: Empirical PDF calculated for simulated samples from stable distribution with parameters $\alpha = 1.98$, $\beta = 0$, $\sigma = 1$ and $\mu = 0$ (black asterisks), and from Gaussian distribution with mean equal zero and variance equal 2 (grey circles), are shown in the log-log scale. Number of observations in both samples is 5000. The straight line shows the asymptotic slope of the stable PDF with $\alpha = 1.98$. Inset: Empirical cumulative fourth moment for the samples considered.

Now, we combine the AD and JB tests with the test based on ECFM. The resulting procedure in presented in the following algorithm.

Algorithm 3.1 *Recognizing α-stable distribution with α close to 2*

1. *Calculate the ECFM for the sample.*
2. *If the ECFM tends to a constant, we check for the Gaussian distribution by using the JB test.*
   (a) *If its p-value exceeds the confidence level (usually 5%), then we can assume the underlying distribution of time series is Gaussian. In this case we estimate its parameters by using the standard maximum likelihood estimation method.*
   (b) *If the JB test shows the data can not be modeled by a Gaussian distribution, then we test them for the stable distribution.*
      i. *If the AD test gives the p-value that exceeds the confidence level, then we can assume the time series can be described by the α-stable distribution. In this case we estimate its parameters via the regression approach.*
      ii. *If the stability is rejected other distributions have to be take into considerations.*

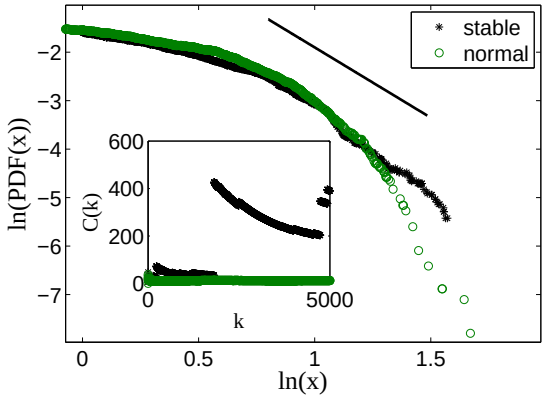

Figure 3.2: Empirical PDF calculated for simulated sample from stable distribution with parameters $\alpha = 1.9$, $\beta = 0$, $\sigma = 1$ and $\mu = 0$ (black asterisks), and from Gaussian distribution with mean equal zero and variance equal 2 (grey circles) in the log-log scale. Number of observations in both samples is 5000. The straight line shows the asymptotic slope of the stable PDF with $\alpha = 1.9$. Inset: Empirical cumulative fourth moment for the samples considered.

3. *If the ECFM exhibits chaotic-like behaviour, then we test for the stable distribution by means of AD test.*
   (a) *If the p-value is greater than the confidence level, then we can assume the data follow the stable law. In this case we estimate its parameters via the regression approach.*
   (b) *If the stability is rejected other distributions have to be take into considerations.*

In order to illustrate the procedure we apply it to the simulated samples presented in Figures 3.1–3.3 and to a generated Gaussian sample. The obtained results are presented in Table 3.1. In the last two columns we demonstrate the values of the JB and AD statistics, respectively, together with the corresponding p-values (in parentheses).

For the sample from the stable distribution with parameters $\alpha = 1.98$, $\beta = 0$, $\sigma = 1$ and $\mu = 0$ we can observe chaotic behaviour of the ECFM, see Figure 3.1.

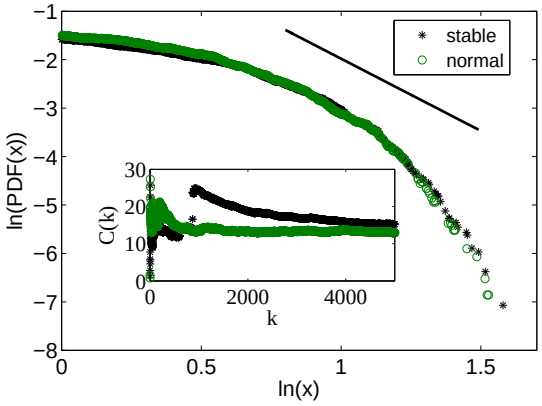

Figure 3.3: Empirical PDF calculated for simulated sample from stable distribution with parameters $\alpha = 1.98$, $\beta = 1$, $\sigma = 1$ and $\mu = 0$ (black asterisks), and from Gaussian with mean equal zero and variance equal 2 (grey circles) in the log-log scale. Number of observations in both samples is 5000. The straight line shows the asymptotic slope of the stable PDF with $\alpha = 1.9$. Inset: Empirical cumulative fourth moment for the samples considered.

Therefore, according to the procedure, we test its stability using the AD test statistic. The obtained p-value indicates the time series constitutes i.i.d. stable random variables.

As we observe in Figure 3.2, the ECFM for the sample from symmetric stable distribution with $\alpha = 1.9$ and $\sigma = 1$ behaves chaotically, therefore we use the AD statistic here. The obtained p-value clearly indicates the underlying stable distribution.

The ECFM for simulated data from the totally skewed stable distribution with $\alpha = 1.98$ for large number of observations tends to a constant (see Figure 3.3), therefore first we test the Gaussian law hypothesis. The JB test indicates the time series can not be modeled by the normal distribution, therefore we check for the Lévy stable by using the AD test that justifies the stable distribution.

For the last simulated sample from the normal distribution we observe that the ECFM tends to a constant. The obtained p-value clearly indicates that the underlying distribution of the analyzed time series is Gaussian.

Table 3.1: Illustration of the introduced procedure for simulated samples. The last two columns contain values of the appropriate statistics, calculated by equations (5.1)–(5.4), and the corresponding p-values (in parentheses).

Sample distribution	**ECFM tends to const?**	**JB stat. (p-value)**	**AD stat. (p-value)**
Stable $\alpha = 1.98$, $\beta = 0$ $\sigma = 1$, $\mu = 0$	NO	–	0.24 (0.64)
Stable $\alpha = 1.90$, $\beta = 0$ $\sigma = 1$, $\mu = 0$	NO	–	0.88 (0.42)
Lévy $\alpha = 1.98$, $\beta = 1$ $\sigma = 1$, $\mu = 0$	YES	61.33 (< 0.01)	0.18 (0.86)
Gaussian $\mu = 0$, $\sigma = \sqrt{2}$	YES	0.55 (0.5)	–

4

Three models for fractional dynamical systems

In this chapter we discuss three models for fractional dynamical systems: fractional Brownian motion (Section 4.1), fractional stable motion (Section 4.2), and FARIMA (Section 4.3). The former two models are self-similar and their increments form long-range dependent processes. The discrete-time FARIMA process is stationary and generalizes both models since aggregated, in the limit, it converges to either fractional Brownian or stable motion. In contrast to them, it allows for different light- and heavy-tailed distributions, and both long (power-like) and short (exponential) dependences. Moreover, as a stationary process, it provides prediction tools. In Section 4.3 we present a new estimation algorithm for the FARIMA parameters in the case of noise belonging to the attraction of the stable law for both positive and negative memory parameters. A new algorithm for fractional differencing is also provided.

4.1 Fractional Brownian motion

Fractional Brownian motion introduced by Kolmogorov in 1940 [121, 141] is a generalization of the classical Brownian motion (BM) [76]. Most of its statistical properties are characterized by the self-similarity (Hurst) exponent $0 < H < 1$.

Definition 4.1 *For any $0 < H < 1$, fractional Brownian motion (FBM) of index H (Hurst exponent) is the mean-zero Gaussian process $B_H(t)$ with the following integral representation [141]:*

$$B_H(t) = \int_{-\infty}^{\infty} \left\{ (t-u)_+^{H-\frac{1}{2}} - (-u)_+^{H-\frac{1}{2}} \right\} dB(u),$$

where $B(t)$ is a standard Brownian motion and $(x)_+ = \max(x, 0)$.

FBM is H-self-similar, namely for every $c > 0$ we have $B_H(ct) \stackrel{D}{=} c^H B_H(t)$ in the sense of all finite dimensional distributions, and has stationary increments. It is the only Gaussian process satisfying these properties. With probability 1, the graph of $B_H(t)$ has both Hausdorff dimension and box dimension of $2 - H$ (for the idea based on the fractal dimension of testing existence of nonlinear I/O relationships on samples of FBM see [176]).

For $H > 1/2$, the increments of the process are positively correlated and exhibit positive long-range dependence, whereas for $H < 1/2$, the increments of the process are negatively correlated and exhibit negative long-range dependence (see Section 5.4). For the second moment of the FBM we have $\mathrm{E}B_H^2(t) = \sigma^2 t^{2H}$, where $\sigma > 0$.

The precise simulation of such process is of great interest. The most commonly used approaches can be split in two categories. The first one, related to theoretically exact methods, has so far been composed only of a matrix factorization technique based on the Cholesky decomposition of the covariance matrix [173]. Unfortunately, this technique has a complexity of $O(N^2)$ and requires high computational resources even for moderate trajectory lengths. The other category is composed of nonexact techniques [139]. All of the above methods have their particular drawbacks and advantages. The choice between them boils down to a tradeoff between speed and accuracy. However, Davies and Harte [65] proposed both fast and exact synthesis method for stationary Gaussian processes, called the circulant embedding method. It can be easily applied for a FBM [50]. Since based on the fast Fourier transform (FFT) algorithm, its complexity is only $O(N \log N)$ (for the Fourier transform and its application to science and engineering, see, e.g., [224]).

4.2 Fractional stable motion

FBM can be generalized to a fractional stable motion [105, 188, 117, 213, 207].

Definition 4.2 *Fractional stable motion (FSM) is the α-stable process $L_H^\alpha(t)$ with the following integral representation:*

$$L_H^\alpha(t) = \int_{-\infty}^{\infty} \left\{ (t-u)_+^d - (-u)_+^d \right\} dL_\alpha(u), \tag{4.1}$$

where $L_\alpha(t)$ is an α-stable motion (SM), $0 < \alpha \leqslant 2$, $0 < H < 1$, and $d = H - 1/\alpha$.

The process for $\alpha = 2$ becomes a FBM, is H-self-similar and has stationary increments. Analogously to the FBM case, we say the increments of the process exhibit positive long-range dependence if $d > 0$ ($H > 1/\alpha$), and negative dependence when $d < 0$ ($H < 1/\alpha$) [188, 195]. This is due to the behaviour of the integrand (kernel function) in (4.1).

When $d < 0$ the kernel function has singularities at $u = 0$ and $u = t$. These singularities magnify the stable noise process $dL_\alpha(t)$, cause large spikes in the paths of the FSM. Their dependence structure resembles that of a negatively correlated process and thus we shall refer to this case as to the negative dependence scenario. In the case when $d > 0$ the kernel is bounded and positive, for all $t > 0$. Thus the jumps in the paths of $L_H^\alpha(t)$ due to the fluctuations of the noise process are not as magnified as in the case $d < 0$. In this case especially for large values of H, the kernel function decays slowly as $s \to -\infty$. This implies that the past fluctuations in the process $dL_\alpha(t)$ influence significantly the present values of the process $L_H^\alpha(t)$. This case is referred to as the positive long-range dependence scenario. Therefore, as in the Gaussian case, the parameter d controls sign of dependence.

Simulation of such process is even more difficult than in the FBM case. This is due the fact that stable processes cannot be characterized by two-dimensional distributions, which is the case for the Gaussian case. Namely, any Gaussian random vector is fully defined by the mean and covariance matrix, whereas stable vectors are defined in terms of so-called spectral measure [105, 188], which is a much more complicated object. Hence, in the literature there are no exact methods of simulation of FSM. The most commonly used approaches apply integral representation of FSM. Stoev and Taqqu [195] proposed a fast method method based on the FFT algorithm with complexity only $O(N \log N)$. It can be also applied to FARIMA processes and more general moving average stable processes, see also [225].

4.3 FARIMA process

Fractional autoregressive integrated moving average (FARIMA) time series were introduced by Granger and Joyeux [84] and Hosking [101]. FARIMA times series serve as a classic example of long-range dependent processes and have been widely studied in the literature, e.g. in telecommunication [104, 61], industry [77], economics [62, 82, 35], astronomy [192], hydrology [125], image processing [103], geoscience [15], medicine [184] and nanoscale biophysics [33, 42].

Definition 4.3 *The FARIMA process* $\{X_t\}$, *denoted by FARIMA*(p, d, q), *is defined by equation*

$$\Phi_p(B)X_t = \Theta_q(B)(1-B)^{-d}Z_t, \tag{4.2}$$

where innovations Z_t *are i.i.d. random variables with either finite or infinite variance.* $\Phi_p(z) = 1 - \phi_1 z - \phi_2 z^2 - \ldots - \phi_p z^p$ *is the autoregressive polynomial and* $\Theta_q(z) = 1 + \theta_1 z + \theta_2 z^2 + \ldots + \theta_q z^q$ *is the moving average polynomial. The operator* $(1-B)^{-d}$ *is the integrating operator and has the infinite binomial expansion*

$$(1-B)^{-d}Z_t = \sum_{j=0}^{\infty} b_j(d)Z_{t-j}, \tag{4.3}$$

where the $b_j(d)'s$ *are the coefficients in the expansion of the function* $f(z) = (1-z)^{-d}$, $|z| < 1$, *i.e.*

$$b_j(d) = \frac{\Gamma(j+d)}{\Gamma(d)\Gamma(j+1)}, \ j = 0, 1, \ldots, \tag{4.4}$$

where Γ *is the gamma function.*

The sequence $\{Z_t\}$ is often called the noise process (sequence) [30, 171]. We assume that innovations Z_t are i.i.d. and belong to the domain of attraction of an α-stable law with $0 < \alpha \leqslant 2$. If $0 < \alpha < 2$, then

$$P(|Z_t| > x) = x^{-\alpha}L(x), \ as \ x \to \infty, \tag{4.5}$$

where L is a slowly varying function at infinity, and

$$\frac{P(Z_t > x)}{P(|Z_t| > x)} \to a, \ \frac{P(Z_t < -x)}{P(|Z_t| > x)} \to b, \ as \ x \to \infty, \tag{4.6}$$

where a and b are nonnegative numbers such that $a + b = 1$. This corresponds to the infinite variance case [188, 192, 196]. We also consider the situation with finite variance, where the noise is in the domain of attraction of a Gaussian law.

The backward operator B satisfies $BX_t = X_{t-1}$ and $B^j X_t = X_{t-j}$, $j \in \mathbb{N}$. When $d = 0$, the integrating operator $(1-B)^{-d}$ reduces to the identity operator and definition (4.2) has a simplified form

$$\Phi_p(B)X_t = \Theta_q(B)Z_t.$$

Therefore, FARIMA$(p, 0, q)$ is equivalent to ARMA(p, q) process [149]. In particular, when $\Phi_p = \Theta_q \equiv 1$, the sequence $\{X_t\}$ is purely the noise process $\{Z_t\}$ and there is no dependence between observations.

For negative integer values of the memory parameter d, i.e, $d = -1, -2, \ldots$, the operator $(1-B)^{-d}$ has a finite extension and a positive exponent $-d$. Therefore, it determines full differencing operation of order $-d$ and is often called the full differencing operator. For example, when $d = -1$ the operator $(1-B)^{-d} = (1-B)$ means taking the increments of the underlying sequence. Moreover, when polynomials Φ_p and Θ_q do not have common roots and Φ_p has no roots in the closed unit disk $\{z : |z| \leqslant 1\}$, formula (4.2) can be rewritten in an equivalent form:

$$\underbrace{X_t}_{\mathrm{FARIMA}(p,\,d,\,q)} = \overbrace{(1-B)^{-d}}^{\text{full differencing of order -d}} \underbrace{\frac{\Theta_q(B)}{\Phi_p(B)} Z_t}_{\mathrm{ARMA}(p,\,q)}.$$

Therefore FARIMA(p, d, q) time series is just a process of increments of order $-d$ of ARMA(p, q) series. In particular, when $\Phi_p = \Theta_q \equiv 1$ and $d = -1$, the sequence $\{X_t\}$ consists purely of the increments of the noise process $\{Z_t\}$.

The series given by (4.3) is convergent almost surely (a.s.) and FARIMA definition (4.2) is correct if and only if

$$\alpha(d-1) < -1 \Leftrightarrow d < 1 - \frac{1}{\alpha}. \tag{4.7}$$

In particular, in the Gaussian case ($\alpha = 2$) we have $d < 1/2$ and that is why analysis of FARIMA series with a positive integer valued memory parameter d is incorrect. Under assumption (4.7) and above mentioned conditions for polynomials Φ_p and Θ_q, the FARIMA(p, d, q) time series defined by (4.2) has a causal moving average form

$$\underbrace{X_t}_{\mathrm{FARIMA}(p,\,d,\,q)} = \overbrace{(1-B)^{-d}}^{d\text{–fractional integrating}} \underbrace{\frac{\Theta_q(B)}{\Phi_p(B)} Z_t}_{\mathrm{ARMA}(p,\,q)} = \sum_{j=0}^{\infty} c_j(d) Z_{t-j}, \tag{4.8}$$

for details see [119]. Therefore FARIMA(p, d, q) time series can be obtained by d-fractional integrating of ARMA(p, q) series. In particular, the FARIMA$(0, d, 0)$ series is just a d-fractionally integrated noise process. The d-fractional integrating operation through the $(1-B)^{-d}$ operator builds the dependence between observations in a FARIMA sequence, even as they are distant in time.

In the case when the memory parameter d is small and $d \approx 0$, all the coefficients $b_j(d)'$s are very close to zero, only $b_0(d) = 1$. In view of the representation (4.3), FARIMA$(0, d, 0)$ observation X_t depends mainly on the present noise observation Z_t. This explains why FARIMA$(0, d, 0)$ series with $d \approx 0$ behave very similarly to a pure noise process.

When the memory parameter d is close to $1/2$, all the coefficients $b_j(d)'$s are positive and converge to zero at a power rate. In view of the series representation (4.3), FARIMA$(0, d, 0)$ observation X_t depends not only on the present noise observation Z_t, but also depends on the whole history of the noise process. Hence, for FARIMA$(0, d, 0)$ series with $d \approx 1/2$ observations that are distant in time still depend significantly on each other.

For Gaussian FARIMA series the rate of decay of the autocorrelation function is t^{2d-1} (see [101]), thus for $d > 0$ the process is long-range dependent in the classical sense, see Section 5.4. For stable FARIMA series the rate of decay of the codifference function is $t^{\alpha(d-1)+1}$ and therefore the long memory property (see Section 5.4) holds in the codifference sense for $d > 1 - 2/\alpha$, for details see [119].

We focus now on the case when the memory parameter d is negative. Then the operator $(1 - B)^{-d}$ has the positive exponent $-d$. It defines the fractional differencing operation. All the coefficients $b_j(d)'$s are negative (when $-1 < d < 0$), except for $b_0(d) = 1$, and converge to zero at a power rate, but slower than in the d-positive case. Therefore, in view of (4.3), FARIMA$(0, d, 0)$ observation X_t is formed as present noise observation Z_t subtracted by the whole history of the noise process. Hence, FARIMA series with negative memory is a d-fractionally differenced ARMA model.

Simulation of FARIMA processes is by no means straightforward. For such processes, except for the special case: ARMA(p, q) model, there are no exact methods. The existing algorithms mostly concentrate on efficient ways to calculate the infinite sum given in (4.8). Stoev and Taqqu [195] proposed a fast method method based on the FFT algorithm with complexity only $O(N \log N)$. It can be also applied to FSM and more general moving average stable processes.

In Figure 4.1 simulated trajectories of symmetric 1.8-stable FARIMA$(0, d, 0)$ (with zero-mean innovations) with different memory parameters ($d \in \{-0.43, 0.03, 0.43\}$) are presented. The crucial difference in the behaviour of trajectories for different signs of memory is the frequency of level-zero crossing. For $d = 0.03$, the FARIMA trajectory looks like a pure noise. For $d = -0.43$, the FARIMA observations cross the level zero and change the sign more often than in the previous case. Finally, for $d = 0.43$ there are long periods of observations of the same sign and the level zero is crossed very rarely. In order to better illustrate this phenomenon, we also show the corresponding cumulative sums of FARIMA trajectories. Now, the difference is pronounced in different vertical scales.

If $d < 1-1/\alpha$ and m is integer number such that $d_0 := d+m \in [-1/\alpha, 1-1/\alpha)$, then

$$X_t(d) = (1 - B)^m X_t(d_0) \ a.s.,$$

where $\{X_t(d)\}$ and $\{X_t(d_0)\}$ are FARIMA processes with memory parameters d and d_0, respectively. Therefore, any FARIMA(p, d, q) time series can be obtained

by a number of full differencings applied to FARIMA(p, d_0, q) with $d_0 \in [-1/\alpha, 1-1/\alpha)$. According to this fact, without any loss of generality, we may consider FARIMA models with memory parameter only in the interval $[-1/\alpha, 1-1/\alpha)$.

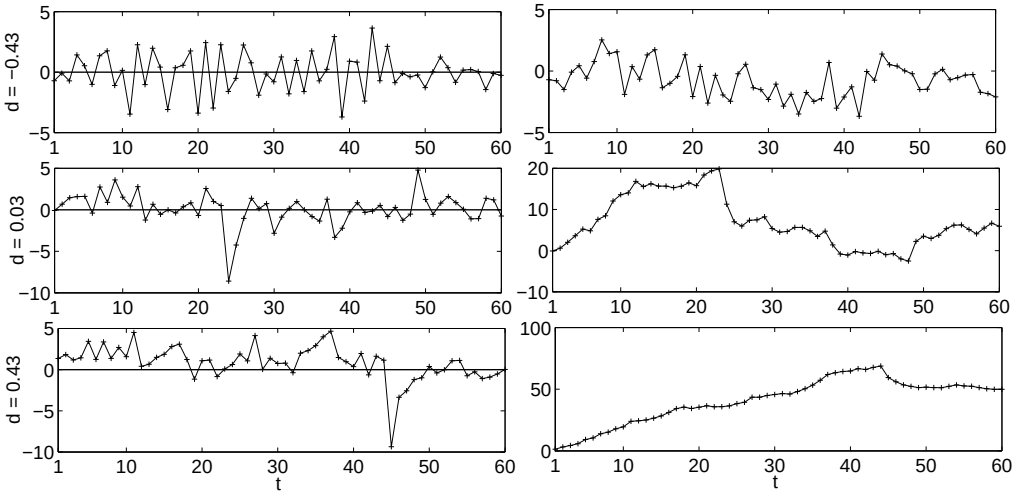

Figure 4.1: Simulated trajectories of symmetric 1.8-stable FARIMA$(0, d, 0)$ series with the memory parameter $d \in \{-0.43, 0.03, 0.43\}$ (left panel), and their cumulative sums (right panel). Notice different vertical scales for figures in the right panel.

Finally, the linear predictor for the FARIMA process based on the finite past $X_n, \ldots, X_0$ takes the form:

$$\hat{X}_{n+h} = \sum_{j=0}^{n} a_j X_{n-j}, \tag{4.9}$$

where the sequence $\{a_0, a_1, \ldots, a_n\}$ is given by $a_j = -\sum_{t=0}^{k-1} c_t h_{j+k-t}$, the c_j's are defined by (4.8) and h_j's are given by

$$\frac{\Phi_p(z)(1-z)^d}{\Theta_q(z)} = \sum_{j=0}^{\infty} h_j z^j, \quad |z| < 1,$$

see [118] for the discussion of the prediction problem in the infinite variance FARIMA case.

4.3.1 Estimation

First, let us recall a method of estimation of FARIMA parameters, which was developed in [120] for FARIMA time series with the positive memory parameter $d \in (0, 1-1/\alpha)$. It is well-defined for FARIMA series with noise satisfying

conditions (4.5) and (4.6). Since $d \in (0, 1 - 1/\alpha)$, the method is only defined for $\alpha \in (1, 2]$.

The method is a variant of Whittle's estimator [90]. In [90] Whittle's method was applied to finite variance ARMA time series. In the Gaussian case the estimator is closely connected to least squares (for the small sample bias least squares estimator see [159]) and maximum likelihood estimators and is therefore a standard estimator for ARMA processes with finite variance [29, 13, 30]. These results were extended in [78] to Gaussian FARIMA processes with positive memory. For ARMA time series with noise in the domain of attraction of non-Gaussian stable law, the estimation techniques were developed in [156]. The method described in [120] is an extension of these results to FARIMA processes with stable noise and positive memory property, i.e. for $d \in (0, 1 - 1/\alpha)$ and $\alpha > 1$.

Following [120], we define the $(p + q + 1)$-dimensional vector

$$\boldsymbol{\beta}_0 := (\phi_1, \phi_2, \ldots, \phi_p, \theta_1, \theta_2, \ldots, \theta_q, d),$$

where $\phi_1, \phi_2, \ldots, \phi_p$ and $\theta_1, \theta_2, \ldots, \theta_q$ are the coefficients of the polynomials Φ_p and Θ_q respectively. The vector $\boldsymbol{\beta}_0$ is from the parameter space $E := \{\boldsymbol{\beta} : \phi_p,\ \theta_q \neq 0, \Phi_p(z)\Theta_q(z) \neq 0 \text{ for } |z| \leqslant 1, \Phi_p,\ \Theta_q \text{ have no common roots}, d \in (0, 1 - 1/\alpha)\}$.

We denote the normalized periodogram by

$$\tilde{I}_n(\lambda) := \frac{\left|\sum_{t=1}^n X_t e^{-i\lambda n}\right|^2}{\sum_{t=1}^n X^2(t)},\ -\pi \leqslant \lambda \leqslant \pi, \tag{4.10}$$

and let

$$\gamma(0) := \sum_{j=0}^{\infty} c_j^2(d), \tag{4.11}$$

where the coefficients $c_j(d)'$s are defined in (4.8).

We also introduce the power transfer function (here we make the dependence on $\boldsymbol{\beta}$ explicit)

$$g(\lambda, \boldsymbol{\beta}) := \left|\frac{\Theta_q(e^{-i\lambda}, \boldsymbol{\beta})}{\Phi_p(e^{-i\lambda}, \boldsymbol{\beta})(1 - e^{-i\lambda})^{d(\boldsymbol{\beta})}}\right|^2.$$

Finally, the estimator $\boldsymbol{\beta}_n$ of the true parameter vector $\boldsymbol{\beta}_0$ is defined as

$$\boldsymbol{\beta}_n := \arg\min_{\boldsymbol{\beta} \in E} \int_{-\pi}^{\pi} \frac{\tilde{I}_n(\lambda)}{g(\lambda, \boldsymbol{\beta})} d\lambda. \tag{4.12}$$

The main result presented in [120] is the following consistency condition. If $\boldsymbol{\beta}_0$ is the true parameter vector and $\boldsymbol{\beta}_n$ is the estimator defined in (4.12), then

$$\boldsymbol{\beta}_n \xrightarrow{P} \boldsymbol{\beta}_0$$

and

$$\int_{-\pi}^{\pi} \frac{\tilde{I}_n(\lambda)}{g(\lambda, \boldsymbol{\beta}_n)} d\lambda \xrightarrow{P} \frac{2\pi}{\gamma(0)},$$

where the coefficient $\gamma(0)$ is defined in (4.11) and $\xrightarrow{P}$ denotes convergence in probability.

Now, we study an extension of the FARIMA estimation method to the d-negative case, for $d \in (-1/2, 0)$ [43]. In view of assumption (4.7) it yields that $\alpha \in (2/3, 2]$. We construct an estimator in this new case which is an extension of the one proposed in the positive memory case. Moreover, we justify the consistency result for the estimator using Monte Carlo simulations. A rigorous proof of the consistency is presented in [43].

First, we start with noticing that the denominator of the formula (4.10) for the periodogram function $\tilde{I}_n(\cdot)$ does not depend on λ. Therefore, we can simplify the form (4.12) of the estimator $\boldsymbol{\beta}_n$ by replacing $\tilde{I}_n(\cdot)$ with

$$I_n(\lambda) := \left| \sum_{t=1}^{n} X_t e^{-i\lambda t} \right|^2, \ -\pi \leqslant \lambda \leqslant \pi.$$

The next simplifying step may be done by noticing the symmetry of $I_n(\cdot)$ (since all considered processes are real-valued) and power transfer function $g(\cdot, \boldsymbol{\beta})$ with respect to λ. Hence the simplified form for the estimator $\boldsymbol{\beta}_n$ can be written as

$$\boldsymbol{\beta}_n := \arg\min_{\boldsymbol{\beta} \in E} \int_0^{\pi} \frac{I_n(\lambda)}{g(\lambda, \boldsymbol{\beta})} d\lambda. \tag{4.13}$$

The subject of our analysis is the estimator $\tilde{\boldsymbol{\beta}}_n$ of the above form, but with the revised parameter space $\tilde{E}$, the same space as E but with $d \in (-1/2, 0)$. The crucial difference between the classical d-positive case and considered here situation with $d \in (-1/2, 0)$ is the fact that function $1/g(\cdot, \cdot)$ is discontinuous on space $[0, \pi] \times \tilde{E}$. More precisely, the function $1/g(\cdot, \boldsymbol{\beta})$ diverges at $\lambda = 0$ for any $\boldsymbol{\beta} \in \tilde{E}$. Therefore, the domain of the function $1/g(\cdot, \cdot)$ is restricted to $(0, \pi] \times \tilde{E}$. This suggests us to define an estimator in d-negative case as

$$\tilde{\boldsymbol{\beta}}_n := \arg\min_{\boldsymbol{\beta} \in E} \int_{\frac{1}{n}}^{\pi} \frac{I_n(\lambda)}{g(\lambda, \boldsymbol{\beta})} d\lambda. \tag{4.14}$$

Notice now that the function $1/g(\cdot, \cdot)$ is continuous on $[1/n, \pi] \times \tilde{E}$ for any $n \in \mathbb{N}$. Therefore, the estimator $\tilde{\boldsymbol{\beta}}_n$ is correctly defined. The main result is the following consistency theorem for the estimator $\tilde{\boldsymbol{\beta}}_n$.

Theorem 4.1 *If $\boldsymbol{\beta}_0 \in \tilde{E}$ is the true parameter, then*

$$\tilde{\boldsymbol{\beta}}_n := \arg\min_{\boldsymbol{\beta}\in\tilde{E}} \int_{\frac{1}{n}}^{\pi} \frac{I_n(\lambda)}{g(\lambda,\boldsymbol{\beta})} d\lambda \quad \xrightarrow{P} \quad \boldsymbol{\beta}_0$$

and

$$\int_{\frac{1}{n}}^{\pi} \frac{I_n(\lambda)}{g(\lambda,\boldsymbol{\beta})} d\lambda \quad \xrightarrow{P} \quad \frac{\pi}{\gamma(0)}.$$

For the proof we refer to [43]. We justify here the consistency of the FARIMA estimator $\tilde{\boldsymbol{\beta}}_n$ by means of the Monte Carlo method [33]. To this end we perform simulations for sample trajectories of FARIMA$(1, d, 1)$ with α-stable innovations. In such a case the parameter space $\tilde{E}$ has three dimensions, i.e. $\boldsymbol{\beta} = (\phi_1, \theta_1, d) \in \tilde{E}$, polynomials $\Phi_p(z) = 1 - \phi_1 z$, $\Theta_q(z) = 1 + \theta_1 z$ and the power transfer function has the form

$$g(\lambda, \boldsymbol{\beta}) = \left| \frac{1 + \theta_1 e^{-i\lambda}}{1 - \phi_1 e^{-i\lambda}} \right|^2 \frac{1}{(2 - 2\cos\lambda)^d}.$$

Therefore, the estimator $\tilde{\boldsymbol{\beta}}_n$ can be rewritten as

$$\arg\min_{\boldsymbol{\beta}\in\tilde{E}} \int_{\frac{1}{n}}^{\pi} I_n(\lambda) \left[\frac{1 - 2\phi_1\cos\lambda + \phi_1^2}{1 + 2\theta_1\cos\lambda + \theta_1^2} \right] (2 - 2\cos\lambda)^d d\lambda.$$

We consider the following sets of parameters:

- $(\alpha, d) \in \{1.8\} \times \{-0.4, -0.2, -0.1\}$ and $\beta = 0.5$.

For the above three sets we choose the following polynomial coefficients:

- $\phi_1 = -0.1$ and $\theta_1 = 0.2$,
- $\phi_1 = 0.6$ and $\theta_1 = -0.3$,
- $\phi_1 = -0.1$ and $\theta_1 = -0.3$,

respectively.

We simulated 100 trajectories of FARIMA$(1, d, 1)$, using the algorithm presented in [195], of length $N = 50000$. Next, we calculated the values of FARIMA estimators (the memory parameter and polynomial coefficients) for increasing subsets of the trajectories, namely for the first $n = 1000$, $n = 10000$, and $n = 50000$ (i.e. the whole trajectory) points. The results are presented in the form of box plots in Figure 4.2. We can see that with the increase of subset's length n, the estimators converge to the true values as their variance decreases.

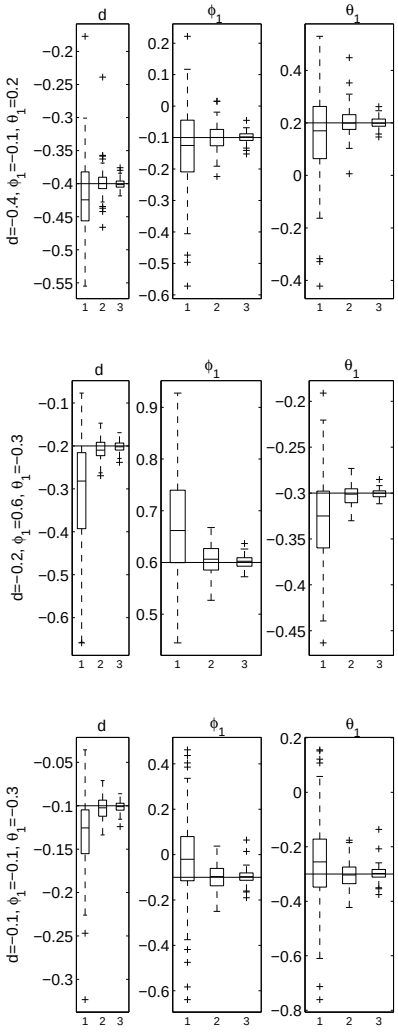

Figure 4.2: The box plots of estimators (d, ϕ_1, θ_1) for FARIMA$(1, d, 1)$ series with skewed α-stable innovations with $\alpha = 1.8$ and $\beta = 0.5$ for three trajectory subsets: (top panel) $n = 1000$, (middle panel) $n = 10000$, and (bottom panel) $n = 50000$ of total length $N = 50000$. The horizontal lines correspond to the true values of parameters (d, ϕ_1, θ_1).

Let us also clarify for which set of values of d and α the described FARIMA estimation method works properly. Assumption (4.7) gives that $d < 1 - 1/\alpha$. Moreover, according to [43], the estimator is correctly defined only when $d > -1/2$. Therefore, we may conclude that we extended the FARIMA estimation technique from the positive memory case with $d \in (0, 1/2)$ and $\alpha \in (1, 2]$ to the

negative memory case with $d \in (-1/2, 0)$ and $\alpha \in (2/3, 2]$. We illustrate the area of applicability of the extended FARIMA estimation method in Figure 4.3. Finally, we should remember that in order to apply the estimation procedure we have to know the order of the model: (p, q). This can done at least in two different ways, see Section 5.1.1 and discussion below.

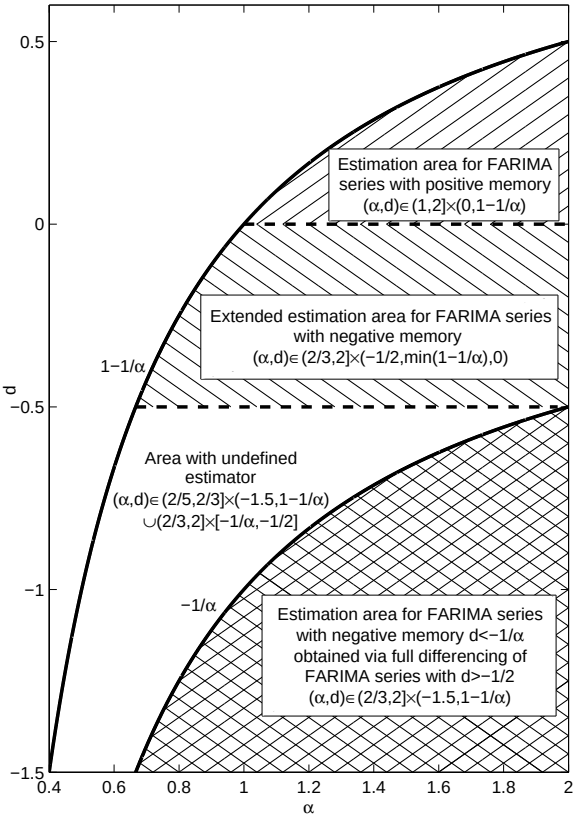

Figure 4.3: The area of applicability of the FARIMA estimation method.

4.3.2 Fractional differencing – from FARIMA to ARMA models

We briefly address here the issue of fractional differencing. Recall that one can d–fractionally differentiate the FARIMA(p, d, q) sequence $\{X_t\}$ to obtain the ARMA(p, q) series $\{(1 - B)^d X_t\}$. This idea can be helpful, e.g., in estimation of FARIMA parameters since estimation methods for ARMA models are well established [156, 30]. Thus, the problem of fitting of the FARIMA model can be reduced to fitting of a simpler model or at least the order of the model (p, q) can

be recovered [34, 44]. Furthermore, this procedure can be applied to calculate residuals of the process, which should be checked for stationarity and independence [33]. However, we have to remember that, first, we need a reliable estimate of d (see Section 5.4).

Our idea of fractional differencing stems from the simulation algorithm of FARIMA time series introduced in [195]. The method involves the discrete Fourier transform (DFT). For a M-periodic sequence $a = \{a_j : j \in \mathbb{Z}\}$, i.e. $a_{j+M} = a_j$, for all $j \in \mathbb{Z}$, the DFT $\mathcal{D}_M(a) = \{\mathcal{D}_M(a)_k : k \in \mathbb{Z}\}$ is defined as

$$\mathcal{D}_M(a)_k := \sum_{j=0}^{M-1} e^{2\pi ijk/M} a(j),\ k \in \mathbb{Z},$$

with the inverse formula

$$a_j = \mathcal{D}_M^{-1}(\mathcal{D}_M(a))_j := \frac{1}{M} \sum_{k=0}^{M-1} e^{-2\pi ijk/M} \mathcal{D}_M(a)_k,\ j \in \mathbb{Z}.$$

Moreover, for any two M-periodic sequences a and b, the following convolution theorem holds

$$\mathcal{D}_M(a)_k \mathcal{D}_M(b)_k = \mathcal{D}_M(a * b)_k,\ k \in \mathbb{Z}, \tag{4.15}$$

where

$$(a * b)_n := \sum_{j=0}^{M-1} a_{n-j} b_j,\ n \in \mathbb{Z}.$$

Furthermore, if M is a power of two, then to compute the DFT of a M-periodic sequence one can apply the efficient FFT (radix–2) algorithm [79].

The fractional differencing algorithm is based on the approximation of the ARMA trajectory of length N

$$Y_n := (1 - B)^d X_n = \sum_{j=0}^{\infty} b_j(-d) Z_{n-j},\ n = 0, 1, \ldots, N-1,$$

by the truncated moving average

$$Y_n \approx Y_n^M := \sum_{j=0}^{M-1} b_j(-d) Z_{n-j},\ n = 0, 1, \ldots, N-1. \tag{4.16}$$

Since the right-hand side of formula (4.16) has a form of a finite discrete convolution, we apply the convolution formula (4.15) and the efficient FFT technique. First, we define $(M + N - 1)$-periodic sequence $\tilde{b}_{j+k(M+N-1)}(-d) := \tilde{b}_j(-d)$, $j = 0, 1, \ldots, M + N - 2$, $k \in \mathbb{Z}$, where

$$\tilde{b}_j(-d) = \begin{cases} b_j(-d), & \text{for } j = 0, 1, \ldots, M-1, \\ 0, & \text{for } j = M, M+1, \ldots, M+N-2, \end{cases}$$

with $b_j(-d)$'s are computed according to formula (4.4).

In order to apply the convolution idea from the simulation algorithm, we assume that the analyzed FARIMA trajectory $\{X_n : n = 0, 1, \ldots, M+N-2\}$ is of length $M+N-2$ and define $(M+N-1)$-periodic sequence $\tilde{X}_{j+k(M+N-1)} := X_j$, $j = 0, 1, \ldots, M+N-2$, $k \in \mathbb{Z}$. Therefore, Y_n can be approximated by the convolution of the two periodic sequences

$$Y_n \approx \sum_{j=0}^{M+N-2} \tilde{b}_j(-d)\tilde{X}_{n-j},\ n = 0, 1, \ldots, N-1$$

and computed by taking the inverse DFT of the product of the DFT's of the two $(M+N-1)$-periodic sequences $\tilde{b} = \{\tilde{b}_j(-d) : j \in \mathbb{Z}\}$ and $\tilde{X} = \{\tilde{X}_j : j \in \mathbb{Z}\}$.

The algorithm can be written as follows.

Algorithm 4.1 *Fractional differencing algorithm*

1. *Choose a truncation parameter $M \in \mathbb{N}$ and expected sample size $N \in \mathbb{N}$, such that $M+N-1$ is the length of the dataset and is a power of two.*
2. *According to relation (4.4), compute $(M+N-1)$-periodic sequence $\tilde{b} = \{\tilde{b}_j(-d) : j = 0, 1, \ldots, M+N-2\}$ and using the radix–2 FFT algorithm, compute the DFT $\mathcal{D}_{M+N-1}(\tilde{b})$ of the sequence $\tilde{b}$.*
3. *Using the radix–2 FFT algorithm, compute the DFT $\mathcal{D}_{M+N-1}(\tilde{X})$ of the $(M+N-1)$-periodic sequence $\tilde{X}$.*
4. *By using the radix–2 FFT algorithm, compute the inverse DFT of the product $\mathcal{D}_{M+N-1}(\tilde{b})\mathcal{D}_{M+N-1}(\tilde{X})$ and approximate Y_n by*

$$Y_n \approx \mathcal{D}^{-1}_{M+N-1}(\mathcal{D}_{M+N-1}(\tilde{b})\mathcal{D}_{M+N-1}(\tilde{X}))_n,$$

for $n = 0, 1, \ldots, N-1$, where the sequence product $\mathcal{D}_{M+N-1}(\tilde{b})\mathcal{D}_{M+N-1}(\tilde{X})$ consists of numbers

$$\mathcal{D}_{M+N-1}(\tilde{b})_j\mathcal{D}_{M+N-1}(\tilde{X})_j,$$

$j = 0, 1, \ldots, M+N-2$.

In the literature different methods of fractional differencing have been proposed, see, e.g. [130, 146, 226]. It may seem that our algorithm has a major drawback from a practical point of view. Note that in order to obtain a series of length N we have to fractionally differentiate a sequence of length $N+M-1$. Therefore, we loose a large part of the data (especially when $M >> N$). However, we believe that it is necessary to truly differentiate the data in a fractional way [44].

5

Selected identification and validation tools

In this chapter we recall closely related notions of self-similarity and long memory, and collect a list of identification and validation tools for fractional dynamical systems. In Table 5.1 the tools are presented along with the related characteristics of the data, which are stationarity, ergodicity, distribution type, self-similarity and long memory. Quantile lines, dynamical functional, goodness of fit statistics and modified RS method stand for examples of both identification and validation tools, whereas FIRT method, sample p-variation, RS method and sample MSD are examples of identification tools.

Table 5.1: List of simple identification and validation tools.

Tools	Characteristic of the data
Autocovariance function and quantile lines	Stationarity
Dynamical functional	Ergodicity
Goodness of fit statistics	Distribution type
FIRT method	Self-similarity
Sample p-variation	Self-similarity
RS method	Memory
Modified RS method	Memory
Sample MSD	Memory

5.1 Stationarity and ergodicity

5.1.1 Autocovariance function and quantile lines

Most tools depicted in Table 5.1 assume that the data are stationary or increment stationary [41]. Stationary and nonstationary processes are very different in their properties, and they require different inference procedures. First, note that a simple and useful method to tell if a process is stationary in empirical studies

is to plot the data. Loosely speaking, if a series does seem to have a trend, or a varying volatility, then very likely, it is not stationary. To make the process stationary sometimes it is enough to calculate its increments.

For the basic diagnostic of dependence and stationarity, autocorrelation function (ACF) and partial autocorrelation function (PACF) are examined. For any ergodic process, both functions, ACF and PACF, should approach to zero as the lag tends to infinity. They also provide information about the order of the underlying ARMA model.

The sample autocorrelation function at lag k is defined as

$$r_k = \frac{\sum_{t=1}^{N-k}(X_t - \overline{X})(X_{t+k} - \overline{X})}{\sum_{t=1}^{N}(X_t - \overline{X})}, \quad \text{for } k = 0, 1, \ldots, K$$

and stands for a measure of the linear dependence between observations with time lag k [30]. For i.i.d. data $r_k = 0$ for all $k \neq 0$. For moving average processes of order q (MA(q)), the autocorrelation function is zero for lags beyond q, hence the function is also useful for estimating the q parameter.

For assessing the order p of an autoregressive processes (AR(p)) we use the partial autocorrelation function, which, at lag k, is defined as a correlation between the predictor errors of values X_n and X_{n+k} represented in terms of $X_{n+1}, \ldots, X_{n+k-1}$. The partial autocorrelation function for AR(p) processes behaves much like the autocorrelation function for MA(q) processes, see [63] and references therein. For i.i.d. series the function should be equal to zero except for the lag $k = 0$.

If a number of realizations of a phenomena is recorded, the stationarity property of the analyzed datasets can be checked by means of so-called quantile lines. Let us assume that we observe M samples of length N and denote their values by $\{Z_n^k\}$, $n = 1, 2, \ldots N$, $k = 1, ..., M$, and $0 < p_j < 1$, $j = 1, ..., J$ are given probabilities. It is possible to derive estimators of the corresponding quantiles $q_j(n) = F_n^{-1}(p_j)$, where $F_n = F_n(x)$ denotes CDF of the random variable Z_n represented by the statistical sample Z_n^k, $k = 1, ..., M$. In this way we obtain the approximation of the so-called quantile lines, i.e., the curves

$$q_j = q_j(n)$$

defined by the condition

$$P\{Z_n \leqslant q_j(n)\} = p_j.$$

In layman terms, the quantile lines represent the value q_j for which $p_j * 100\%$ of the data are below at a certain time point n. For a stationary process the quantile lines $q_j(n) = \text{const}$, whereas for a self-similar process they behave like n^H [219].

5.1.2 Dynamical functional

We describe now another two fundamental properties of the data: ergodicity and mixing. Ergodicity of the stationary process $Z(n)$ means that its phase space cannot be divided into two nontrivial sets such that a point starting in one set will never get to the second set. Let us emphasize that for every stationary and ergodic process the Boltzmann ergodic hypothesis, enabling better analysis of the data characteristics is satisfied, i.e. the temporal and ensemble averages coincide [23, 20].

Another fundamental property is mixing, i.e., the asymptotic independence of the random variables $Z(n)$ and $Z(0)$ as n goes to infinity. It is well-known that mixing is a stronger property than ergodicity [127]. Thus to show ergodicity it is enough to prove mixing, which is easier in many cases.

To this end, we use the dynamical functional (DF) test recently developed in [136]. It is based on a concept of the dynamical functional [105]

$$D(n) = \mathrm{E}\exp\{i(Z(n) - Z(0))\}.$$

Note that $D(n)$ is actually a Fourier transform of the process increment evaluated for the Fourier-space variable $k = 1$. Denote by $E(n) = D(n) - \mathrm{E}^2\exp\{iZ(0)\}$. It turns out that stationary process $Z(n)$ is mixing if and only if

$$\lim_{n\to\infty} E(n) = 0.$$

Similarly, $Z(n)$ is ergodic if and only if

$$\lim_{n\to\infty} 1/n \sum_{k=0}^{n-1} E(k) = 0.$$

The DF test holds for all infinitely divisible stationary processes [136, 137]. In [38] it was applied to telomere data. The test confirmed that the increments of telomere motion have the ergodicity (and mixing) property.

5.2 Goodness of fit tests

A natural estimate for the distribution is the observed (empirical) distribution. For a sample of observations $\{x_1, \ldots, x_n\}$ the empirical distribution function (EDF) is defined as:

$$F_n(x) = \frac{1}{n}\#\{i : x_i \leqslant x\},$$

i.e. it is a piecewise constant function with jumps of size $1/n$ at points x_i. Very often, especially if the sample is large, the EDF is approximated by a continuous, piecewise linear function with the "jump points" connected by linear functions.

A statistics measuring the difference between the empirical $F_n(x)$ and the fitted $F(x)$ distribution function, called an EDF statistic, is based on the vertical difference between the distributions. This distance is usually measured either by a supremum or a quadratic norm [64, 40, 37].

The most popular supremum statistic:

$$D = \sup_x |F_n(x) - F(x)|, \tag{5.1}$$

is known as the Kolmogorov or Kolmogorov-Smirnov (KS) statistic. It can also be written in terms of two supremum statistics:

$$D^+ = \sup_x \{F_n(x) - F(x)\} \quad \text{and} \quad D^- = \sup_x \{F(x) - F_n(x)\},$$

where the former is the largest vertical difference when $F_n(x)$ is larger than $F(x)$ and the latter is the largest vertical difference when it is smaller. The Kolmogorov statistic is then given by $D = \max(D^+, D^-)$. A closely related statistic proposed by Kuiper is simply a sum of the two differences, i.e. $V = D^+ + D^-$.

The second class of measures of discrepancy is given by the Cramér-von Mises family

$$Q = n \int_{-\infty}^{\infty} \{F_n(x) - F(x)\}^2 \psi(x) \mathrm{d}F(x), \tag{5.2}$$

where $\psi(x)$ is a suitable function which gives weights to the squared difference $\{F_n(x) - F(x)\}^2$. When $\psi(x) = 1$ we obtain the W^2 statistic of Cramér-von Mises. When $\psi(x) = [F(x)\{1 - F(x)\}]^{-1}$ formula (5.2) yields the A^2 statistic of Anderson and Darling (AD). From the definitions of the statistics given above, suitable computing formulas must be found. This can be done by utilizing the transformation $Z = F(X)$. When $F(x)$ is the true distribution function of X, the random variable Z is uniformly distributed on the unit interval.

Suppose that a sample $x_1, \ldots, x_n$ gives values $z_i = F(x_i),\ i = 1, \ldots, n$. It can be easily shown that, for values z and x related by $z = F(x)$, the corresponding vertical differences in the EDF diagrams for X and for Z are equal. Consequently, EDF statistics calculated from the empirical distribution function of the z_i's compared with the uniform distribution will take the same values as if they were calculated from the empirical distribution function of the x_i's, compared with $F(x)$. This leads to the following formulas given in terms of the order statistics $z_{(1)} < z_{(2)} < \cdots < z_{(n)}$:

$$
\begin{aligned}
D^+ &= \max_{1\leqslant i\leqslant n}\left\{\frac{i}{n}-z_{(i)}\right\},\\
D^- &= \max_{1\leqslant i\leqslant n}\left\{z_{(i)}-\frac{(i-1)}{n}\right\},\\
D &= \max(D^+,D^-),\\
V &= D^+ + D^-,\\
W^2 &= \sum_{i=1}^{n}\left\{z_{(i)}-\frac{(2i-1)}{2n}\right\}^2+\frac{1}{12n},
\end{aligned}
$$

$$
\begin{aligned}
A^2 &= -n-\frac{1}{n}(2i-1)\sum_{i=1}^{n}\left\{\log z_{(i)}+\log(1-z_{(n+1-i)})\right\} =\\
&= -n-\frac{1}{n}\sum_{i=1}^{n}\left\{(2i-1)\log z_{(i)}+(2n+1-2i)\log(1-z_{(i)})\right\}.
\end{aligned}
$$

For testing the Gaussianity we propose to use the standard Jarque–Bera (JB) test [106], which is different from the KS and AD. It is a goodness of fit test of whether sample data have the skewness and kurtosis matching a normal distribution (for testing bivariate independence and normality see [110]). The JB statistic is defined as:

$$
J=\frac{n}{6}\left(S^2+\frac{(K-3)^2}{4}\right), \tag{5.3}
$$

where S and K are the sample skewness and kurtosis respectively, namely

$$
S=\frac{1/n\sum_{i=1}^{n}(x_i-\bar{x})^3}{\left(\sqrt{1/n\sum_{i=1}^{n}(x_i-\bar{x})^2}\right)^3},
$$

$$
K=\frac{1/n\sum_{i=1}^{n}(x_i-\bar{x})^4}{\left(\sqrt{1/n\sum_{i=1}^{n}(x_i-\bar{x})^2}\right)^2}. \tag{5.4}
$$

Equations (5.3)–(5.4) can be used for any sample from an arbitrary probability distribution. The value of the JB statistic given by equation (5.3) forms a random variable which converges to zero if the underlying distribution has skewness zero and kurtosis 3 (e.g., Gaussian). Any deviation from zero skewness and kurtosis equal 3 increases the JB statistic. For distributions with infinite kurtosis (e.g., α-stable with $\alpha<2$) it diverges to infinity.

To employ any of the tests, first, we need to estimate the parameters of the hypothetical distribution. In the Gaussian case the standard method is the maximum likelihood, whereas in the non-Gaussian stable case we suggest the use of the regression method estimates [122, 216].

The general test of fit is structured as follows [64, 66]. The null hypothesis is that a specific distribution is acceptable, whereas the alternative is that it is not:

$$\begin{aligned} H_0: &\quad F_n(x) = F(x;\zeta), \\ H_1: &\quad F_n(x) \neq F(x;\zeta), \end{aligned}$$

where ζ is a vector of known parameters. Small values of the test statistic T are evidence in favor of the null hypothesis, large ones indicate its falsity. To see how unlikely such a large outcome would be if the null hypothesis was true, we calculate the p-value by:

$$p\text{-value} = P(T \geqslant t),$$

where t is the test value for a given sample. It is typical to reject the null hypothesis when a small p-value is obtained.

However, we are in a situation where we want to test the hypothesis that the sample has a common distribution function $F(x;\zeta)$ with unknown ζ. To employ any of the EDF tests we first need to estimate the parameters. It is important to recognize that when the parameters are estimated from the data, the critical values for the tests of the uniform distribution (or equivalently of a fully specified distribution) must be reduced. In other words, if the value of the test statistics T is d, then the p-value is overestimated by $P_U(T \geqslant d)$. Here P_U indicates that the probability is computed under the assumption of a uniformly distributed sample. Hence, if $P_U(T \geqslant d)$ is small, then the p-value will be even smaller and the hypothesis will be rejected. However, if it is large then we have to obtain a more accurate estimate of the p-value.

Ross [179] advocates the use of Monte Carlo simulations in this context. First the parameter vector is estimated for a given sample of size n, yielding $\hat{\zeta}$, and the EDF test statistics is calculated assuming that the sample is distributed according to $F(x;\hat{\zeta})$, returning a value of d. Next, a sample of size n of $F(x;\hat{\zeta})$-distributed variates is generated. The parameter vector is estimated for this simulated sample, yielding $\hat{\zeta}_1$, and the EDF test statistics is calculated assuming that the sample is distributed according to $F(x;\hat{\zeta}_1)$. The simulation is repeated as many times as required to achieve a certain level of accuracy. The estimate of the p-value is obtained as the proportion of times that the test quantity is at least as large as d.

Finally, for the consistent estimation of the distributions based on incomplete data via the in-sample goodness of fit procedures and forecasting see [58].

5.3 Self-similarity

Self-similar processes, introduced by Lamperti [126], are the ones that are invariant under suitable translations of time and scale. In the last 20 years there has been an explosive growth in the study of self-similar processes, cf., e.g., [205, 72, 70, 186]. They are important in probability theory because of their connection to limit theorems and they are of great interest in modelling heavy-tailed and long-memory phenomena. In fact, Lamperti used the term "semi-stable" in order to underline the role of self-similar processes among stochastic processes which is analogous to the role of stable distributions among all distributions. The term self-similarity was coined by Mandelbrot, who used it also in the context of the scaling of non-random objects [140, 139].

A process $\{X(t)\}_{t\geqslant 0}$ is called self-similar (ss) [126] if for some $H > 0$,

$$X(at) \stackrel{D}{=} a^H X(t) \text{ for every } a > 0, \tag{5.5}$$

where $\stackrel{D}{=}$ denotes equality of all finite-dimensional distributions of the processes on the left and right side. The parameter H is called the self-similarity index or exponent and $\{X(t)\}_{t\geqslant 0}$ is also called H-self-similar. If we interpret t as "time" and $X(t)$ as "space" then (5.5) tells us that every change of time scale $a > 0$ corresponds to a change of space scale a^H. The bigger H, the more dramatic the change of the space co-ordinate [47].

Notice that (5.5), indeed, means a "scale-invariance" of the finite-dimensional distributions of $X(t)$. This property of a self-similar process does not imply scale-invariance for sample paths. Therefore, pictures trying to explain self-similarity by zooming in or out on one sample path, are by definition misleading because in contrast to the deterministic self-similarity, this property does not mean that the picture repeats itself as we enlarge it. It is rather the general impression that remains the same. A convenient mathematical tool to observe self-similarity are so-called quantile lines, see Section 5.1.1.

Self-similarity is convenient for simulations: a sample path of $X(t)$ on $[0, 1]$ multiplied by c^H and re-scaling of the time axis by c immediately provide a sample path on $[0, c]$ for any $c > 0$.

Many interesting self-similar processes have stationary increments. A process $\{X(t)\}_{t\geqslant 0}$ is said to have stationary increments (si) if for any $b > 0$,

$$(X(t + b) - X(b)) \stackrel{D}{=} (X(t) - X(0)) .$$

For the introduction of the fundamental concept of a stationary random process and its autocorrelation structure, see, e.g., [223].

Fix now $0 < H < 1$. Since the function $\{|t_1|^{2H}+|t_2|^{2H}-|t_1-t_2|^{2H},\ t_1, t_2 \in \mathbb{R}\}$ is non-negative definite, there exist a Gaussian process $(X(t))_{t\geqslant 0}$ with mean zero and autocovariance function

$$R(t_1,t_2) = \mathrm{Cov}(X(t_1),X(t_2)) = \frac{1}{2}\left\{|t_1|^{2H}+|t_2|^{2H}-|t_1-t_2|^{2H}\right\}\mathrm{Var}X(1).$$

It is easy to check that this process is H-ss, si. It is called a fractional Brownian motion (see Section 4.1) and is often denoted by $B_H(t)$. Are there any other Gaussian H-ss si processes for $0 < H < 1$? The answer is negative [134].

Various estimators of the self-similarity parameter H has been proposed in the literature, see, e.g., [205, 218, 194, 150, 155, 158]. In this section we concentrate on two methods, namely the well-known FIRT method which is considered in the literature as very accurate and a new one based on the notion of p-variation.

5.3.1 FIRT estimator

Recently introduced an estimation procedure of the self-similarity parameter H is based on the so-called finite impulse response transformation (FIRT) [194]. FIRT is widely used in signal processing (see, e.g., [165]) and also appears in time series analysis (see, e.g., [5]). The estimation method was proposed in [194] for a FSM but as it can be used for general self-similar (even asymptotically) processes with stationary increments belonging to the domain of attraction of the stable law. Hence, it can be also applied to FARIMA processes [194, 195]. FIRT is characterized by a filter $u = (u_0, u_1, \ldots, u_p)$ of real numbers u_k, $k = 1, \ldots, p$. It is defined for a discrete ss si process $\{X_k\}_{k\in\mathbb{Z}}$ by

$$D_{n,k}^{FIRT} = \sum_{i=1}^{p} u_i X_{n(i+k)},\quad n \geqslant 1,\ k \in \mathbb{Z}, \tag{5.6}$$

where $n \geqslant 1$ and the indices n and k can be interpreted as "scale" and "location", respectively.

If we assume that that the filter u has at least one zero moment the following scaling property can be proved [194]:

$$\text{for any } n \geqslant 1,\ \left\{D_{n,k}^{FIRT}\right\}_{k\in\mathbb{Z}} \stackrel{D}{=} n^H \left\{D_{1,k}^{FIRT}\right\}_{k\in\mathbb{Z}}. \tag{5.7}$$

We suppose also that $\left\{D_{n,k}^{FIRT}\right\}_{k\in\mathbb{Z}}$'s are available for some fixed scales n_j, $j = 1, \ldots, m$, and locations $k = 0, \ldots, N_j - 1$ at the scale n_j. Since we want to estimate the self-similarity parameter of the process from its discrete-time observations, suppose that these observations are $X_0, \ldots, X_N$. From (5.6) it is

clear that only a finite number, say N_j, of the FIRT coefficients will be available at the scale n_j and that $N_j \sim N/n_j$, as $N \to \infty$, that is, $\lim_{N\to\infty} n_j N_j = 1$.

Observe that, by (5.7), we have

$$\mathrm{E}\log\left|D_{n_j,0}^{FIRT}\right| = H\log(n_j) + \mathrm{E}\log\left|D_{1,0}^{FIRT}\right|,$$

where $\mathrm{E}\log\left|D_{n_j,0}^{FIRT}\right|$ can be approximated by the statistic

$$d_{FIRT}(n) := \frac{1}{N_j}\sum_{k=0}^{N_j-1}\log\left|D_{n_j,k}^{FIRT}\right|.$$

Finally, the estimator H_{FIRT} is obtained by performing a log-linear regression on the coefficients $d_{FIRT}(n)$ and measuring the slope, which is equal to H. It is unbiased for all $0 < \alpha < 2$. Moreover, for $1 < \alpha < 2$, under certain technical conditions, the estimator is consistent and asymptotically normal. It is closely related to the wavelet transform estimator which was also introduced in [194]. They have been proven to be both reliable and computationally efficient tools [195]. The FIRT estimator is illustrated in Figure 5.1 for the simulated data from a fractional stable motion with $H = 0.4$ and $\alpha = 1.85$. The filter $U = (1, -3, 3, -1)$.

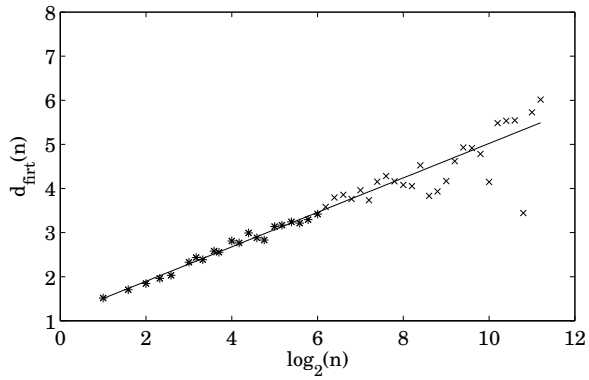

Figure 5.1: FIRT estimator for a time series simulated from FSM with $H = 0.4$ and $\alpha = 1.85$. The estimator H_{FIRT} is obtained by fitting a least-squares line to the values of the coefficients $d_{FIRT}(n)$ on a log-linear scale. The points at the high end (marked by "x") are not used to fit the line. The estimated slope of the line for the original data is $H_{FIRT} = 0.39$.

5.3.2 New estimator based on sample p-variation

Let us now discuss the idea of p-variation, $p > 0$. The concept of p-variation generalizes the well-known notions of total or quadratic variations, which have found applications in various areas of physics [138, 214], mathematics and engineering [54]. Let $X(t)$ be a stochastic process analyzed on the time interval $[0, T]$. Then, the p-variation of $X(t)$ is defined as the limit of sum of increments of $X(t)$ taken to the p-th power over all partitions P of the interval $[0, T]$, when the mesh of the partitions goes to zero. When $p = 1$ it reduces to the total variation, whereas $p = 2$ leads to the notion of quadratic variation.

In practice, one calculates sample p-variation [48] taking differences between every m-th element of the data.

Definition 5.1 *Let $\{X_i,\ i = 0, \ldots, N\}$ be a sample of length $N + 1$. Sample p-variation $V_m^{(p)}$ for lag m is defined as*

$$V_m^{(p)} = \sum_{k=0}^{N/m-1} \left| X_{(k+1)m} - X_{km} \right|^p . \tag{5.8}$$

The following theorem describes the behaviour of the p-variation for the partial sum process of the FARIMA time series.

Theorem 5.1 *Let $\{Y_i,\ i = 0, \ldots, N\}$ be a FARIMA(p, d, q) time series with α-stable noise and $1 < \alpha \leqslant 2$. We define its partial sum process $\{X_k = \sum_{i=1}^{k} Y_i,\ k = 0, \ldots, N\}$. Then for large N/m:*

(i) if $\alpha = 2$, then $V_m^{(p)} \sim m^{Hp-1}$;

(ii) iff $1 < \alpha < 2$, then $V_m^{(p)} \sim m^{Hp-1}$ for $d > 0$ and $V_m^{(p)} \sim m^{p(H-1/\alpha)}$ for $d < 0$.

The symbol $\sim$ denotes similarity in distribution and $H = d + 1/\alpha$.

Proof. Since the process $\{N^{-H} \sum_{k=1}^{[Nt]} Y_k,\ t \geqslant 0\}$, for $H = d + 1/\alpha$, converges (with respect to all finite dimensional distributions) to a FSM $\{X(t),\ t \geqslant 0\}$ [195], which is H-selfsimilar, we can rewrite sample p-variation as

$$V_m^{(p)} \overset{D}{\sim} \sum_{k=0}^{N/m-1} |m^H (X_{k+1} - X_k)|^p = m^{Hp} \sum_{k=0}^{N/m-1} |Y_{k+1}|^p .$$

(i) Gaussian case ($\alpha = 2$ and $\mathrm{E}|Y_i|^p < \infty$ for all $p > 0$). By the law of large numbers, for large N/m $V_m^{(p)} \sim m^{Hp}(N/m)\mathrm{E}|Y_1|^p = C_1(N)m^{Hp-1}$, where $C_1(N)$ is a constant depending only on N.

(ii) Stable non-Gaussian case ($1 < \alpha < 2$ and $\mathrm{E}|Y_i|^2 = \infty$).

(a) If $p < \alpha$, by repeating arguments from the Gaussian case [51] we obtain that for large N/m: $V_m^{(p)} \sim m^{Hp}(N/m)\mathrm{E}|Y_1|^p = C_1(N)m^{Hp-1}$, where $C_1(N) = N$.

(b) If $p > \alpha$, first we rewrite the sample p-variation as

$$V_m^{(p)} = m^{Hp}(N/m)^{p/\alpha}\frac{1}{(N/m)^{p/\alpha}}\sum_{k=0}^{N/m-1}|Y_{k+1}|^p.$$

By [120], for large N/m we obtain that

$$V_m^{(p)} \stackrel{D}{\sim} m^{Hp}(N/m)^{p/\alpha}S_{\alpha/p} = C_2(N)m^{pd}S_{\alpha/p},$$

where $S_{\alpha/p}$ is an α/p-stable random variable, $C_2(N)$ a constant depending only on N and $d = H - 1/\alpha$. □

Since the normalized partial sum FARIMA process converges to a FSM with $H = d + 1/\alpha$ [195], we obtain the following fact.

Corollary 5.1 *Let $\{X_i,\ i = 0, \ldots, N\}$ be a FSM with $0 < H < 1$ and $1 < \alpha \leqslant 2$. Then for large N/m:*

(i) if $\alpha = 2$, then $V_m^{(p)} \sim m^{Hp-1}$;

(ii) if $1 < \alpha < 2$, then $V_m^{(p)} \sim m^{Hp-1}$ for $H > 1/\alpha$ and $V_m^{(p)} \sim m^{p(H-1/\alpha)}$ for $H < 1/\alpha$.

The symbol $\sim$ denotes similarity in distribution.

This implies that in the case of Gaussian FARIMA or FBM for $p > 1/H$ sample p-variation is a strictly increasing function of m (it tends to zero as m gets smaller), whereas for $p < 1/H$ it is a strictly decreasing function of m (it diverges to infinity when m gets smaller). For a stable non-Gaussian FARIMA or FSM the situation differs from the Gaussian case and depends on whether d is positive or negative. It appears that the sample p-variation is always a decreasing function with respect to m when $d < 0$. If $d > 0$, the situation is the same as in the Gaussian case: if $p > 1/H$, then sample p-variation is an increasing function of m, if $p < 1/H$ it is a decreasing function of m.

This suggests a p-variation dynamics test for checking both weather a given sample follows the self-similar dynamics and, in the the negative memory case, distinguishing between a Gaussian FARIMA or FBM (light-tailed models) and a stable non-Gaussian FARIMA or FSM (heavy-tailed models) [48].

Algorithm 5.1 *p-variation dynamics test*

1. *For the positive memory case, sample p-variation as a function of m should be monotonic and change its behaviour from an increasing function to a decreasing one for some $p = 1/H$, $0 < H < 1$.*
2. *In the negative memory case, light-tailed and heavy-tailed dynamics differ.*
   (a) *If the underlying model is a Gaussian FARIMA or FBM, then sample p-variation should behave exactly as in the positive memory case, namely change its monotonic character from a decreasing to an increasing one for some $p = 1/H$, $0 < H < 1$.*
   (b) *If the underlying model is a non-Gaussian FARIMA or FSM, then sample p-variation as a function of m should be a decreasing function for all $p = 1/H$, $0 < H < 1$ (in practice, for simulated samples from these processes one can observe that for large p the function has a quite chaotic character for moderate sample sizes).*

In Figures 5.2–5.4 we illustrate the behaviour of sample p-variation for positive and negative cases for both FBM and FSM. The simulated samples are of length 2^{10}. In the positive memory case (Figures 5.3 and 5.4), sample p-variation as a function of m is monotonic and changes its behaviour from an increasing function to a decreasing one for some $p = 1/H$, $0 < H < 1$ for both FBM and FSM.

In the negative memory case (Figures 5.2 and 5.5) p-variation dynamics test allows to distinguish between a FBM (a light-tailed model) and a FSM (a heavy-tailed model). If the underlying model is a FBM, then sample p-variation behaves exactly as in the positive memory case, namely changes its monotonic character from a decreasing to an increasing one for some $p = 1/H$, $0 < H < 1$. The behaviour within a FSM model is different, sample p-variation as a function of m should be a decreasing function for all $p = 1/H$, $0 < H < 1$. However, we can observe in Figure 5.5, that for simulated time series from a FSM, for large p, the function has a quite chaotic character for such a moderate sample size.

Estimation algorithm

Now, we introduce a method of estimation of the self-similarity index H based on sample p-variation for the FARIMA process and fractional stable motion. For

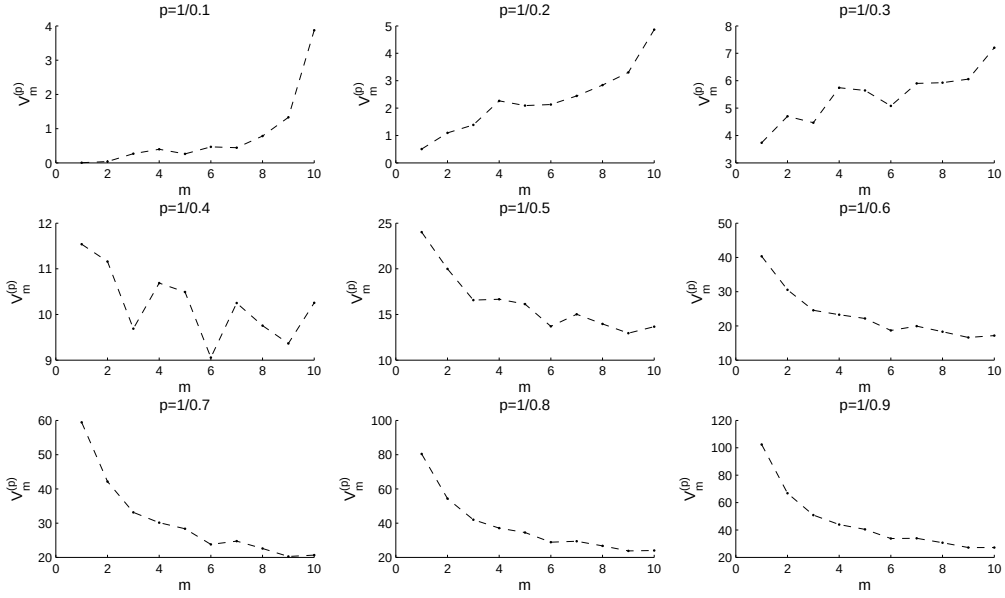

Figure 5.2: Sample p-variation for $m = 1, 2, \ldots, 10$, $H = 0.1, 0.2, \ldots, 0.9$ for a simulated trajectory of the FBM with $H = 0.4$ (negative memory case). We observe that in the first three panels, sample $1/H$-variation increases, it stabilizes in the fourth panel (corresponding to $H = 0.4$), and it decreases in the subsequent panels.

$d > 0$ the method applies to both Gaussian ($\alpha = 2$) and non-Gaussian ($\alpha < 2$) cases. For $d < 0$ it is defined only for the Gaussian case. The idea of the method is to find $p = 1/H$ for which $V_m^{(p)}$ as a function of m changes its monotonic behaviour, i.e. becomes constant. To this end, as a simple tool, we propose to calculate difference between $V_m^{(p)}$ for the smallest m, namely $m = 1$, and $V_m^{(p)}$ for a larger m (but not too large to make sure N/m is sufficiently high), e.g., $m = 8$. This is done for different p's. The smallest distance defines the value of the estimator $\hat{p}$ and consequently $\hat{H}$. The procedure can be summarized as follows.

Algorithm 5.2 *Estimation of the self-similarity index based on sample p-variation*

1. *Calculate* $V_m^{(p)}$ *for* $p = 1./(0.01 : 0.01 : 0.99)$ *and* $m = 1$ *and* 8.

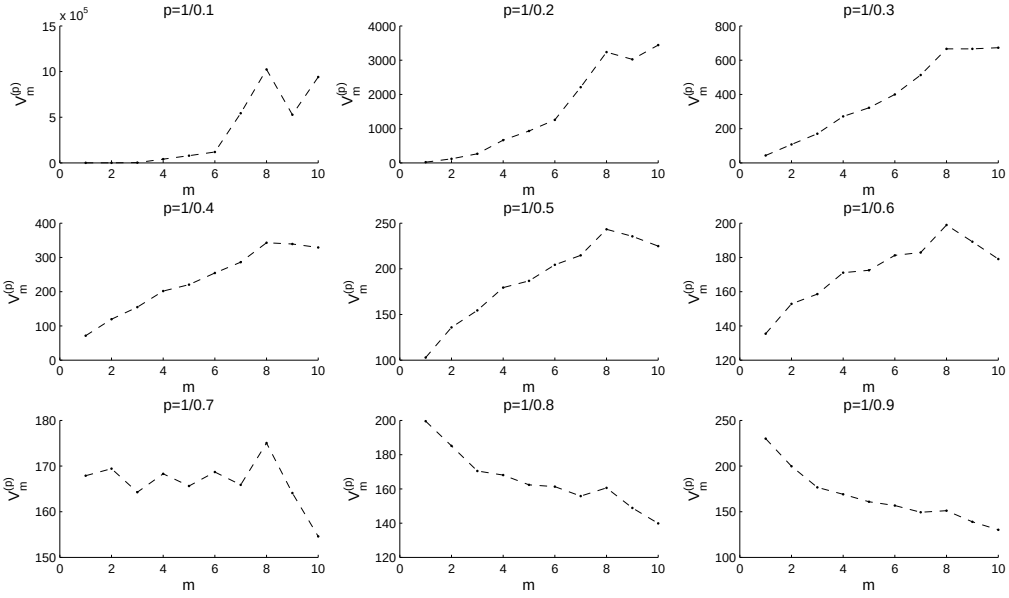

Figure 5.3: Sample p-variation for $m = 1, 2, \ldots, 10$, $H = 0.1, 0.2, \ldots, 0.9$ for a simulated trajectory of the FBM with $H = 0.7$ (positive memory case). We observe that in the first six panels, sample $1/H$-variation increases, it stabilizes in the seventh panel (corresponding to $H = 0.7$), and it decreases in the subsequent panels.

2. Find $\hat{p}$ that minimizes $\left(V_8^{(p)} - V_1^{(p)}\right)^2$.

3. The estimated value $\hat{H} = 1/\hat{p}$.

Let us illustrate the behaviour of the function $\left(V_8^{(p)} - V_1^{(p)}\right)^2$ for a sample trajectory of FBM with $H = 0.65$, see Figure 5.6. We can see that the function has a parabolic shape with vertex around the point 0.65. This justifies the idea behind the estimator.

Next, we study the consistency of the estimator by means of the Monte Carlo method. To this end we performed simulations for sample trajectories of the FBM with $H = 0.65$. We simulated 100 trajectories of length $N = 60000$. Next, we calculated values of the self-similarity estimator for increasing subsets of the trajectories, namely for the first $n = 100$, $n = 1000$, $n = 10000$, $n = 30000$, and $n = 60000$ (i.e. the whole trajectory) points. The results are presented in the form of box plots in Figure 5.7. We can see that with the increase of subset's length n,

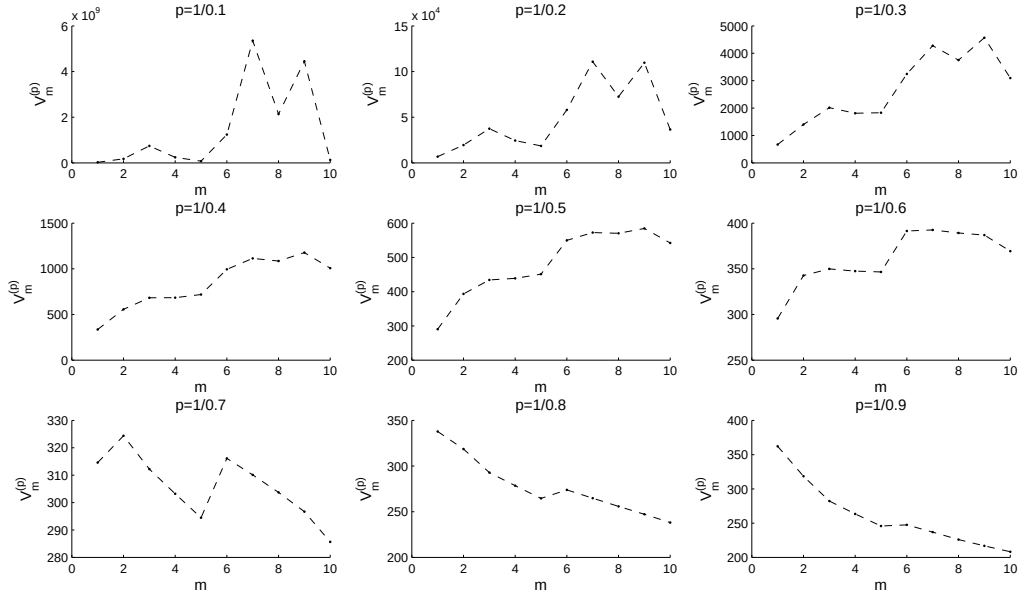

Figure 5.4: Sample p-variation for $m = 1, 2, \ldots, 10$, $H = 0.1, 0.2, \ldots, 0.9$ for a simulated trajectory of the FSM with $H = 0.7$ and $\alpha = 1.85$ (positive memory case). We observe that in the first six panels, sample $1/H$-variation increases, it stabilizes in the seventh panel (corresponding to $H = 0.7$), and it decreases in the subsequent panels.

the estimator converges to the true value as its variance decreases. This suggests that the estimator is consistent.

Finally, by means of Monte Carlo simulations we study efficiency of the method for both Gaussian and non-Gaussian stable cases. Moreover, we compare the results with another method of estimation of H, namely FIRT. First, we conduct the analysis for 1000 simulated trajectories of FBM of length $N = 1000$ with $H = 0.35$ and $H = 0.65$, see Figures 5.8–5.9. We notice that the p-variation method works remarkably well for the FBM. The medians (and means which were additionally checked) are close the true values and the variance of the estimator is very low.

Next, we performed similar calculations for the FARIMA$(0, -0.25, 0)$ process with Gaussian noise, see Figure 5.10. We can notice that the self-similarity parameter calculated by the p-variation method is overestimated. The same is true for the FIRT estimator. However, we should remember that the FARIMA process is self-similar only in the limit. Moreover, the simulation method of FARIMA is not exact as it was for the FBM case. With respect to variance, the p-variation

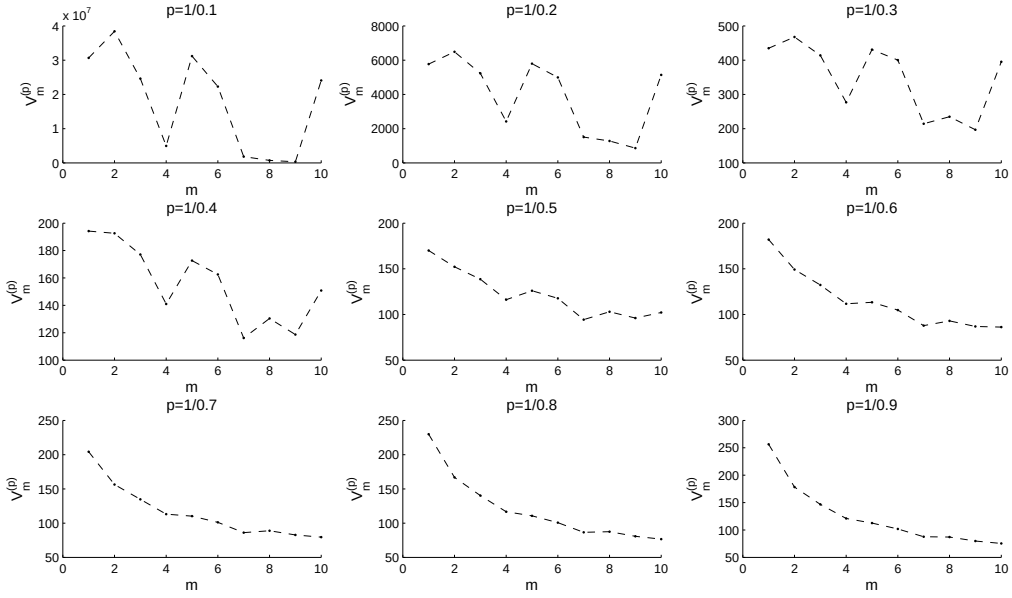

Figure 5.5: Sample p-variation for $m = 1, 2, \ldots, 10$, $H = 0.1, 0.2, \ldots, 0.9$ for a simulated trajectory of the FSM with $H = 0.4$ and $\alpha = 1.85$ (negative memory case). We observe that in the first three panels, sample $1/H$-variation does not show any clear trend, and it decreases in the subsequent panels.

method is, as before, the clear winner. Similar conclusions were drawn for other values of d and for the general Gaussian FARIMA(p, d, q) case. Monte Carlo analysis of the stable case, namely FSM and FARIMA with non-Gaussian stable noise for $\alpha > 1$ and $d > 0$, led to similar conclusions but overestimation changed to underestimation, e.g. see Figure 5.11, where FARIMA$(0, 0.25, 0)$ process with stable noise with $\alpha = 1.8$ was considered.

To sum up, the introduced p-variation estimator seems to be consistent and works remarkably well at least for the FBM case. We believe it is also true for other analyzed cases but they need more attention due to approximation errors caused by simulation methods and not exact but limiting results. What is definitely true is the fact that the variance of the estimator is really low when compared with the well-known FIRT method of estimation of H.

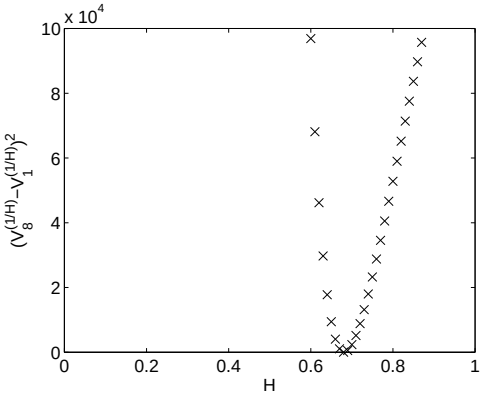

Figure 5.6: Plot of $(V_8^{(p)} - V_1^{(p)})^2$ for $H = 0.1, 0.2, \ldots, 0.9$ ($p = 1./(0.01 : 0.01 : 0.99)$) for a simulated trajectory of the FBM $\{B_H(t),\ t = 1, 2, \ldots, 1000\}$ with $H = 0.65$ (positive memory case). We observe that it resembles a parabola and changes from a decreasing function to the increasing around $H = 0.65$.

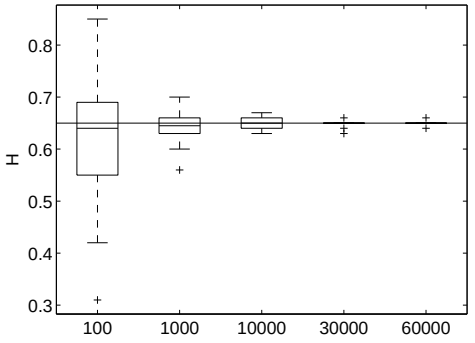

Figure 5.7: The box plots of the p-variation estimator for FBM with $H = 0.65$ for five trajectory subsets: $n = 100$, $n = 1000$, $n = 10000$, $n = 30000$, and $n = 60000$ of total length $N = 60000$. The horizontal line corresponds to the true value of H.

5.3.3 BMW^2 test

We note that an estimator of H gives information on the self-similarity and not on the distribution of the process. However, applying the BMW^2 test [213, 150] one can identify the stability index α of the data. The test applies the concept of surrogate data [56], which refers to data that preserve certain linear statistic properties of the experimental time series, without the deterministic component.

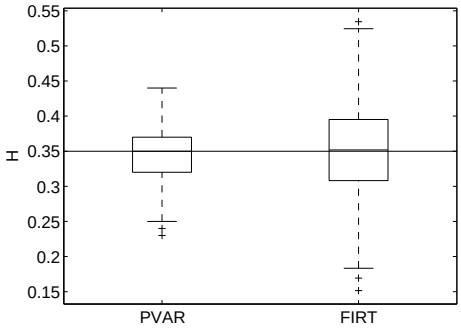

Figure 5.8: Box plots of sample p-variation and FIRT estimators for 1000 simulated trajectories of the FBM $\{B_H(t),\ t = 1, 2, \ldots, 1000\}$ with $H = 0.35$ (negative memory case). We observe that medians of the estimators are similar and close to $H = 0.35$ but the former method is vastly superior with respect to variance.

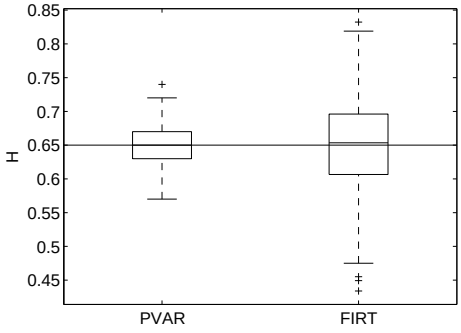

Figure 5.9: Box plots of sample p-variation and FIRT estimators for 1000 simulated trajectories of the FBM $\{B_H(t),\ t = 1, 2, \ldots, 1000\}$ with $H = 0.65$ (positive memory case). We observe that medians of the estimators are similar and close to $H = 0.65$ but the former method is vastly superior with respect to variance.

The surrogate data can be obtained by several different ways. Here, we obtain them by random shuffling of the original data positions.

According to [213] we have the following BMW^2 computer test.

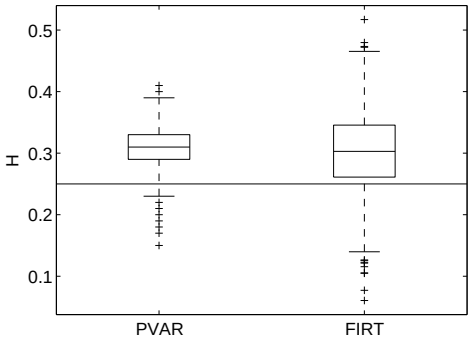

Figure 5.10: Box plots of sample p-variation and FIRT estimators for 1000 simulated trajectories of the FARIMA$(0, -0.25, 0)$ $\{X_t,\ t = 1, 2, \ldots, 1000\}$ with Gaussian noise (negative memory case). We observe that medians of the estimators are similar but the self-similarity parameter is now overestimated. This is mostly due to the fact that the FARIMA process is self-similar but only asymptotically. The variance of the former method is much lower.

Algorithm 5.3 *BMW^2 computer test to distinguish between classical and fractional dynamics*

1. *If the data come from FBM, then the values of the applied self-similarity estimator should change to 1/2 for the surrogate data independently on the initial values.*
2. *If the data come from BM or SM, then the estimator values should be the same for the original and surrogate data and equal to* $1/2$ *and* $1/\alpha$*, respectively.*
3. *If the data come from FSM, then we should observe a change to* $1/\alpha$ *in the estimators values.*

The test can be used to distinguish between diffusion (BM and SM) and fractional (anomalous) diffusion (FBM and FSM) models. This is illustrated in Figure 5.12. Comparing Figures 5.1 and 5.12 we can observe a change of the value of the estimator from 0.39 to 0.56. This clearly suggests FSM as the underlying model, which is true since the data come from FSM with $H = 0.4$ and $\alpha = 1.85$. Moreover, the FIRT estimator for the surrogate data yields a good approximation of the $1/\alpha$ value.

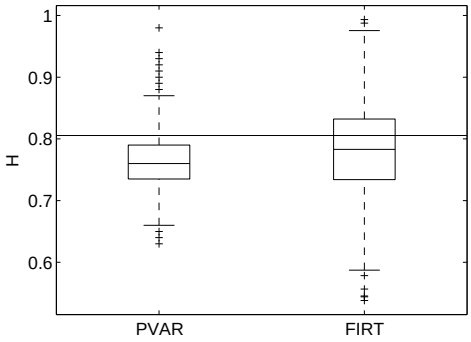

Figure 5.11: Box plots of sample p-variation and FIRT estimators for 1000 simulated trajectories of the FARIMA$(0, 0.25, 0)$ $\{X_t,\ t = 1, 2, \ldots, 1000\}$ with stable noise with $\alpha = 1.8$ (positive memory case). We observe that medians of the estimators are similar but the self-similarity parameter is now underestimated. This is mostly due to the fact that the FARIMA process is self-similar but only asymptotically. The variance of the former method is much lower.

5.4 Long memory

The notion of long memory (or long-range dependence) has intrigued many at least since Mandelbrot brought it to the attention of the scientific community in the 1960s in a series of papers, see, e.g., [140, 141], that, among other things, explained the so-called Hurst phenomenon, having to do with unusual behaviour of the water levels in the Nile river [102]. Today this notion has become especially important as potentially crucial applications arise in new areas such as communication networks [14, 164, 74, 222, 61] and finance [132].

In modelling long-memory phenomena, the stationary increments of H-self-similar processes are of interest [72, 186]. Any H-self-similar process with stationary increments $\{X(t)\}_{t\geqslant 0}$ induces a stationary sequence $\{Y_j\}_{j\in\mathbb{Z}}$ where $Y_j = X(j+1) - X(j);\ j = 0, 1, \ldots$ which is called the fractional noise.

The fractional noise has some remarkable properties. Let us concentrate now on the finite variance (e.g., Gaussian) case. In this case, if $H = 1/2$, then its autocovariance function $r(k) = 0$ for $k \neq 0$ and hence it is the sequence of uncorrelated random variables (in the Gaussian case this is equivalent to independence). The situation is quite different when $H \neq 1/2$, namely the Y_j's are correlated (dependent) and the time series has the autocovariance function of the power-law form:

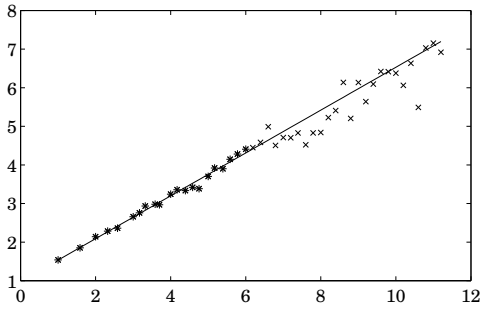

Figure 5.12: FIRT estimator for a time series simulated for the surrogate date from FSM with $H = 0.4$ and $\alpha = 1.85$ (for the original data see Figure 5.1). The estimator H_{FIRT} is obtained by fitting a least-squares line to the values of the coefficients $d_{FIRT}(n)$ on a log-linear scale. The points at the high end (marked by "x") are not used to fit the line. The estimated slope of the line for the surrogate data is $H_{FIRT} = 0.56$, which corresponds to $1/\alpha \sim 0.54$.

$$r(k) \sim \mathrm{Var} Y_1 \, H(2H-1)k^{2H-2}, \quad \text{as } \; k \to \infty. \tag{5.9}$$

The autocovariance function $r(k)$ tends to 0 as $k \to \infty$ for all $0 < H < 1$ as a power function hence much slower than exponentially. For example, the common linear (ARMA) processes, GARCH processes, finite state Markov chains all lead to exponentially fast decaying correlations [30]. Moreover, when $1/2 < H < 1$ it tends to zero so slowly that the sum $\sum_{k=1}^{\infty} r(k)$ diverges. We say that in this case the increment process exhibits long memory (long-range dependence, persistence) [13]. We also note that the coefficient $H(2H-1)$ is positive, so the $r(j)$'s are positive for all large j, a behaviour referred to as "positive dependence". Furthermore, formula (5.9) by the Wiener Tauberian theorem [230] implies that the spectral density $h(\lambda)$ of the stationary process has a pole at zero. A phenomenon often referred to as "$1/f$ noise".

If $0 < H < 1/2$, then $\sum_{k=1}^{\infty} |r(k)| < \infty$ and the spectral density tends to zero as $|\lambda| \to 0$. Furthermore, the coefficient $H(2H-1)$ is negative and the $r(j)$'s are negative for all large j. We say in that case that the sequence displays negative long-range (power-like) dependence (antipersistence).

For processes with $\alpha < 2$ the covariance can be replaced by the codifference

$$\tau(Y_t, Y_0) = \ln \mathrm{E} e^{i(Y_t - Y_0)} - \ln \mathrm{E} e^{iY_t} - \ln \mathrm{E} e^{-iY_0},$$

which reduces to the covariance in Gaussian case, for details, see [188].

In the literature different methods of assessing long-range dependence and estimating the memory parameter d have been developed [13, 205, 150, 70]. It is important to realize what are the assumptions and limitations of various tools and what is the exact output of different estimators. This also refers to self-similarity index estimators. For example, a very well-know RS method, in the general stable case, does not return H, which is true only in the Gaussian case, but the value $d + 1/2$, where $d = H - 1/\alpha$ and α is the index of stability [150].

5.4.1 RS estimator

Our study of memory estimators begins with the RS (R/S) method [102, 218]. It is the most classical and well-known method for estimation of the self-similarity parameter H. In the case of Gaussian, i.e. finite variance, time series the result of the method is indeed an estimate of H. The relationship with the memory parameter d is $H = d + 1/2$. For α-stable (infinite variance) time series the method does not return H, only $d + 1/2$, where $d = H - 1/\alpha$ [205]. Hence, in the general stable case it is a method of estimation of d. Naturally, if $\alpha = 2$, this reduces to H.

We begin with dividing a stationary series of length N into blocks of equal length n. Then for each subseries ($m = 1, 2, \ldots, [N/n]$) the $(RS)_m$ statistic is computed according to the formula:

$$(RS)_m = \frac{\max_{1 \leqslant k \leqslant n} \sum_{i=1}^{k} (X_i - \overline{X}_m) - \min_{1 \leqslant k \leqslant n} \sum_{i=1}^{k} (X_i - \overline{X}_m)}{S_m}, \tag{5.10}$$

where $\overline{X}_m$ and S_m are sample mean and sample standard deviation of the mth subseries respectively. Finally, the sample mean of the $(RS)_m$ statistics over blocks is calculated and denoted by $(RS)_n$. After such procedure, we plot the $(RS)_n$ statistic against n on a log-log scale with the fitted least squares line. The slope should be equal to $d + 1/2$. Efficiency of the method is studied in Section 5.4.3.

5.4.2 Modified RS estimator

Next, we consider a Lo's modified RS statistic (MRS) V_q [133] defined by

$$V_q = \sqrt{N}(RS_q)_N,$$

which is a version of the formula (5.10) for $n = N$ with the same numerator and modified denominator

$$S_q = \sqrt{S^2 + 2 \sum_{i=1}^{q} \omega_i(q) \hat{\gamma}_i},$$

where $\hat{\gamma}_i$ are the sample autocovariances and weights $\omega_i(q)$ are given by

$$\omega_i(q) = 1 - \frac{i}{q+1}, \qquad q < N.$$

Plotting V_q with respect to q on a log-log scale and fitting the least squares line, leads, for the finite variance case, to an estimate of H, namely the slope of the line equals $1/2 - H$. One may check that, as it was the case for RS statistic, that the formula for the α-stable distribution can be generalized to $-d$, where $d = H - 1/\alpha$.

In the case of absence of long memory, Lo showed in [133] that for the finite variance case, under the right choice of q,

$$V_q \xrightarrow{D} W_1 = \max_{0 \leqslant t \leqslant 1} W_0(t) - \min_{0 \leqslant t \leqslant 1} W_0(t),$$

where W_0 is a standard Brownian bridge. Since the distribution of the random variable W_1 is known, Lo proposed testing the null hypothesis

$$\mathcal{H}_0 = \{\text{no long memory, i.e. } d = 0\}$$

with 95% (asymptotic) acceptance region $[0.809, 1.862]$. Therefore, MRS method can serve as a validation procedure. We note here that the right choice of q is an open problem and the test statistic V_q has a strong bias toward accepting the null hypothesis [206].

For the infinite variance case, there are yet no analytical results for the asymptotics of V_q. However, under the assumption the analyzed time series is a sequence of i.i.d. stable random variables, one can calculate the confidence interval applying Monte Carlo simulations [35]. Efficiency of the method is studied in Section 5.4.3.

5.4.3 New estimator based on sample mean-squared displacement

The sample mean-squared displacement (MSD) was introduced in [48].

Definition 5.2 *Let $\{X_i,\ i = 0, \ldots, N\}$ be a sample of length $N+1$. Sample MSD $M_N(\tau)$ for lag τ is defined as*

$$M_N(\tau) = \frac{1}{N - \tau + 1} \sum_{k=0}^{N-\tau} (X_{k+\tau} - X_k)^2. \tag{5.11}$$

Note the essential difference between increments of X_t in equations (5.8) and (5.11). The former is a moving window, whereas the latter in defined on disjoint

subintervals (blocks). The sample MSD is a time average MSD on a finite sample regarded as a function of difference τ between observations. It is a random variable in contrast to the ensemble average which is deterministic. The following theorem describes the behaviour of the sample MSD for the partial sum process of the FARIMA time series.

Theorem 5.2 *Let $\{Y_i,\ i=0,\dots,N\}$ be a FARIMA(p,d,q) time series with α-stable noise and $1<\alpha\leqslant 2$. We define its partial sum process $\{X_k=\sum_{i=1}^{k}Y_i,\ k=0,\dots,N\}$. Then for large N/τ:*

$$M_N(\tau)\overset{D}{\sim}\tau^{2d+1}.$$

Proof. We rewrite the sample MSD as

$$\begin{aligned}M_N(\tau) &= \frac{1}{N-\tau+1}\Bigg\{\sum_{k=0}^{N/\tau-1}\Big(X_{(k+1)\tau}-X_{k\tau}\Big)^2+\sum_{k=0}^{N/\tau-2}\Big(X_{(k+1)\tau+1}-X_{k\tau+1}\Big)^2\\ &+ \dots+\sum_{k=0}^{N/\tau-2}\Big(X_{(k+1)\tau+\tau-1}-X_{k\tau+\tau-1}\Big)^2\Bigg\}.\end{aligned}$$

The process $\{N^{-H}\sum_{k=1}^{[Nt]}Y_k,\ t\geqslant 0\}$, for $H=d+1/\alpha$, converges (with respect to all finite dimensional distributions) to a FSM $\{X(t),\ t\geqslant 0\}$ [195], which is H-selfsimilar. Hence, if $\alpha=2$, then by the law of large numbers [195], for large N/τ

$$M_N(\tau)\overset{D}{\sim}\tau^{2H}\mathrm{E}Y_1^2=\tau^{2(H-1/2)+1}\mathrm{E}Y_1^2=\tau^{2d+1}\mathrm{E}Y_1^2.$$

For $\alpha<2$, by [120], for large N/τ,

$$\begin{aligned}M_N(\tau) &\overset{D}{\sim} \frac{1}{N-\tau+1}\left\{\tau^{2H}(N/\tau)^{2/\alpha}S_{\alpha/2}+(\tau-1)\tau^{2H}(N/\tau-1)^{2/\alpha}S_{\alpha/2}\right\}\\ &\overset{D}{\sim} \tau^{2H}\tau(N/\tau)^{2/\alpha}S_{\alpha/2}=C(N)\tau^{2d+1}S_{\alpha/2},\end{aligned}$$

where $C(N)=N^{2/\alpha}$ and $S_{\alpha/2}$ is an $\alpha/2$-stable random variable with the skewness parameter $\beta=1$. Therefore, $S_{\alpha/2}$ is only one-sided. □

Since the normalized partial sum FARIMA process converges to a FSM with $H=d+1/\alpha$ [195], we obtain the following fact [48].

Corollary 5.2 *Let* $\{X_i,\ i = 0, \ldots, N\}$ *be a FSM with* $0 < H < 1$ *and* $1 < \alpha \leqslant 2$. *Then, for large* N,

$$M_N(\tau) \overset{D}{\sim} \tau^{2d+1},$$

where $d = H - 1/\alpha$.

In particular, for a FBM we obtain the well-known result that $M_N(\tau) \sim \tau^{2H}$, and for both BM and LSM we arrive at the diffusion case, namely $M_N(\tau) \sim \tau$ since $d = 0$.

As a consequence, we see that the memory parameter d controls the type of anomalous diffusion [48]. If $d < 0$ ($H < 1/\alpha$), so in the negative dependence case, the process follows the subdiffusive dynamics, if $d > 0$ ($H > 1/\alpha$), the character of the process changes to superdiffusive. Moreover, it appears that α-stable processes for $\alpha < 2$ can serve both as examples of subdiffusion and superdiffiusion. This is illustrated in Figure 5.13. The subdiffusion pattern arises when the dependence is negative, so possible large positive jumps are quickly compensated by large negative jumps, and on average the process travels shorter distances than the light-tailed Brownian motion.

Estimation algorithm

Now, we introduce a method of estimation of the memory parameter d based on the notion of sample MSD. It can be applied to FARIMA time series and fractional stable noise for both Gaussian ($\alpha = 2$) and non-Gaussian ($\alpha < 2$) cases. For the FARIMA case see also [35].

Algorithm 5.4 *Estimation of the memory parameter based on sample MSD*

1. *Calculate* $M_N(\tau)$ *for* $\tau = 1, 2, \ldots, 10$.
2. *Fit the linear regression line according to* $\ln(M_N(\tau)) = \ln(C) + a \ln(\tau)$, $\tau = 1, 2, \ldots 10$, *where* C *and* a *are constants.*
3. *The estimated value* $\hat{d} = (a - 1)/2$.

Next, we study the consistency of the estimator by means of the Monte Carlo method. To this end we performed simulations for sample trajectories of the FBM with $d = 0.2$ ($H = 0.7$). We simulated 100 trajectories of length $N = 60000$. Next, we calculated values of the memory estimator for increasing subsets of the trajectories, namely for the first $n = 100$, $n = 1000$, $n = 10000$, $n = 30000$, and

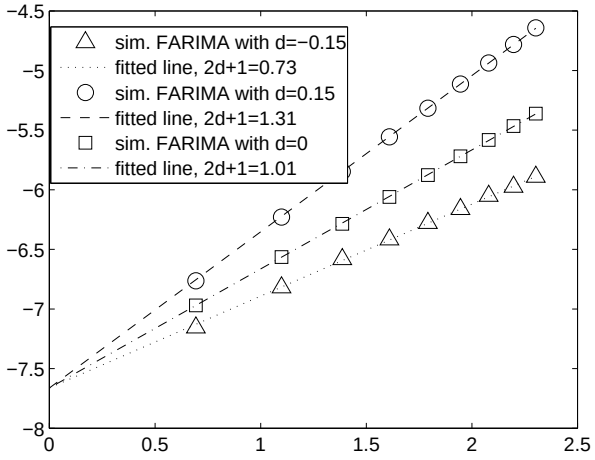

Figure 5.13: Sample MSD for a simulated trajectory from FARIMA with $d = -0.15$ (triangles) – negative memory case, a simulated trajectory from FARIMA with $d = 0.15$ (circles) – positive memory case, and a simulated trajectory from FARIMA with $d = 0$ (squares) – no memory case in double logarithmic scale. Estimated exponents equal 0.73 (corresponding dotted line), 1.31 (dashed line), and 1.01 (dash-dot line), respectively.

$n = 60000$ (i.e. the whole trajectory) points. The results are presented in the form of box plots in Figure 5.14. We can see that with the increase of subset's length n, the estimator converges to the true value as its variance decreases. This suggests that the estimator is consistent.

Finally, by means of Monte Carlo simulations we study efficiency of the method for both Gaussian and non-Gaussian stable cases. Moreover, we compare the results with another methods of estimation of H, namely RS and MRS. First, we conduct the analysis for 1000 simulated trajectories of FBM of length $N = 1000$ with $d = -0.2$ ($H = 0.3$) and $d = 0.2$ ($H = 0.7$), see Figures 5.15–5.16. We can see that the sample MSD method works remarkably well for the FBM. The medians (and means which were additionally checked) of the estimator coincide well with the true values. MRS method seems superior to RS in the negative memory case, whereas for positive d it is quite opposite. The variance of the sample MSD and RS estimators are the lowest.

Next, we performed similar calculations for the FARIMA$(0, -0.25, 0)$ process with Gaussian noise, see Figure 5.17. We can notice that the memory parameter estimates are now overestimated. However, we should remember that the FARIMA process is self-similar only in the limit. Moreover, the simulation method of

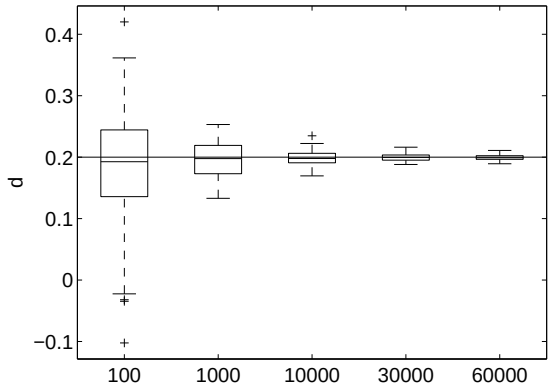

Figure 5.14: The box plots of the sample MSD estimator for FBM with $d = 0.2$ ($H = 0.7$) for five trajectory subsets: $n = 100$, $n = 1000$, $n = 10000$, $n = 30000$, and $n = 60000$ of total length $N = 60000$. The horizontal line corresponds to the true value of d.

FARIMA is not exact as it was for the FBM case. MRS method seem to be the winner closely followed by the sample MSD. With respect to variance, sample MSD and RS methods are superior to the MRS. Similar conclusions were drawn for other values of d and for the general Gaussian FARIMA(p, d, q) case. Monte Carlo analysis of the stable case, namely FSM and FARIMA with non-Gaussian stable noise, for negative d, led to similar conclusions, see Figure 5.18, where FARIMA$(0, -0.2, 0)$ process with stable noise with $\alpha = 1.8$ was considered. For positive d the best methods are sample MSD and RS, see Figure 5.19, where FARIMA$(0, 0.2, 0)$ process with stable noise with $\alpha = 1.8$ was considered.

Summarizing, the introduced sample MSD estimator seems to be consistent and works remarkably well at least for the FBM case. Other analyzed cases need more attention due to approximation errors caused by simulation methods and not exact but limiting results. What is definitely true is the fact that the variance of the estimator is really low and compared with the well-known method RS. Moreover, for positive d RS method works better than the MRS, whereas for negative d the situation is opposite.

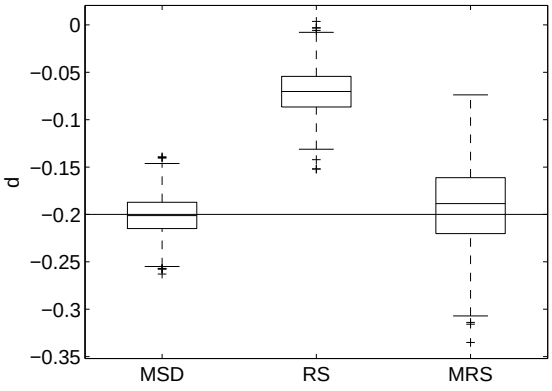

Figure 5.15: Box plots of sample MSD, RS and MRS estimators for 1000 simulated trajectories of the FBM $\{B_H(t),\ t = 1, 2, \ldots, 1000\}$ with $H = 0.3$ (negative memory case). We observe that sample MSD and MRS estimators are accurate but the former method is superior with respect to variance.

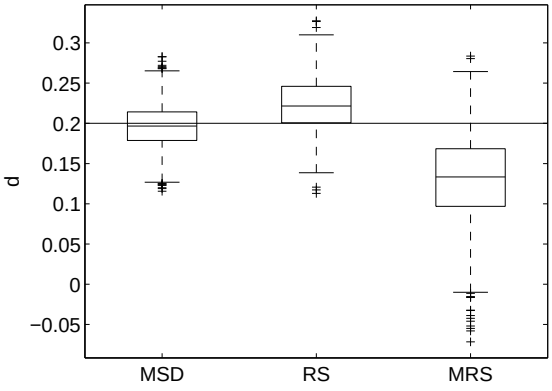

Figure 5.16: Box plots of sample MSD, RS and MRS estimators for 1000 simulated trajectories of the FBM $\{B_H(t),\ t = 1, 2, \ldots, 1000\}$ with $H = 0.7$ (positive memory case). We observe that the sample MSD method is the clear winner.

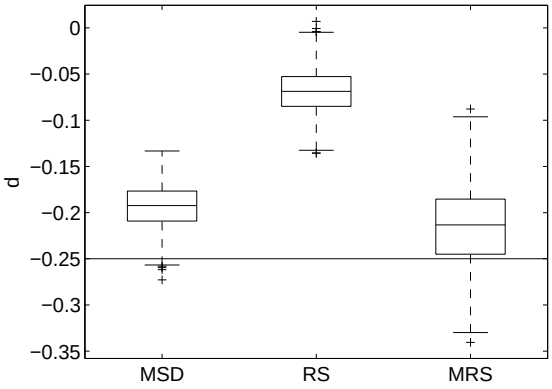

Figure 5.17: Box plots of sample MSD, RS and MRS estimators for 1000 simulated trajectories of the FARIMA$(0, -0.25, 0)$ $\{X_t,\ t = 1, 2, \ldots, 1000\}$ with Gaussian noise (negative memory case). We observe that estimators are now overestimated. This is mostly due to the fact that the FARIMA process is self-similar but only asymptotically. MRS method seems to produce the best results but the sample MSD is only slightly worse. Variances of the sample MSD and RS methods are lower than that of the MRS.

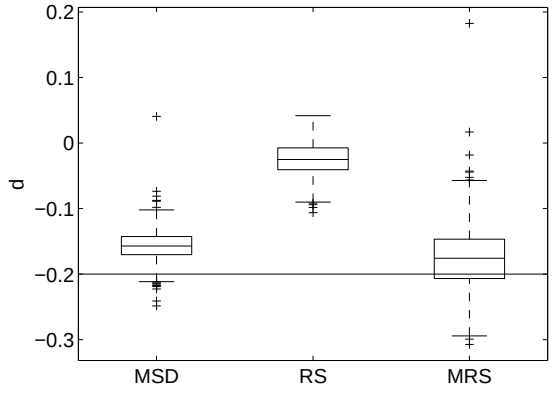

Figure 5.18: Box plots of sample MSD, RS and MRS estimators for 1000 simulated trajectories of the FARIMA$(0, -0.2, 0)$ $\{X_t,\ t = 1, 2, \ldots, 1000\}$ with stable noise with $\alpha = 1.8$ (positive memory case). We observe that estimators are now overestimated. This is mostly due to the fact that the FARIMA process is self-similar but only asymptotically. MRS method seems to produce the best results but the sample MSD is only slightly worse. Variances of the sample MSD and RS methods are lower than that of the MRS.

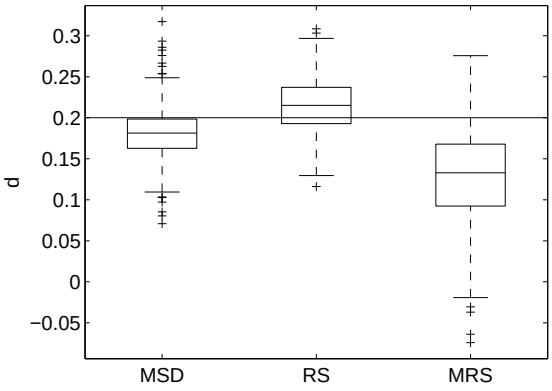

Figure 5.19: Box plots of sample MSD, RS and MRS estimators for 1000 simulated trajectories of the FARIMA$(0, 0.2, 0)$ $\{X_t,\ t = 1, 2, \ldots, 1000\}$ with stable noise with $\alpha = 1.8$ (positive memory case). We observe that, this time, RS method seems to produce the best results but the sample MSD is only slightly worse. Variances of the sample MSD and RS methods are lower than that of the MRS.

6

Modelling of fractional dynamical systems. Case studies

In this chapter we present solutions to the three sample problems outlined in Sections 2.1–2.3 which are related to the environmental protection risk areas listed in the project "Detectors and sensors for measuring factors hazardous to environment – modeling and monitoring of threats" (POIG.01.03.01-02-002/08-00) (for the detailed information about the project see Chapter 2).

6.1 Modelling of electromagnetic field

In this section we analyse a set of UMTS data which is plotted in Figure 6.1. The electromagnetic field intensity was measured in Wrocław in an urban area every minute from 12.01.2011 22:40 to 19.01.2011 21:18 (9999 observations). It has been normalized with respect to its minimum.

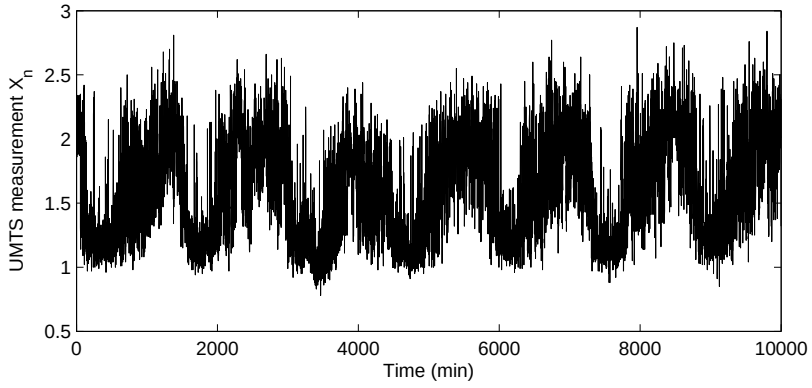

Figure 6.1: The UMTS data $\{X_n : n = 1, 2, \ldots, 9999\}$ recorded in Wrocław in 2011.

6.1.1 Seasonality and volatility in UMTS data

We denote the collected time series by $\{X_n : n = 1, 2, \ldots, 9999\}$. One can clearly notice a strong seasonal structure in time domain of the data. Therefore our first task is to remove seasonal components. In order to check seasonal periods we apply the periodogram plot, see Figure 6.2. It is a sample analog of the spectral density [30].

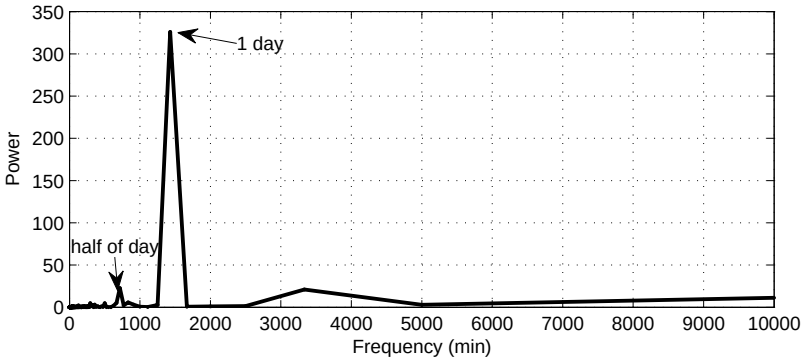

Figure 6.2: The periodogram of UMTS data. The half–day (714), day (1428) and 56 hours (3332) frequencies are strongly visible.

In Figure 6.2 we can clearly see peaks at frequencies 714, 1428 and 3332. They corresponds to cycles with periods of 12, 24 and 56 hours. Since the seasonality structure of data seems to have a sinusoidal form, we try to remove it by fitting a sum of sine functions. The proposed functions have the following form

$$S_k(x) = p_1 + \sum_{i=1}^{k} p_{2i} \sin\left[\frac{2\pi}{d_i}\,(x + p_{2i+1})\right], \quad k = 1, 2, 3,$$

where p_j's are the parameters and d_j's are the selected frequencies of the corresponding cycles.

We define an error of the fit of the function to the data $\{X_n : n = 1, 2, \ldots, 9999\}$ as

$$ERR(S_k) = \frac{1}{9999} \sum_{n=1}^{9999} (S_k(X_n) - X_n)^2.$$

In Table 6.1 we present the errors of fitting of all possible functions $S_k(x)$ for the three discovered periods: $d_1 = 714$, $d_2 = 1428$ and $d_3 = 3332$. From Table 6.1 we can see that the lowest error is for the case with sum of two sines having periods 12 and 24 hours. Hence, we remove seasonality by subtracting the values

Table 6.1: Fitted functions $S_k(x)$ with their errors $ERR(S_k)$.

Function $S_k(x)$	Periods	$ERR(S_k)$
S_1	$d_1 = 714$	0.1540
$\mathbf{S_2}$	$\mathbf{d_1 = 714}$, $\mathbf{d_2 = 1428}$	**0.0888**
S_3	$d_1 = 714$, $d_2 = 1428$, $d_3 = 3332$	0.1058

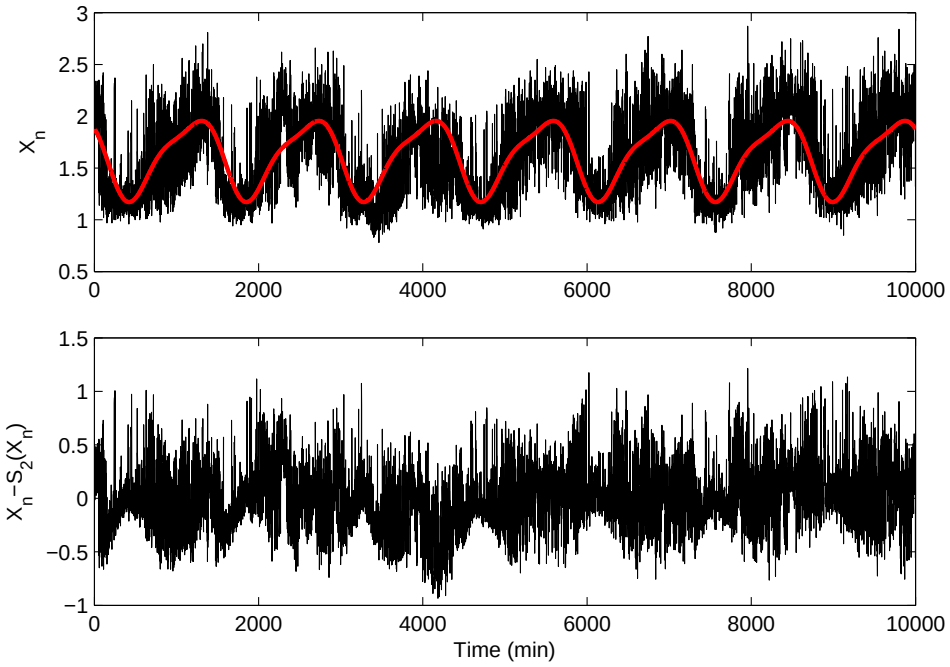

Figure 6.3: (Top panel) The UMTS data $\{X_n : n = 1, 2, \ldots, 9999\}$ with the fitted sum of sines $S_2(x)$ (grey line). (Bottom panel) The UMTS data $\left\{\hat{X}_n : n = 1, 2, \ldots, 9999\right\}$ after removal of $S_2(x)$.

of the function $S_2(x)$ with $d_1 = 714$ and $d_2 = 1428$ from the analyzed series, see Figure 6.3. We denote the new series by $\left\{\hat{X}_n : n = 1, 2, \ldots, 9999\right\}$.

After removing seasonality we check stationarity by studying volatility of the data $\left\{\hat{X}_n : n = 1, 2, \ldots, 9999\right\}$. A fluctuating variance can be observed in a squared data plot, especially when the data have mean close to 0. In our set $\left\{\hat{X}_n : n = 1, 2, \ldots, 9999\right\}$ the sample mean equals -2.515×10^{-8} and we present

the squared series $\left\{\hat{X}_n^2 : n = 1, 2, \ldots, 9999\right\}$ in Figure 6.4.

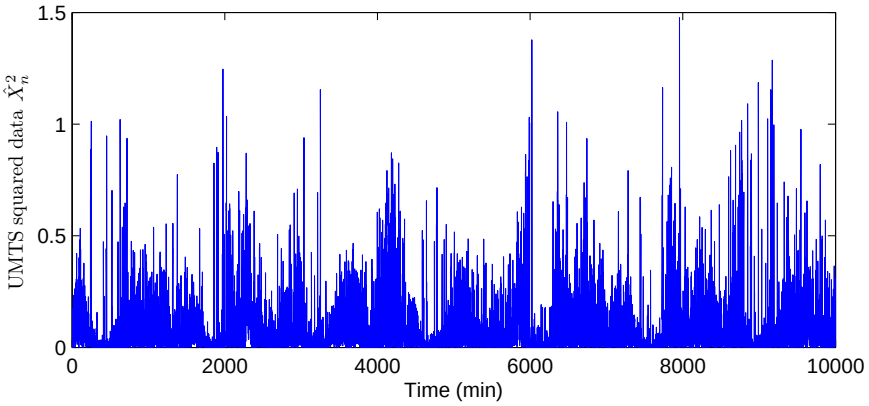

Figure 6.4: The squared UMTS data $\left\{\hat{X}_n^2 : n = 1, 2, \ldots, 9999\right\}$ with removed seasonality.

We can clearly see that volatility changes in time. In order to remove this effect and make the volatility constant we propose the following procedure based on the moving sample variance. We calculate the sample moving variance corresponding to time intervals of 61 minutes (61 observations) along the whole series $\left\{\hat{X}_n^2 : n = 1, 2, \ldots, 9999\right\}$, i.e.

$$S_k^2 := \frac{1}{60} \sum_{i=k}^{k+60} \left(\hat{X}_k - \bar{X}\right)^2,$$

for $k = 1, 2, \ldots, 9939$, see Figure 6.5 (top panel). Next, we compute the series $\left\{\tilde{X}_n := \hat{X}_{n+30}/\sqrt{S_n^2} : n = 1, 2, \ldots, 9939\right\}$. We notice that the observation $\hat{X}_n$ (for $n = 31, 32, \ldots, 9969$) is divided by the square root of sample variance S_{n-30}^2 from the time interval with 30 observations measured before and after the measurement $\hat{X}_n$. That is why we focus on observations from $n = 31$ to $n = 9969$ and therefore the length of the series $\left\{\tilde{X}_n\right\}$ is 9939. It seems that such data are closer to constant volatility, see the squared data plot in Figure 6.5 (bottom panel). For further studies we choose the data which are plotted in Figure 6.6.

In order to fit a FARIMA model to the data $\left\{\tilde{X}_n : n = 1, 2, \ldots, 9939\right\}$, first, we study the memory structure of the series. To this end we apply three methods described in Section 5.4 of estimating the memory parameter d, namely the RS, MRS and sample MSD. We obtained the following estimates: $d_{RS} = 0.3275$, $d_{MRS} = 0.3317$ and $d_{MSD} = 0.3104$. All estimators returned similar positive

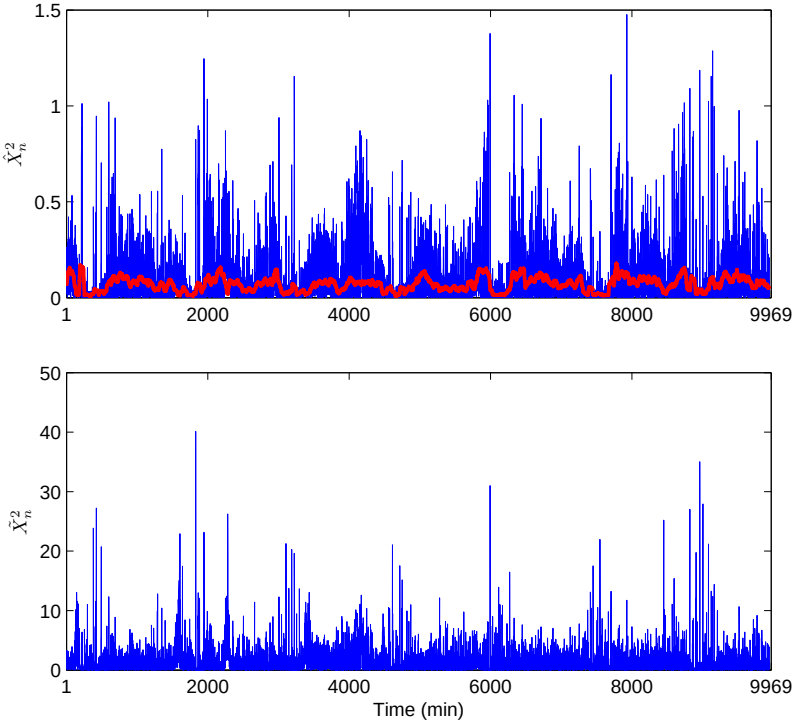

Figure 6.5: (Top panel) The squared UMTS data $\left\{\hat{X}_n^2 : n = 31, 32, \ldots, 9969\right\}$ with removed seasonality and sample variances $\left\{S_n^2 : n = 1, 2, \ldots, 9939\right\}$ (grey line). (Bottom panel) The squared UMTS data with removed seasonality and smoothed variance: $\left\{\tilde{X}_n^2 = \hat{X}_{n+30}^2 / S_n^2 : n = 1, 2, \ldots, 9939\right\}$.

values, which indicates the long memory property. This can be also observed from the ACF, see Figure 6.7.

For the model orders $p \leqslant 1$ and $q \leqslant 1$, we applied the FARIMA estimator defined in Section 4.3. The results of the estimation procedure are presented in Table 6.2. Next, we calculated corresponding residuals. It appeared that only in the case of $p = 1$ and $q = 0$ the residuals can be treated as i.i.d. (residuals were tested for randomness with the Ljung-Box test [131, 35]). Hence this suggests that the FARIMA$(1, 0.34, 0)$ model describes the data well.

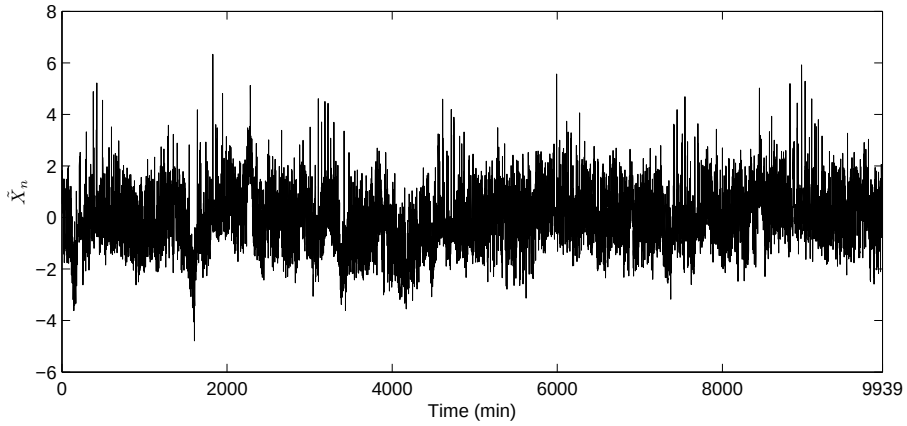

Figure 6.6: The series $\left\{\tilde{X}_n = \hat{X}_{n+30}/\sqrt{S_n^2} : n = 1, 2, \ldots, 9939\right\}$.

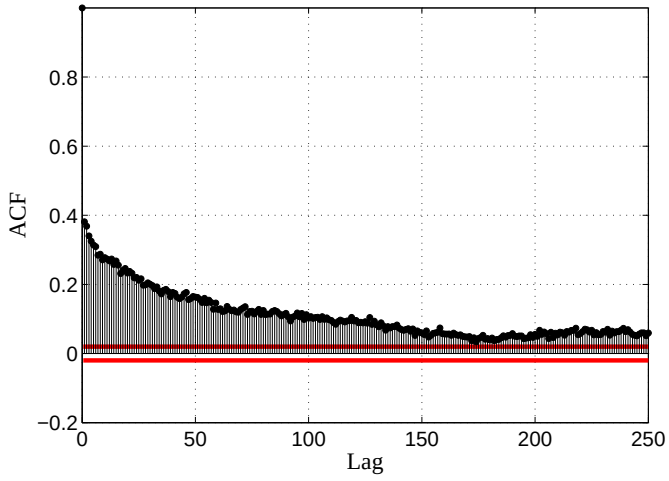

Figure 6.7: Autocorrelation function of the series $\left\{\tilde{X}_n : n = 1, 2, \ldots, 9939\right\}$ with 95% confidence interval.

6.2 Modelling of high solar flare activity

Observations of solar flare phenomena in X-rays became possible in the 1960s with the availability of space-borne instrumentation. Since 1974 broad-band soft X-ray emission of the Sun has been measured almost continuously by the meteorology satellites operated by the National Oceanic and Atmospheric Admini-

Table 6.2: Fitted FARIMA model to the UMTS data.

Model order	$\boldsymbol{\beta}_n$	FARIMA
$(p,q)=(1,0)$	(d,ϕ_1)	$(0.34,-0.17)$
$(p,q)=(0,1)$	(d,θ_1)	$(0.39,0.23)$
$(p,q)=(1,1)$	(d,ϕ_1,θ_1)	$(0.44,0.17,0.44)$

stration (NOAA) so as the Synchronous Meteorological Satellite (SMS) and the Geostationary Operational Environment Satellite (GOES).

From 1974 to 1986 the soft X-ray records are obtained by at least one from the series of GOES-type satellites. Since 1983 data from two and even three co-operating GOES are generally available. The X-ray sensor, part of the space environment monitor system aboard GOES, consists of two ion chamber detectors which provide whole-sun X-ray fluxes in the 0.05-0.3 and 0.1-0.8 nm wavelength bands. The solar soft X-ray flares data is widely available from the NOAA Space Environment Center site (http://goes.ngdc.noaa.gov/data/avg/).

At present, the accurate solar data is available for the most recent three cycles only. Individual solar cycles are different in form, amplitude and length. Hence understanding the long-term solar variability and predicting the solar activity is an actual problem for solar physics. It is associated with a variety of space weather effects [142, 197]. Solar activity is known to correlate with flare activity, and a variety of flare properties (time and strength) could be incorporated into predictions. Moreover, these disturbances pose serious threats to man-made spacecrafts, disrupt electronic communication channels and can even set up huge electrical currents in power grids [60].

Our aim is to present a proper statistical model for analyzing and predicting soft X-ray solar emission activity in the period of solar maximum. The complete solar flare data analysis for the strong activity intervals: 1978–1981, 1988–1992 and 1999–2003 was presented in [46, 192]. Interrelations among statistical flare parameters, such as the tail index α, the memory parameter d, and the self-similarity (Hurst) exponent H for X-ray flux and their evolution during solar cycles, were analyzed. It was shown that in the period of strong solar activity all parameters are nearly constant. This feature can be used for predicting the power of soft X-ray emission for the 24-th solar cycle near its solar activity maximum in future.

For the last high activity period, when energy was transmitted by X-rays emitted during blasts on a solar surface from 2000 January 1 to 2002 December 31, see Figure 6.8. Our aim is to employ identification and validation methods from Chapter 5 to check whether the FARIMA model can describe the data well.

The FARIMA process is a discrete-time analog of the fractional Langevin

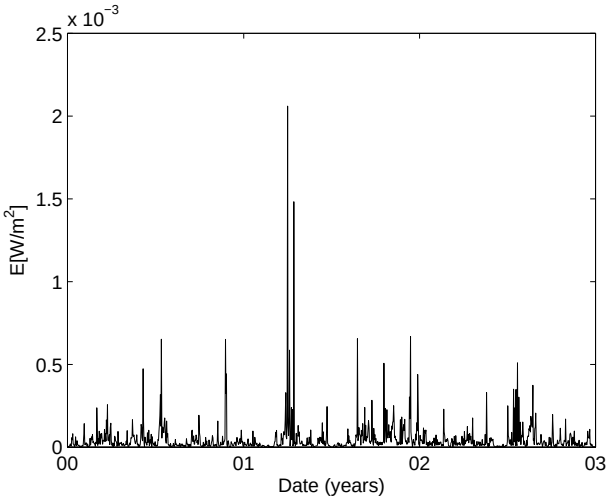

Figure 6.8: Energy-time series of solar flares from 2000 January 01 to 2002 December 31.

equation [135] that takes into account the non-Gaussian statistics and the long-range dependence (long-term memory), i.e. the events are arbitrarily distant still influence each other exceptionally strong. Briefly, the physical arguments for our subsequent consideration are the following. Processes resulting in solar flares are associated with a complicated motion of charged particles (plasma) in magnetic and electric fields. It generates the fields, accumulates their energy and transforms into the energy of flares. The classical problem for a motion of the charged Brownian particle into magnetic and electric fields can be described in the framework of the Langevin equations. However, in the strongly non-equilibrium plasmas, where turbulence is a prevailing phenomenon, the non-Gaussian stable statistics of random force becomes dominant [57]. As a consequence the solar data provide information on anomalous (non-Gaussian) diffusion and non-Maxwell stationary states.

6.2.1 Calibration of parameters for FARIMA model

The first estimation procedure of the self-similarity exponent H is the FIRT, see Section 5.3.1. For the last high activity period we have found that $H_{FIRT} = 1.14$. For the RS method (see Section 5.4.1) we have found that $d_{RS} = 0.24$.

The analysis of the data in [46] showed that the tail of the underlying distribution of the energy of solar flares conforms to the power law. Hence, we suggest to describe the data by a FARIMA model with Pareto variables since the analyzed data are positive. As the power-law distributions belong to the domain of at-

traction of stable law (see, e.g., [105]), the resulting distribution of the FARIMA process should be close to the stable one. We applied the McCulloch quantile fit [147] to obtain the parameters of the distribution. The value of α was estimated to be 1.21. One may check that the estimated value of α for simulated FARIMA times series with Pareto variables is usually underestimated. Therefore, we have assumed that the variables in our model follow the Pareto law with $\alpha = 1.25$.

According to Section 5.3.3, using an arbitrary estimator of H, in order to recover both the self-similarity exponent H and the tail index α (hence, the memory parameter d) we can apply a procedure based on the concept of surrogate data. In our case the surrogate data have been obtained by random shuffling of the original data positions. Taking into account that:

- if the process is FARIMA with finite variance (e.g., Gaussian) noise, then the values of the estimator should change to $1/2$ for the surrogate data independently on the initial values;
- if the process is FARIMA with α-stable or Pareto noise for $\alpha < 2$, then the values of the estimator should change to $1/\alpha$ for the surrogate data independently on the initial values;

we have studied the surrogated data of the empirical time series recorded from the system describing the energy of solar flares. The obtained values of the parameters are listed in Table 6.3. From the results for the surrogate data, the corresponding estimate for the parameter $1/\alpha$ are: $1/\alpha_{FIRT} = H_{FIRT} = 0.8452$. We observe that the estimators are close to the one assumed in our model: $1/\alpha_{MC} = 0.8$. Moreover, we choose $d = 0.19$ as the highest admissible value of $d < 1/\alpha$ for the FARIMA model which is close to the one obtained via the RS method for the original data.

One may notice that the estimator of H obtained via the FIRT method, see Table 6.3, is greater than theoretically admissible in the FARIMA model, i.e. it exceeds one. As stated in [39], this can be justified by performing simulations of the FARIMA processes with H close to one and estimating the parameter H on the simulated time series via different methods. It appears that a great percentage of the values of H is higher than 1. Hence, one can not reject the hypothesis that the underlying model is a FARIMA$(0, d, 0)$ process.

Finally, we look for an enhanced model FARIMA(p, d, q) which would describe better the behaviour of different estimators obtained for the original and shuffled data, and, in a consequence, would improve the fit in terms of the prediction error. We propose a slight generalization of the model incorporating the short-dependence component, namely the FARIMA$(2, d, 0)$ model. We have estimated the AR(2) coefficients: ϕ_1 (linear term) and ϕ_2 (quadratic term) via a mean-square error (MSE) minimalization scheme taking into account two statistics: FIRT and RS, see Figure 6.9. The estimated values are: $\phi_1 = 0.03$ and $\phi_2 = 0.03$.

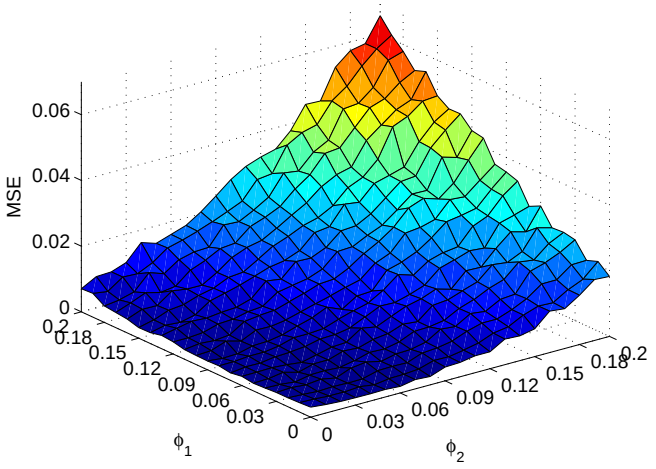

Figure 6.9: Combined mean squared error of the calculated FIRT and RS estimators for the simulated FARIMA$(2, 0.19, 0)$ time series with respect to the ones calculated for the solar flare data for different linear and quadratic coefficients of the AR(2) part. The parameter α is equal to 1.25. One may see the error decreases slightly as ϕ_1 and ϕ_2 grow, it reaches a minimum around the values: $\phi_1 = 0.03$ and $\phi_2 = 0.03$, and then quickly increases when ϕ_1 or ϕ_2 get large.

The FARIMA processes were generated according to the algorithm presented in [195]. The H_{FIRT} and d_{RS} estimators for the simulated FARIMA$(2, 0.19, 0)$ model and the corresponding shuffled data have been obtained following the procedures described in Chapter 5. We have generated 1000 trajectories of size 2^{10}, which is close to the length of the original solar data, i.e. 1089, calculated the estimators for each trajectory, and compared the results with the values in Table 6.3 with the help of box plots [39]. Box plot is a basic tool in descriptive statistics and, moreover, allows to check whether observed values are statistically admissible within a given model. We checked that the results from the simulated model coincide with the values in Table 6.3. Thus, our FARIMA model reconstructs well the behaviour of the original solar flares data.

We conclude that the proper model is FARIMA$(2, d, 0)$ with Pareto noise and the parameters $\phi_1 = 0.03$, $\phi_2 = 0.03$, $d = 0.19$, and $\alpha = 1.25$. Since $d > 1 - 2/\alpha$ [39] the proposed model recovers the long-term dependence property in the investigated solar data in the codifference sense. The Pareto distribution is a proper choice for the present analysis because it gives a description of positive random variables whereas the stable one for $1 < \alpha < 2$ is related to both positive

Table 6.3: Values of the FIRT and RS estimators for the original time series and the shuffled (surrogate) solar flare data.

Data set	H_{FIRT}	d_{RS}
Original time series		
Solar flares	1.1424	0.2408
Surrogate data (single sample shuffling)		
Solar flares	0.8452	0.0507

and negative random variables, but a time series of X-ray flare energy is quite positive by definition.

Finally, we have calculated the 1-day-ahead prediction for the FARIMA time series applying formula (4.9) and setting $h = 1$. The results are depicted in Figure 6.10.

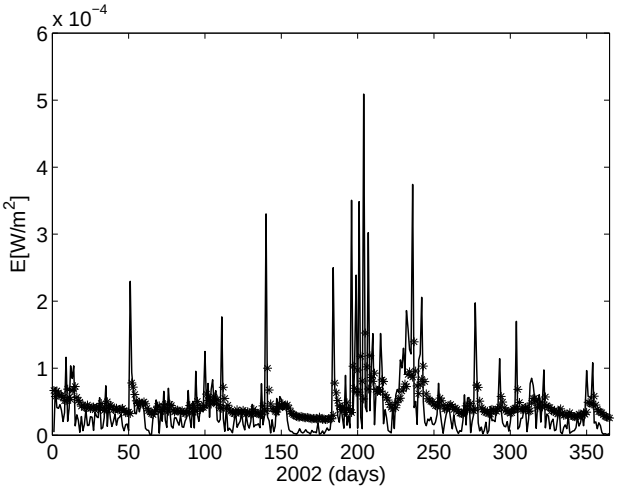

Figure 6.10: Solar flare data (solid line) and 1-day-ahead prediction (asterisks) in the FARIMA$(2, 0.19, 0)$ model. The prediction applies the linear predictor based on the previous observations. The error of the prediction is not low but we have to take into account there is no strong deterministic component in the data, only stochastic one with strongly persistent noise.

The analysis of soft X-ray emission observations shows that this series is enough complicated in nature. It contains both heavy-tailed effects and long-range dependence. The first creates a random number of strong flares on a background,

and the second forms their persistence between each other. The most convenient model for their joint description is the FARIMA time series. The model permits one to predict a time series of soft X-ray solar flares, when the solar activity will be again near its maximum in future. While we do not claim that this model provides the only possible solution, it does provide a rigorous statistical picture of the expected dynamics of X-ray solar flares.

6.3 Modelling of single particle tracking dynamics in molecular biology

As a universal candidate suitable for extensive statistical analysis of the subdiffusive dynamics in biological cells we propose here the FARIMA model which was described in Section 4.3. It is a discrete time analog of FLE [135] that takes into account Gaussian and non-Gaussian α-stable laws and the long-range dependence (long memory). A close relation to FBM and its α-stable extension FSM is explained in detail in Section 4.3.

6.3.1 Fitting FARIMA to the mRNA data

We study here the data of Golding and Cox describing the motion of individual fluorescently labelled mRNA molecules inside live *E. coli* cells [83]. The data clearly follow the subdiffusive character and consist of 27 two-dimensional trajectories [138]. We identified 4 Gaussian and 9 α-stable trajectories among them. Here, we analyse one Gaussian trajectory (no. 12x), the longest trajectory (no. 4y) which follows the non-Gaussian stable law, and a trajectory (no. 2y) which is neither stable with $\alpha \leqslant 2$ nor NIG. Since these trajectories have missing data at some time points, which would influence the FARIMA estimation procedure, we concentrate on their longest continuous parts: $\{(X_n, Y_n) : n = 1, 2, \ldots, N\}$ with $N = 327$, $N = 970$, and $N = 339$, respectively. The data and their increments are presented in Figures 6.11–6.13. First, we estimated FARIMA parameters by applying the estimation procedure given in Section 4.3. The results along with MSD estimates (d_{MSD}) are presented in Table 6.4. We also compared them with the estimates calculated with the standard ITSM package, which was introduced for light-tailed distributions [30]. The ITSM algorithm leads to different results. This is due to the fact that the procedures use different statistics for estimation, but since the MSD estimates (d_{MSD}) are closer to the memory parameter values obtained by the former procedure and most of the data is not Gaussian, we choose the values depicted in Table 6.4 as the FARIMA model parameters.

The goodness of fit of FARIMA processes is illustrated in Figures 6.14–6.16. We plotted sample MSD for three analyzed trajectories along with the 95% con-

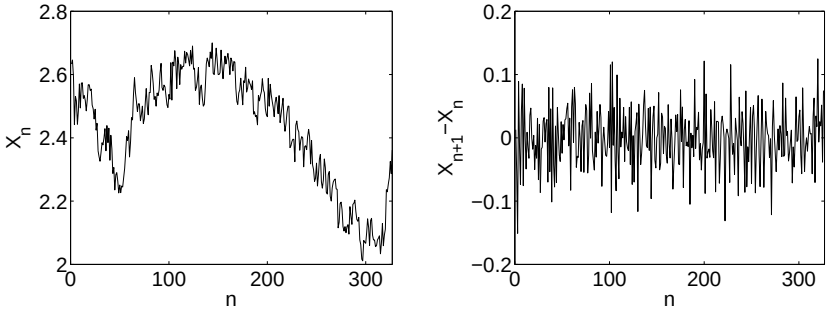

Figure 6.11: A plot of the x-coordinate of particle position during a motion of a tagged RNA molecule inside an *E. coli* cell (left panel) and its increments (right panel). Trajectory no. 12 with Gaussian ($\alpha = 2$) noise.

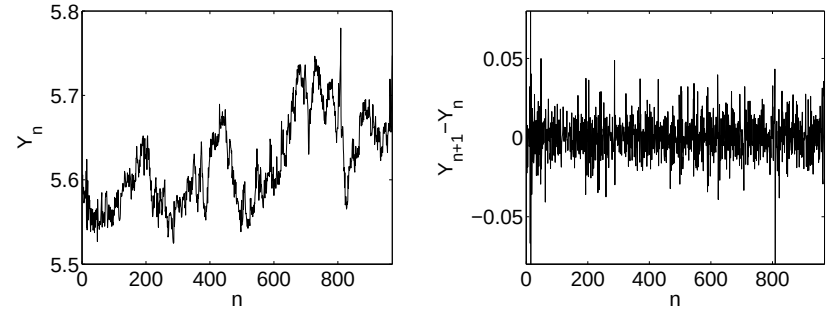

Figure 6.12: A plot of the y-coordinate of particle position during a motion of a tagged RNA molecule inside an *E. coli* cell (left panel) and its increments (right panel). Trajectory no. 4 with stable noise with $\alpha = 1.81$.

fidence intervals calculated under the assumption of the FARIMA model. We can see the empirical trajectiories lie within the bounds except for several values which indicates the model is well-fitted. We also note that the $M_N(\tau)$ statistic used to calculate the last column in Table 6.4 is reliable only for large N/τ, see Section 5.4.3. We set similar minimal values for N/τ for different trajectories.

6.3.2 Fitting FARIMA to a time series from the Golding and Cox microscopy video

First, we describe the method used to extract single mRNA molecule time series from the Golding and Cox microscopy video [83]. The analyzed video has 1801 frames of size 59×76 pixels, which present location of mRNA molecule at the time $t_k = k$ sec, for $k = 1, \ldots, 1801$. The centre of pixel in the upper-left and bottom-

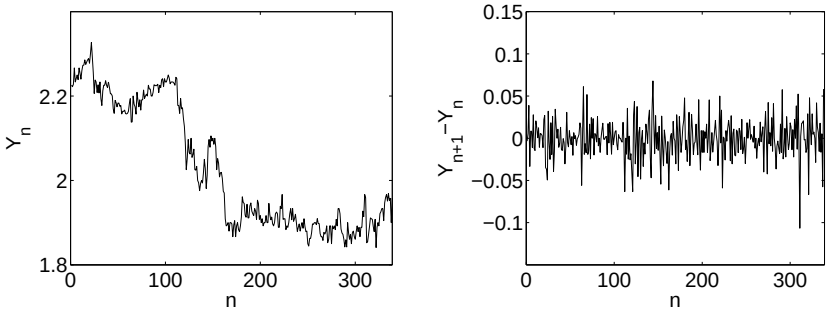

Figure 6.13: A plot of the y-coordinate of particle position during a motion of a tagged RNA molecule inside an *E. coli* cell (left panel) and its increments (right panel). Trajectory no. 2 with unknown noise.

Table 6.4: Results of the proposed estimation procedure for three sample trajectories of the Golding and Cox data.

Traj.	Model	ϕ_1	θ_1	α	d_{MSD}
12x	FARIMA$(0, -0.23, 1)$	0	0.12	2	-0.21
4y	FARIMA$(0, -0.14, 1)$	0	0.16	1.81	-0.17
2y	FARIMA$(1, -0.08, 1)$	0.29	0.48	?	-0.16

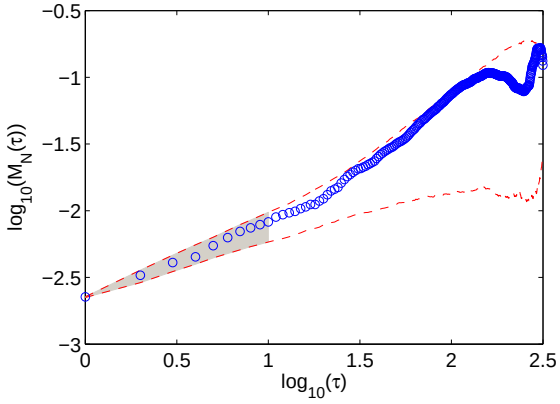

Figure 6.14: Sample MSD (circles) for x-coordinate of the trajectory no. 12 with Gaussian ($\alpha = 2$) noise and estimated 95% confidence intervals obtained via Monte Carlo simulations under the assumption of the Gaussian FARIMA model (dashed lines) in double logarithmic scale. The values of the statistic lie within the confidence interval for FARIMA model supporting its goodness of fit. The memory parameter $d_{MSD} = -0.21$ was obtained by fitting the line to the values in the grey region.

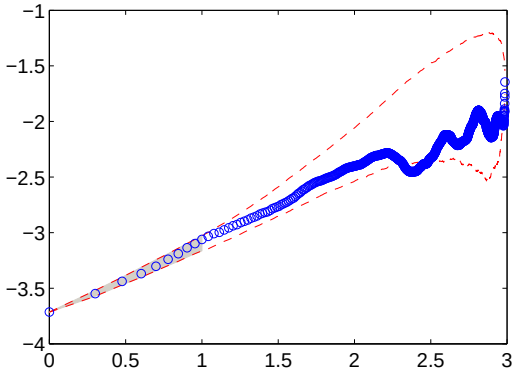

Figure 6.15: Sample MSD (circles) for y-coordinate of the trajectory no. 4 with stable noise with $\alpha = 1.81$ and estimated 95% confidence intervals obtained via Monte Carlo simulations under the assumption of the stable FARIMA model (dashed lines) in double logarithmic scale. The values of the statistic lie within the confidence interval for FARIMA model supporting its goodness of fit. The memory parameter $d_{MSD} = -0.17$ was obtained by fitting the line to the values in the grey region.

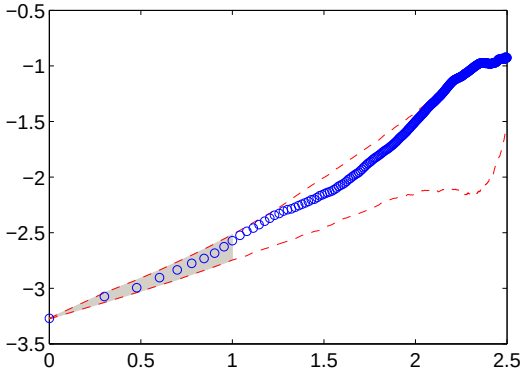

Figure 6.16: Sample MSD (circles) for y-coordinate of the trajectory no. 2 with unknown noise and estimated 95% confidence intervals obtained via Monte Carlo simulations under the assumption of the FARIMA model with noise generated from empirical distribution function (dashed lines) in double logarithmic scale. The values of the statistic lie within the confidence interval for FARIMA model supporting its goodness of fit. The memory parameter $d_{MSD} = -0.16$ was obtained by fitting the line to the values in the grey region.

right corner of each frame has coordinates $(1, 1)$ and $(59, 76)$, respectively. Here we propose to identify the position of a molecule at the time t_k with the position of its mass centre $(\bar{x}(t_k), \bar{y}(t_k))$, which is calculated based on image segmentation results obtained by application of the approach introduced in [55]. For other image processing and filtering methods, see, e.g., [183, 177].

Let Ω be an open and bounded subset of $\mathbb{R}^2$. We denote by $f_k : \Omega \to \mathbb{R}$, an image in the k-th frame of analyzed video. Each value $f_k(x)$ corresponds to grey intensity of this image at the point $x \in \Omega$. In order to find a contour of molecule presented in f_k, we base on the well-known active contour model of Chan and Vese [55]. This model is based on a simple observation that contours of meaningful objects (here a single molecule) in an image f_k must satisfy a minimum of the functional

$$F(C) := \alpha|C| + \lambda_1 \int_{in(C)} (f_k(x) - c_1)^2 dx + \lambda_2 \int_{out(C)} (f_k(x) - c_2)^2 dx,$$

where c_1 and c_2 denote average intensity value of f_k inside and outside of the contour C, respectively. Here $|C|$ is the length of C and α, λ_1, λ_2 are some positive parameters.

In order to be able to find a minimizer of the functional F, we use the level set formulation, in which, the contour C is represented by the zero level set of a function $\phi : \Omega \to \mathbb{R}$, such that

$$\begin{cases} C := \{x \in \Omega : \phi(x) = 0\}, \\ in(C) := \{x \in \Omega : \phi(x) > 0\}, \\ out(C) := \{x \in \Omega : \phi(x) < 0\}. \end{cases}$$

With this definition, we can replace the unknown variable C by ϕ and reformulate the functional F to

$$\begin{aligned} F_\varepsilon(\phi) &:= \alpha \int_\Omega \delta_\varepsilon(\phi(x))|\nabla\phi(x)|\, dx + \lambda_1 \int_\Omega H_\varepsilon(\phi(x))(f_k(x) - c_1(\phi))^2\, dx \\ &+ \lambda_2 \int_\Omega (1 - H_\varepsilon(\phi(x)))(f_k(x) - c_2(\phi))^2\, dx, \end{aligned}$$

where the function $H_\varepsilon : \mathbb{R} \to \mathbb{R}$, defined by

$$H_\varepsilon(z) := \frac{1}{2}\left(1 + \frac{2}{\pi}\arctan\left(\frac{z}{\varepsilon}\right)\right),$$

is a slight regularization (with the parameter $\varepsilon > 0$) of the Heaviside function

$$H(z) := \begin{cases} 1 & \text{if } z \geqslant 0, \\ 0 & \text{if } z < 0, \end{cases}$$

and $\delta_\varepsilon = H'_\varepsilon$. The variables c_1 and c_2 can be calculated with the help of the function H by formulas

$$c_1(\phi) := \frac{\int_\Omega f_k(x) H(\phi(x))\,dx}{\int_\Omega H(\phi(x))\,dx}, \quad c_2(\phi) := \frac{\int_\Omega f_k(x)(1 - H(\phi(x)))\,dx}{\int_\Omega (1 - H(\phi(x)))\,dx}\,.$$

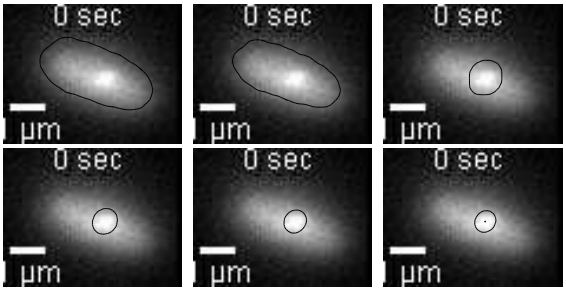

Figure 6.17: Evolution of the active contour in the first frame.

Minimization of the functional F_ε by using standard results of variational calculus, yields to the Euler-Lagrange equation for a function that we denote by ϕ_k. That is, the function ϕ_k satisfies minimum of F_ε with the image f_k given and its zero level set corresponds to an optimal contour of a molecule at time t_k. Parametrization of the descent direction by an artificial time $t \geqslant 0$ gives us the following evolution equation

$$\frac{\partial \phi_k}{\partial t} = \delta_\varepsilon(\phi_k)\left(\alpha \nabla \cdot \left(\frac{\nabla \phi_k}{|\nabla \phi_k|}\right) - \lambda_1 (f_k - c_1(\phi_k))^2 + \lambda_2 (f_k - c_2(\phi_k))^2\right),$$

defined for all $x \in \Omega$ and $t \in (0, T]$, with the associated Neumann boundary condition equal zero and the initial condition $\phi_k(x, 0) = \phi^0(x)$ for all $x \in \Omega$.

Having the solution ϕ_k to the above evolution equation for some $T > 0$, we may calculate position of the mass centre $(\bar{x}(t_k), \bar{y}(t_k))$ of a molecule at the time t_k using the formula

$$(\bar{x}(t_k), \bar{y}(t_k)) := \frac{\int_\Omega x\, H(\phi_k(x))\,dx}{\int_\Omega H(\phi_k(x))\,dx}\,,$$

for $k = 1, \ldots, 1801$.

In Figure 6.17 evolution of the active contour in the image f_0 is presented. As an initial condition for all ϕ_k, with $k = 1, \ldots, 1801$, we have taken the function ϕ^0, which the zero level set is presented in the first image in the upper row of Figure 6.17. In order to improve quality of analyzed images, we have increased

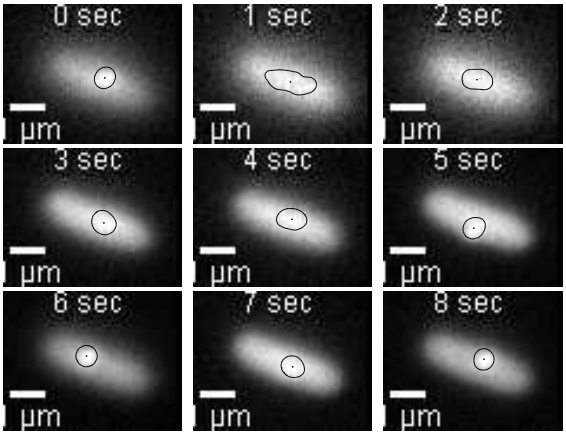

Figure 6.18: Results of mRNA molecule segmentation in nine successive frames and positions of the mass centre $(\bar{x}(t_k), \bar{y}(t_k))$ for $k = 1, \ldots, 9$.

contrast and performed pre-smoothing by convolution with the Gaussian kernel with the standard deviation $\sigma = 2$. In all cases we have used the same set of parameters $\alpha = 1$, $\lambda_1 = 2$, $\lambda_2 = 1$ and $\varepsilon = 1$. The evolution equation for ϕ_k has been solved using the explicit finite difference scheme with 10 iterations and the time step 1. After each iteration, we have performed reinitialization of ϕ_k to the signed distance function to the zero level set of ϕ_k.

In Figure 6.18 we present final results of image segmentation in the first nine frames. The point inside of a contour is the mass centre $(\bar{x}(t_k), \bar{y}(t_k))$ of the detected molecule, for $k = 1, \ldots, 9$.

Finally, we obtain 2D data set of $(\bar{x}(t_k), \bar{y}(t_k)$, for $k = 1, \ldots, 1801$. These data form time series of a single mRNA molecule in the analyzed microscopy video [42]. We analyse a 2D trajectory $\{(X_n, Y_n) := (\bar{x}(t_n), \bar{y}(t_n)) : n = 1, \ldots, 1801\}$ and concentrate on the increments of x and y coordinates. All further analysis is based on two sets of increments $\{X_{n+1} - X_n : n = 1, \ldots, 1800\}$ and $\{Y_{n+1} - Y_n : n = 1, \ldots, 1800\}$ of coordinates x and y, respectively.

Trajectories and their increments of both coordinates of the data are presented in Figure 6.19. A simple examination of the presented increments tells us that the data are not stationary in general. They demonstrate two different regimes. Hence, by using the variance change point test [80] we split the data for two stationary subsets. The changing points are 702th and 715th observation for x and y coordinates, respectively. In sum, we obtain four subsets which we denote by $X1$, $X2$, $Y1$ and $Y2$.

To check the subdiffusive behaviour of the data we calculated the sample MSD, see Section 5.4.3. We calculated sample MSD for all four subsets of the

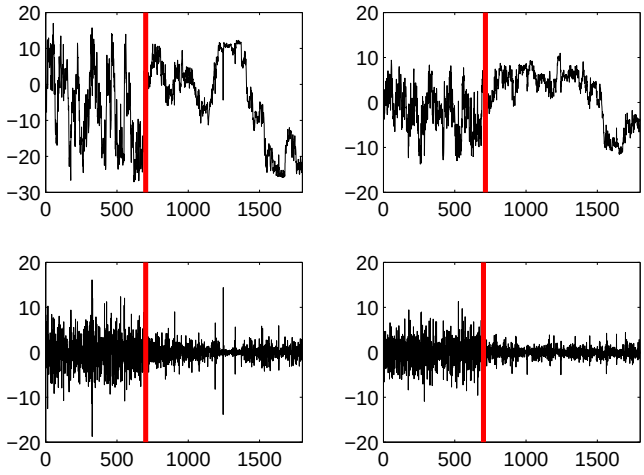

Figure 6.19: A plot of the trajectories (top) and their increments (bottom) with a change point in variance (grey line), for x (left panel) and y coordinates (right panel). The variance changing points are 702th and 715th observations for x and y coordinates, respectively.

Table 6.5: Values of the memory parameter for all four stationary subsets of data and for the 2D counterparts (in Euclidean norm).

$1D$	$X1$	$X2$	$Y1$	$Y2$
d_{MSD}	-0.1	-0.19	-0.28	-0.22
$2D$	$(X1, Y1)$	$(X2, Y2)$	(X, Y)	
d_{MSD}	-0.15	-0.20	-0.16	

data: $X1$, $X2$, $Y1$ and $Y2$ and also for entire coordinates $X = X1 \cup X2$ and $Y = Y1 \cup Y2$. The results are shown in Table 6.5. We can see that in all cases $d < 0$ (the negative dependence case), hence the process clearly follows the subdiffusive dynamics.

Next, we performed various statistical tests to identify the underlying distribution of the increments for all four subsets of data. To this end we employed tests described in Section 5.2. We took into account three well-known possible probability laws, namely Gaussian, NIG and non-Gaussian stable (see Section 3). In Table 6.6 we summarize the results of our statistical procedure. The NIG distribution seems to be the best choice.

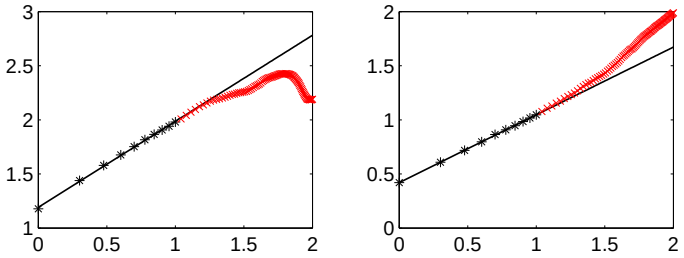

Figure 6.20: Sample MSD for $X1$ (left panel) and $X2$ (right panel) subsets. The line was fitted to the first ten values of sample MSD (black asterisks). The slope $2d+1$ equals 0.8 (left panel) and 0.62 (right panel), respectively.

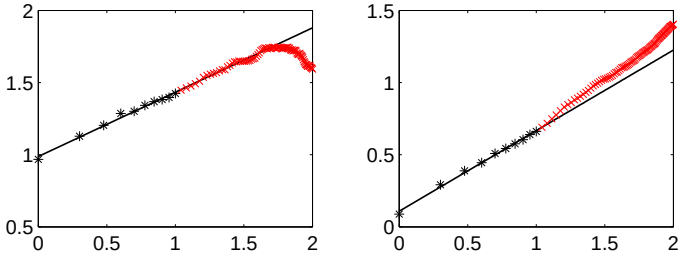

Figure 6.21: Sample MSD for $Y1$ (left panel) and $Y2$ (right panel) subsets. The line was fitted to the first ten values of MSD (black asterisks). The slope $2d+1$ equals 0.44 (left panel) and 0.56 (right panel), respectively.

Statistical tests demonstrate that increments of the data are not Gaussian, but are best fitted with the NIG distribution, see Table 6.6. This implies the FBM can not be considered as the proper model. We fitted a FARIMA model (see Section 4.3) to the four time series [42]. It appears that FARIMA$(1, d, 1)$ is sufficient to describe the data well. The results of the FARIMA procedure for the four studied data sets are presented in Table 6.7.

After fitting the FARIMA model we calculated its residuals, i.e. noise. Their independence could not be rejected for Ljung–Box, turning points, difference-sign and rank tests (for the information about the tests, see, e.g., [30, 35, 33]). We also repeated the same distribution testing procedure as for the increments. The results are presented in Table 6.8. We can see that the NIG law is the recommended choice as it is the only one that cannot be rejected for any part of the data. To check the goodness of fit of the model, we also calculated sample MSD for 1000 simulated trajectories of FARIMA models with parameters given

Table 6.6: Possible distributions for four subsets of the analyzed data. "–" refers to the case where all four statistical tests rejected the distribution, whereas "+" refers to the case where all statistical tests could not reject the distribution.

Distribution	$X1$	$X2$	$Y1$	$Y2$
Gaussian	$-$	$-$	$-$	$-$
Stable	$+$	$-$	$+$	$-$
NIG	$+$	$+$	$+$	$+$

Table 6.7: Results of the estimation procedure for four subsets of the analyzed data.

Subset	Model	ϕ_1	θ_1
X1	FARIMA$(0, -0.13, 0)$	0	0
X2	FARIMA$(0, -0.14, 1)$	0	0.13
Y1	FARIMA$(0, -0.38, 0)$	0	0
Y2	FARIMA$(1, -0.07, 1)$	0.48	0.67

in Table 6.7 and compared the results with the MSD values depicted in Table 6.5. We can see in Figure 6.22 that the fitted FARIMA processes reproduce the sample MSD well.

The proposed FARIMA model is universal, nevertheless, it turns out, that the four stationary subsets of the 2D trajectory have different parameters (ϕ_1, d, θ_1), see Table 6.7. This is due to the fact that our single particle tracking analysis is not a mean-value approach as it was done in [83]. Observe that the shape of the cells and crowded fluid characteristic of the cytoplasm influence the dynamics of the labelled mRNA molecules. In particular, we believe that the parameters of the fitted FARIMA models can provide some insight into the physical reasons for subdiffusive motion of the molecule. Namely, the parameters d in both directions x and y are influenced by the shape of the cell. Simulations show that, e.g., as the width of the biological cell gets smaller then the memory parameter becomes "more negative".

Recall that both the self-similarity index H and type of distribution $1/\alpha$ have impact on d since $d = H - 1/\alpha$. Two additional parameters ϕ_1 and θ_1 which are responsible for short-time effects are influenced by short-distance interactions in a crowded fluid environment in the cytoplasm [42].

Table 6.8: Possible distributions for noise of the fitted FARIMA model for four subsets of the analyzed data. "–" refers to the case where all four statistical tests rejected the distribution, whereas "+" refers to the case where all statistical tests could not reject the distribution.

Distribution	$X1$	$X2$	$Y1$	$Y2$
Gaussian	−	−	+	−
Stable	+	−	+	+
NIG	+	+	+	+

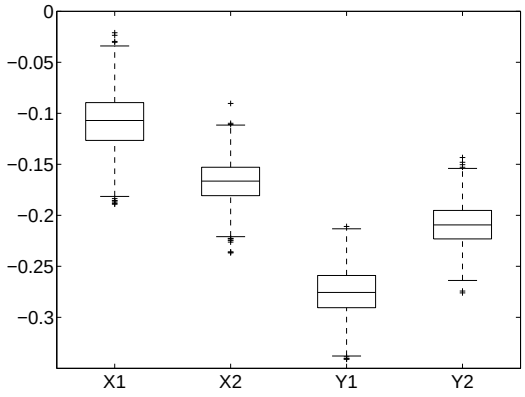

Figure 6.22: Boxplots of sample MSD calculated for 1000 simulated trajectories of the fitted FARIMA processes for X1, X2, Y1 and Y2 subsets of the analyzed data. We can observe that the MSD values for the empirical data: -0.1, -0.19, -0.28 and -0.22 (see Table 6.5) fall into the confidence regions.

Table 6.9: Results of the ITSM estimation procedure for four sample trajectories of the telomere data.

Traj.	Model	ϕ_1	θ_1	α	d_{MSD}
$0707x$	FARIMA$(0, -0.36, 1)$	0	0.30	2	-0.39
$0707y$	FARIMA$(0, -0.36, 1)$	0	0.26	2	-0.37
$47x$	FARIMA$(1, -0.31, 0)$	-0.28	0	2	-0.29
$3y$	FARIMA$(0, -0.36, 0)$	0	0	2	-0.27

6.3.3 Fitting FARIMA to the telomere data

We investigate here the data describing transient anomalous diffusion of telomeres in the nucleus of living human cells (U2OS cancer) presented in [115], where the diffusion properties of telomeres in a broad time range of almost 6 orders of magnitude by combining different imaging setups on the same microscope were examined.

In this context let us mentioned that the 2009 Nobel Prize in Physiology or Medicine was awarded to Elizabeth H. Blackburn, Carol W. Greider and Jack W. Szostak for the discovery of how chromosomes are protected by telomeres and the enzyme telomerase. The chromosomes contain our genome in their DNA molecules. As early as the 1930s, Hermann Muller (Nobel Prize 1946) and Barbara McClintock (Nobel Prize 1983) had observed that the structures at the ends of the chromosomes, the so-called telomeres, seemed to prevent the chromosomes from attaching to each other. They suspected that the telomeres could have a protective role, but how they operate remained an enigma.

We concentrate on the data which clearly follow the subdiffusive character and consist of 180 trajectories of length 1000 observed over time range $10^{-2} - 10$ [s] (CCD) and 2D confocal dataset of 94 trajectories of length 200 measured over $1 - 10^2$ [s] (2D confocal) [115]. All of them follow the Gaussian distribution. We study in detail four different trajectories: (no. 0707x) and (no. 0707y) from the CCD dataset as well as (no. 47y) and (no. 3y) from the 2D confocal dataset. The data and their increments are presented in Figures 6.23–6.26. Since the data are Gaussian we estimated FARIMA parameters with the standard ITSM package, which was proposed for light-tailed distributions [30]. The estimated parameters along with MSD estimates (d_{MSD}) are presented in Table 6.9. We also compared them with the estimates calculated by the estimation procedure given in Section 4.3. It appears that the results are similar. Hence, we choose the values depicted in Table 6.9 as the FARIMA model parameters.

The goodness of fit of FARIMA processes is illustrated in Figures 6.27–6.30. We plotted sample MSD function for four analyzed trajectories along with the

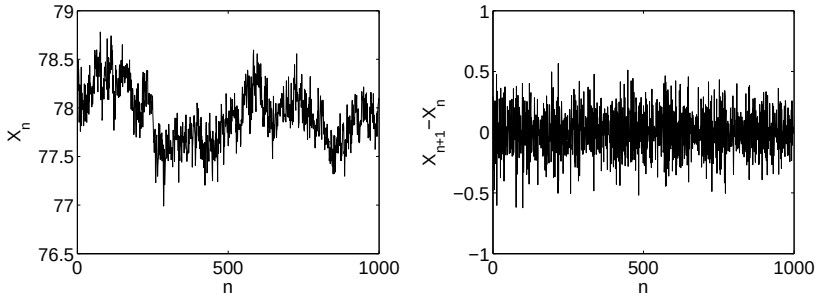

Figure 6.23: A plot of the x-coordinate of telomere data (left panel) and its increments (right panel). Trajectory no. 0707 with Gaussian noise.

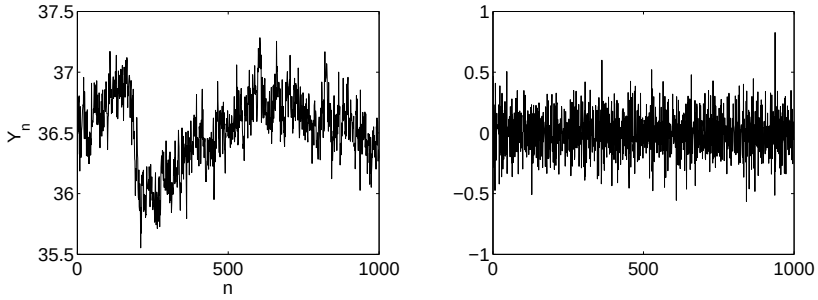

Figure 6.24: A plot of the y-coordinate of telomere data (left panel) and its increments (right panel). Trajectory no. 0707 with Gaussian noise.

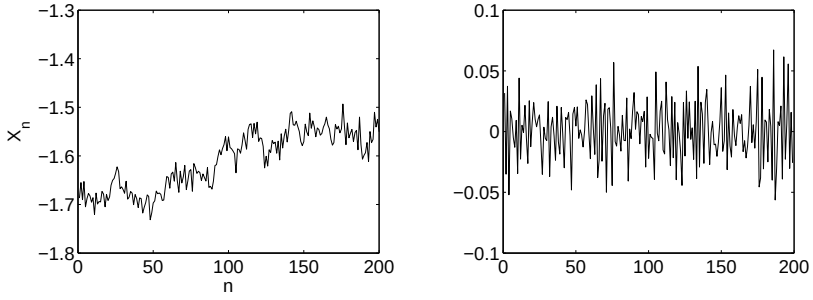

Figure 6.25: A plot of the x-coordinate of telomere data (left panel) and its increments (right panel). Trajectory no. 47 with Gaussian noise.

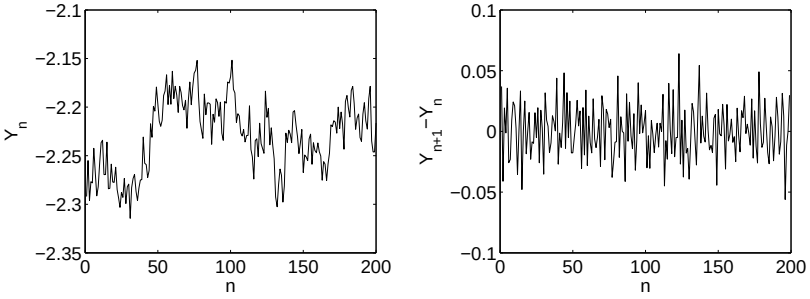

Figure 6.26: A plot of the y-coordinate of telomere data (left panel) and its increments (right panel). Trajectory no. 3 with Gaussian noise.

95% confidence intervals calculated under the assumption of the FARIMA model. We can see the empirical trajectiories lie within the bounds except for several values which indicates the model is well-fitted. We also note that the $M_N(\tau)$ statistic is reliable only for large N/τ [48]. In all cases we set similar minimal values for N/τ.

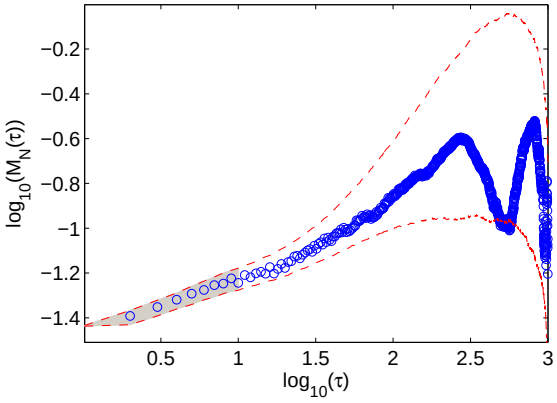

Figure 6.27: Sample MSD (circles) for x-coordinate of the trajectory no. 0707 from telomere data with Gaussian noise and estimated 95% confidence intervals obtained via Monte Carlo simulations under the assumption of the Gaussian FARIMA model (dashed lines) in double logarithmic scale. The values of the statistic lie within the confidence interval for FARIMA model supporting its goodness of fit. The memory parameter $d_{MSD} = -0.39$ was obtained by fitting the line to the values in the grey region.

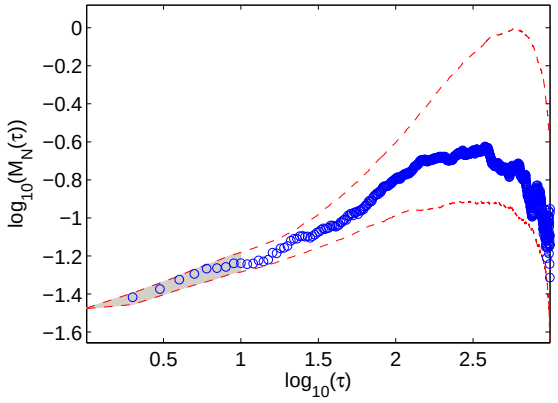

Figure 6.28: Sample MSD (circles) for y-coordinate of the trajectory no. 0707 from telomere data with Gaussian noise and estimated 95% confidence intervals obtained via Monte Carlo simulations under the assumption of the Gaussian FARIMA model (dashed lines) in double logarithmic scale. The values of the statistic lie within the confidence interval for FARIMA model supporting its goodness of fit. The memory parameter $d_{MSD} = -0.37$ was obtained by fitting the line to the values in the grey region.

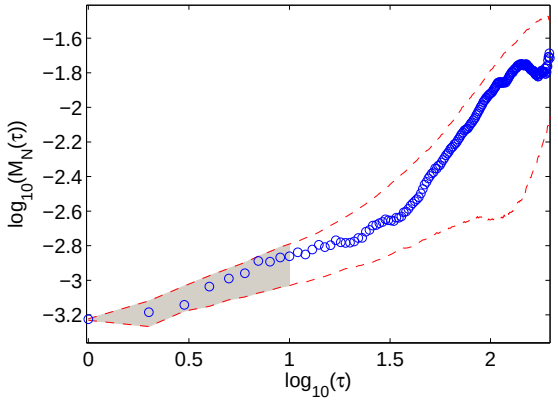

Figure 6.29: Sample MSD (circles) for x-coordinate of the trajectory no. 47 from telomere data with Gaussian noise and estimated 95% confidence intervals obtained via Monte Carlo simulations under the assumption of the Gaussian FARIMA model (dashed lines) in double logarithmic scale. The values of the statistic lie within the confidence interval for FARIMA model supporting its goodness of fit. The memory parameter $d_{MSD} = -0.29$ was obtained by fitting the line to the values in the grey region.

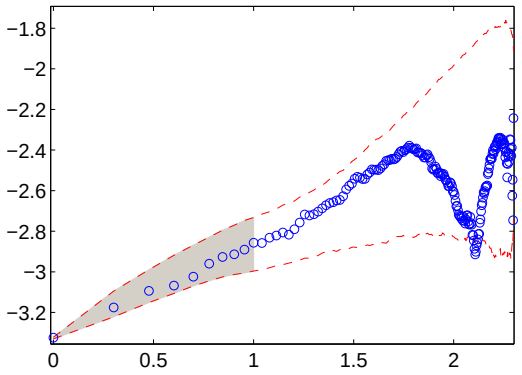

Figure 6.30: Sample MSD (circles) for y-coordinate of the trajectory no. 3 from telomere data with Gaussian noise and estimated 95% confidence intervals obtained via Monte Carlo simulations under the assumption of the Gaussian FARIMA model (dashed lines) in double logarithmic scale. The values of the statistic lie within the confidence interval for FARIMA model supporting its goodness of fit. The memory parameter $d_{MSD} = -0.27$ was obtained by fitting the line to the values in the grey region.

7

Algorithm for identification, validation and prediction. ASP module

In this chapter we present a computer algorithm for identification, validation and prediction, which was implemented in the Analysis, Simulation and Prediction (ASP) module. The module was developed within the Project: "Detectors and sensors for measuring factors hazardous to environment – modeling and monitoring of threats" (for the details of the project see Chapter 2). The algorithm incorporates the tools presented in Chapter 5. The usefulness of the algorithm is illustrated on a case study related to the wind speed data collected by the telecommunication group.

7.1 Description

The ASP module allows to perform analyses both in time and frequency domain. The models implemented in the module are: ARMA-GARCH, FARIMA and hidden Markov. The first model is ARMA (for this special case of the FARIMA process see Section 4.3) with generalized autoregressive conditional heteroskedastic (GARCH) errors introduced by Bollerslev [22]. Bollerslev [22] generalized the autoregressive conditional heteroskedastic model, named ARCH, proposed in the seminal Nobel Prize winning paper by Engle [73]. It takes into account for heteroskedastic effects of the time series process typically observed in form of fat tails, as clustering of volatilities, and the leverage effect. The FARIMA model is treated in detail in Section 4.3. The underlying idea behind the hidden Markov models (HMM) is to describe the observed phenomena by two (or more) separate phases or regimes with different statistical properties. The mathematics behind the HMM was developed by Baum and co-workers [10, 9, 12, 11, 8], for the applications, see, e.g., [175, 16]. The functions of the module are listed below.

ASP module in time domain

1. Stochastic analysis of the data, removing seasonality, checking for long-range dependence.
2. Calibration of ARMA-GARCH and FARIMA models.
3. Goodness of fit testing by backtesting with confidence intervals.
4. Simulation of the fitted time series.
5. Prediction for the fitted time series.

ASP module in frequency domain

1. Calibration of the hidden Markov model.
2. Identification of two regimes.
3. Simulation of the fitted model.

In Figures 7.1 and 7.2 we present schematic algorithms for the flow of processes in the ASP module and FARIMA submodule, respectively.

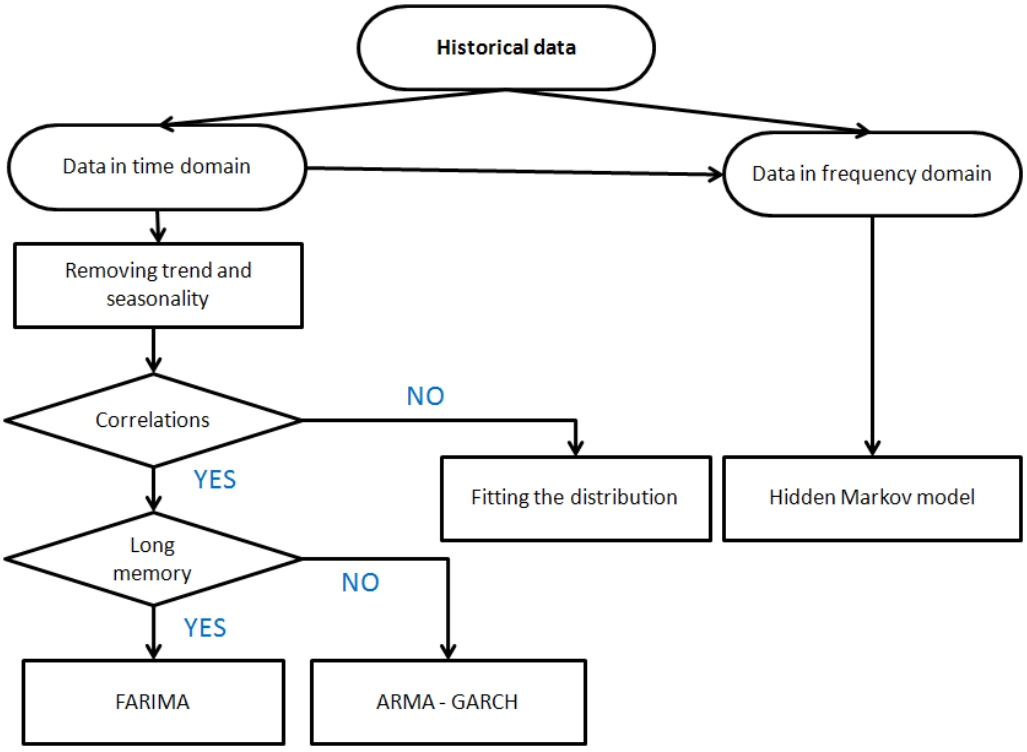

Figure 7.1: Flow chart of the processes in the ASP module.

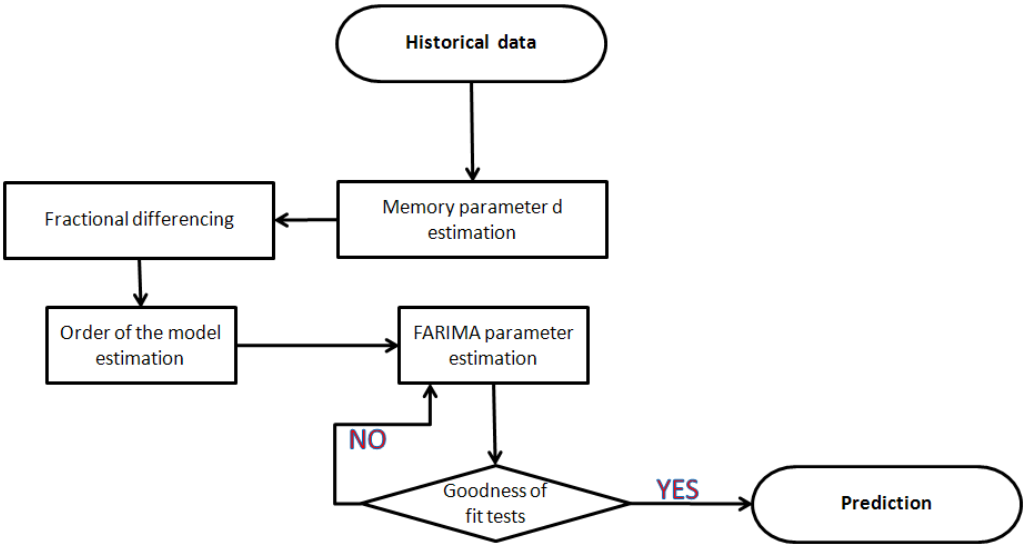

Figure 7.2: Flow chart of the processes in the FARIMA submodule.

7.2 Module contents

The ASP ver. 3.0 consists of 51 functions and scripts. Many functions include internally used routines, hence, the total number of functions is much larger. The files are grouped into four categories:

1. ARMA-GARCH time series,
2. FARIMA time series,
3. Hidden Markov model,
4. GUI functions.

The ASP functions were written in MATLAB 7.11.0 (R2010b), and should therefore work in that or any later version of MATLAB. Note, that some of the results may differ from those presented in the book if a later version of MATLAB is used.

To install the software, you should copy the ASP folder as a folder in your MATLAB "works" folder, and then activate the path to this new folder so that MATLAB knows it is there. All functions will work if you have the Statistics and System Identification Toolboxes in additional to the basic MATLAB package.

7.3 Using the ASP module functions

MATLAB functions or scripts (i.e. series of MATLAB commands) can be run from the MATLAB command line, which is given by the prompt ">>" in the MATLAB Command Window.

Scripts can be executed by simply typing in the relevant file name in the command line. For example, if you wish to run Panel 1 script, you can do so by typing in the following command:

```
>> panel_1
```

Functions, on the other hand, generally require input parameters. Care must be taken to ensure that the input data is in appropriate form. To check data format requirements type ">> help <function name>" at the command line. For instance, to see the help screen of the "sin_tr.m" function type in the following command:

```
>> help sin_tr
```

In return MATLAB will display:

```
SIN_TR estimates sinusoidal trend.
    SIN_TR(DATA,D) fits a sum of sines to the points in vector DATA.
    D is a vector containing period of each component of the sum.
    [PARAM,Y,ERR]=SIN_TR(DATA,D) returns a vector of sines parameters
    PARAM, a vector Y with values of a fitted trend:
    Y(x)=PARAM(1)+PARAM(2)*sin(2pi/D(1)*(x+PARAM(3)))
    + PARAM(4)*sin(2pi/D(2)*(x+PARAM(5)))
    +...+PARAM(2k)*sin(2pi/D(k)(x+PARAM(2k+1)))
    and mean squared error ERR of the fit.
```

7.4 Case Study: Wind speed

We study here data describing wind speed. The data were collected from the pilot network consisting of a meteo station situated on the building D-20 on the main campus of the Wrocław University of Technology. The data describing wind speed in $[m/s]$ were measured on an hourly basis from 20:00 of 2009/12/11 till 8:00 of 2010/02/05. We analyse the data with the ASP module.

First, we load the data using ASP, see Figure 7.3. In Panel 1 we execute the "ACF" command, which plots the autocorrelation function for the considered data, see Figure 7.4. By analyzing the ACF depicted in the lower panel we can observe its quick linear decrease. This suggests a linear trend in the data. We remove it by differencing the time series, cf. Figure 7.5 (lower panel).

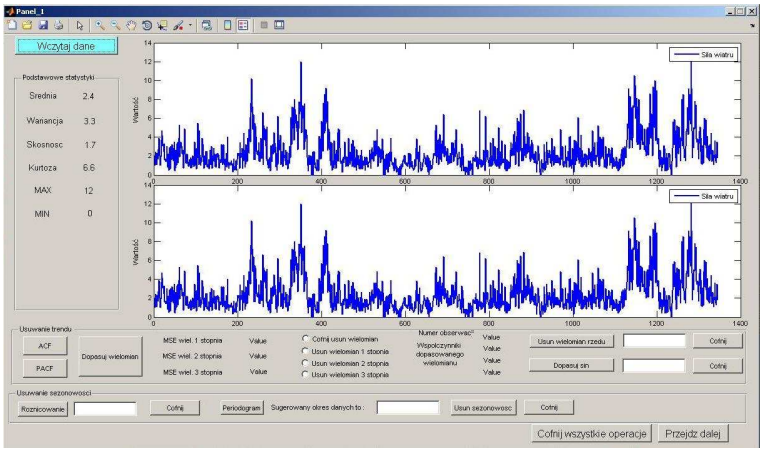

Figure 7.3: Wind speed data in the ASP module.

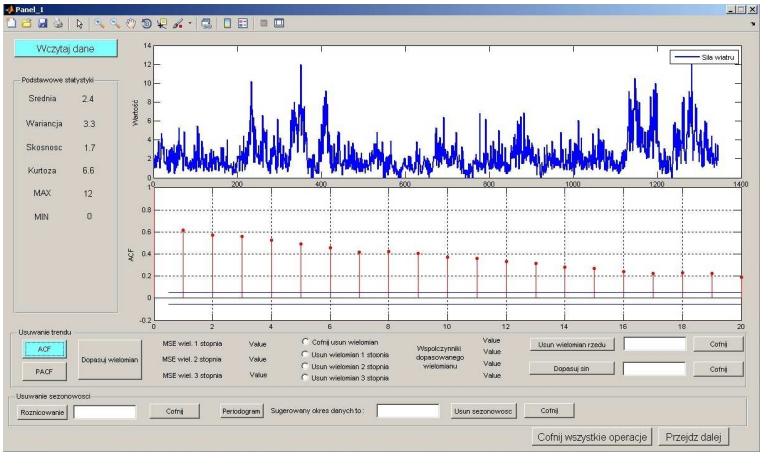

Figure 7.4: Autocorrelation function (lower panel) suggests a linear trend in the data.

Having differenced the data, we have to check whether the data are stationary. This property has to satisfied if we want to proceed to the next step – fitting a time series model. In Figure 7.6 we can observe the ACF for the differenced data. We can see that the values of the function lies within the confidence interval which indicates stationarity of the data.

Next, we choose between two time-series models implemented in ASP, namely ARMA-GARCH and FARIMA processes. The ARMA-GARCH processes are the family of ARMA models [149] with GARCH [22] errors. In contrast to FARIMA processes they do not exhibit power-law long-range dependence. In order to check for long memory we employ sample MSD (see Section 5.4.3). The estimated

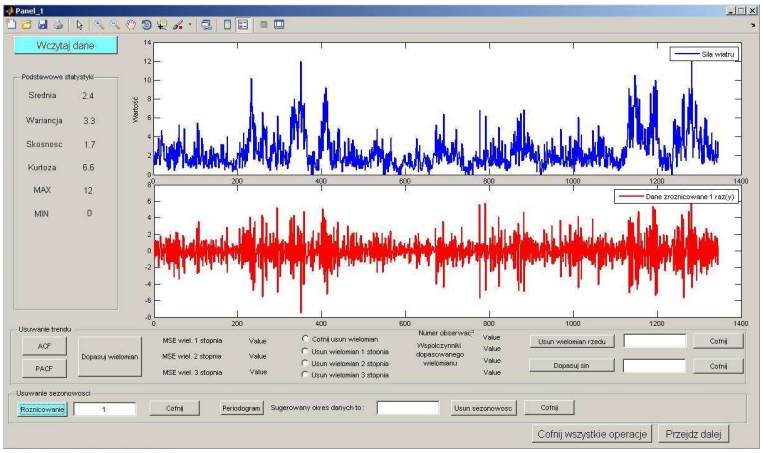

Figure 7.5: First difference of the wind speed data (lower panel).

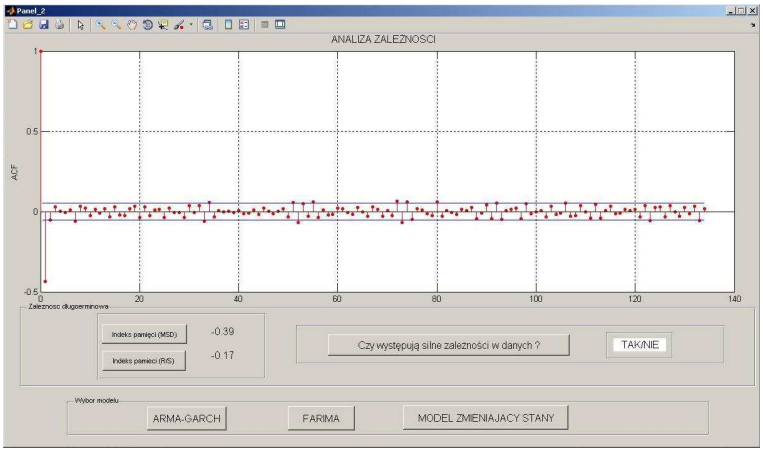

Figure 7.6: Autocorrelation function suggests the differenced data are stationary.

memory parameter equals -0.39 which clearly points to the FARIMA model. In order to estimate the order of the model and other parameters we perform fractional differencing (see Section 4.3) and employ the standard estimation procedure for ARMA time series implemented in Matlab (function "armax"). It is based on the maximum likelihood method. As described in Section 4.3 we can also estimate FARIMA parameters applying the variant of Whittle's estimator. We only have to know the order of the model. It should be equal to the order of the corresponding (via fractional differencing) ARMA process. The results of the estimation procedure are depicted in Figure 7.7.

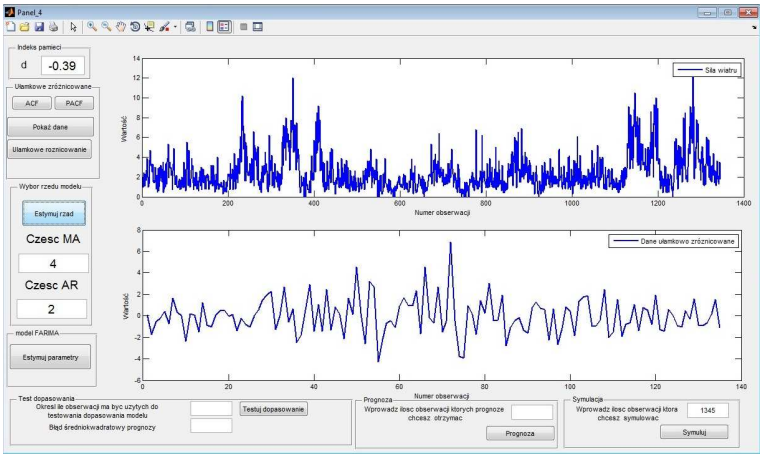

Figure 7.7: The fitted FARIMA$(2, -0.39, 4)$ process.

Finally, we check the goodness of fit by so-called backtesting, namely we remove the last 20 observations, fit the FARIMA model, calculate predictions along with the 90% confidence interval and check whether the removed values lie within the interval. As we can see in Figure 7.8, this condition is fulfiled and we may claim the model is well-fitted. The mean-squared error is equal to 0.93. In Figure 7.9 we also illustrate 20-steps-ahead prediction for the fitted FARIMA$(2, -0.39, 4)$ process.

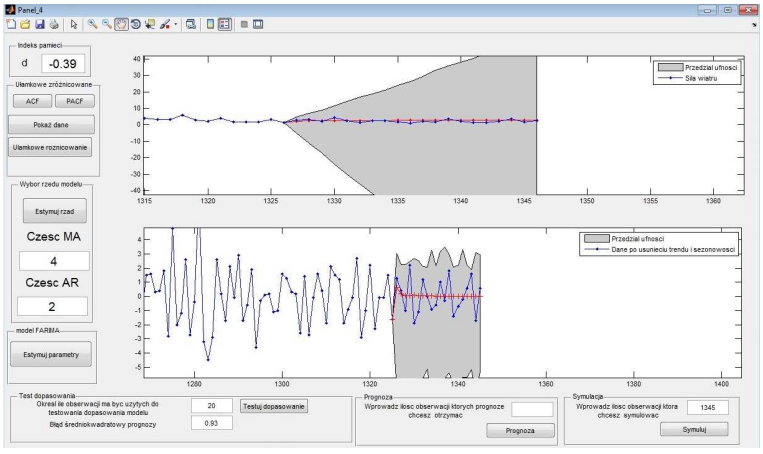

Figure 7.8: Backtesting of the fitted FARIMA$(2, -0.39, 4)$ process on the last 20 observations.

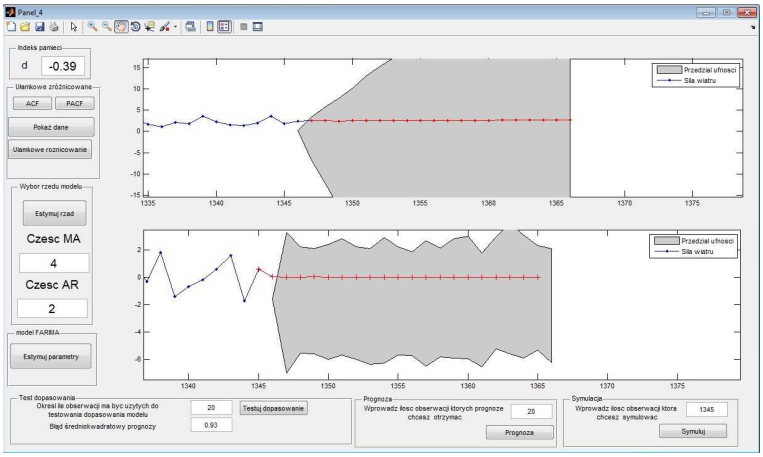

Figure 7.9: 20-steps-ahead prediction for the fitted FARIMA$(2, -0.39, 4)$ process.

Bibliography

[1] Internet page of the Project: "Detectors and sensors for measuring factors hazardous to environment – modeling and monitoring of threats". http://sensory.pwr.wroc.pl/.

[2] E. Alvarez-Lacalle, B. Dorow, J.-P. Eckmann, and E. Moses. Hierarchical structures induce long-range dynamical correlations in written texts. *PNAS*, 103:7956–7961, 2006.

[3] G. R. Arce, editor. *Nonlinear Signal Processing: A Statistical Approach*. Wiley, New York, 2005.

[4] R. T. Baillie. Long memory processes and fractional integration in econometrics. *Journal of Econometrics*, 73:5–59, 1996.

[5] J.-M. Bardet. Testing for the presence of self-similarity of Gaussian time series having stationary increments. *J. Time Ser. Anal.*, 21:497–515, 2000.

[6] O. E. Barndorff-Nielsen. (1995). Normal//Inverse Gaussian Processes and the Modelling of Stock Returns. Research Report 300, Department of Theoretical Statistics, University of Aarhus.

[7] P. Barthelemy, J. Bertolotti, and D. Wiersma. A Lévy flight for light. *Nature*, 453:495–498, 2008.

[8] L. Baum. An inequality and associated maximization technique in statistical estimation of probabilistic functions of a Markov process. *Inequalities*, 3:1–8, 1972.

[9] L. Baum and J. Eagon. An inequality with applications to statistical estimation for probabilistic functions of Markov processes and to a model for ecology. *Bull. Amer. Math. Soc.*, 73:360–363, 1967.

[10] L. Baum and T. Petrie. Statistical inference for probabilistic functions of finite state Markov chains. *Ann. Math. Stat.*, 37:1554–1563, 1966.

[11] L. Baum, T. Petrie, G. Soules, and N. Weiss. A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *Ann. Math. Stat.*, 41:164–171, 1970.

[12] L. Baum and G. Sell. Growth transformations for functions on manifolds. *Pac. J. Math.*, 27:211–227, 1968.

[13] J. Beran. *Statistics for Long-memory Processes*. Chapman & Hall, New York, 1994.

[14] J. Beran, R. Sherman, M. Taqqu, and W. Willinger. Long-range dependence in variable-bit-rate video traffic. *IEEE Trans. Commun.*, 43:1566–1579, 1995.

[15] M. Bertacca, F. Berizzi, and E. Mese. A FARIMA-based technique for oil slick and low-wind areas discrimination in sea sar imagery. *IEEE Trans. Geosci. Remote Sens.*, 43:2484–2293, 2002.

[16] P. Bieńkowski, K. Burnecki, J. Janczura, R. Weron, and B. Zubrzak. A new method for automated noise cancellation in electromagnetic field measurement. *JEMWA*, 26:1226–1236, 2012.

[17] P. Bieńkowski. Electromagnetic fields measurements – methods and accuracy estimation. In *Studies in Applied Electromagnetics and Mechanics*, pages 229–237, Amsterdam, 2008. IOS Press.

[18] P. Bieńkowski. Monitoring of electromagnetic field in the environment – concept of the sensoric network. *Przegląd Elektrotechniczny*, 85:37–39, 2009.

[19] P. Bieńkowski, H. Trzaska, and B. Zubrzak. Methods of the electromagnetic fields intensity level assessment and monitoring in the environment. In I. Dufek et al., editor, *Radioelektronika 2011*, pages 259–262, 2011.

[20] G. D. Birkhoff. Proof of the ergodic theorem. *Proc. Natl. Acad. Sci. USA*, 17:656–660, 1931.

[21] G. Boffetta, V. Carbone, P. Giuliani, P. Veltri, and A. Vulpiani. Power laws in solar flares: Self-organized criticality or turbulence? *Phys. Rev. Lett.*, 83:4662–4665, 1999.

[22] T. Bollerslev. Generalised autoregressive conditional heteroskedasticity. *J. Econometrics*, 31:307–327, 1986.

[23] L. Boltzmann. *Theoretical Physics and Philosophical Problems*. Dover, New York, 1949.

[24] S. Borak, A. Misiorek, and R. Weron. Models for heavy-tailed asset returns. In P. Čižek, W. Härdle, and R. Weron, editors, *Statistical Tools for Finance and Insurance. 2nd Ed.*, pages 21–55. Springer, Berlin, 2011.

[25] S. Borak and R. Weron. Stable distributions. In P. Čižek, W. Härdle, and R. Weron, editors, *Statistical Tools for Finance and Insurance. 2nd Ed.*, pages 21–44. Springer, Berlin, 2005.

[26] G. Box and M. Muller. A note on the generation of random normal deviates. *Ann. Math. Statist.*, 29:610–611, 1958.

[27] R. F. Brcich, D. R. Iskander, and A. M. Zoubir. The stability test for symmetric alpha-stable distributions. *IEEE Trans. Signal. Process.*, 53:977–986, 2005.

[28] D. Brockmann, L. Hufnagel, and T. Geisel. The scaling laws of human travel. *Nature*, 439:462–465, 2006.

[29] P. Brockwell and R. Davis. *Time Series: Theory and Methods*. Springer-Verlag, New York, 1991.

[30] P. Brockwell and R. Davis. *Introduction to Time Series and Forecasting*. Springer-Verlag, New York, 2002.

[31] I. Bronstein, Y. Israel, E. Kepten, S. Mai, Y. Shav-Tal, E. Barkai, and Y. Garini. Transient anomalous diffusion of telomeres in the nucleus of mammalian cells. *Phys. Rev. Lett.*, 103:018102, 2009.

[32] K. Burnecki. Self-similar processes as weak limits of a risk reserve process. *Probab. Math. Statist.*, 20:261–272, 2000.

[33] K. Burnecki. FARIMA processes with application to biophysical data. *J. Stat. Mech.*, 2012. P05015.

[34] K. Burnecki, J. Gajda, and G. Sikora. Modelowanie danych o natężeniu pola elektromagnetycznego szeregami FARIMA. In W. Grzebyk, editor, *Czujniki i sensory do pomiarów czynników stanowiących zagrożenia w środowisku: monografia projektu POIG.01.03.01-02-002/08*, pages 375–386. Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław, 2011.

[35] K. Burnecki, J. Gajda, and G. Sikora. Stability and lack of memory of the returns of the Hang Seng index. *Phys. A*, 390:3136–3146, 2011.

[36] K. Burnecki, J. Janczura, M. Magdziarz, and A. Weron. Can one see a competition between subdiffusion and Lévy flights? : a case of geometric-stable noise. *Acta Phys. Polon. B*, 39:1043–1053, 2008.

[37] K. Burnecki, J. Janczura, and R. Weron. Building loss models. In P. Čižek, W. Härdle, and R. Weron, editors, *Statistical Tools for Finance and Insurance. 2nd Ed.*, pages 293–328. Springer, Berlin, 2011.

[38] K. Burnecki, E. Kepten, J. Janczura, I. Bronshtein, Y. Garini, and A. Weron. Universal algorithm for identification of fractional Brownian motion. A case of telomere subdiffusion. *Biophys. J.*, 103:1839–1847, 2012.

[39] K. Burnecki, J. Klafter, M. Magdziarz, and A. Weron. From solar flare time series to fractional dynamics. *Phys. A*, 387:1077–1087, 2008.

[40] K. Burnecki, G. Kukla, and R. Weron. Property insurance loss distributions. *Phys. A*, 287:269–278, 2000.

[41] K. Burnecki, M. Magdziarz, and A. Weron. Identification and validation of fractional subdiffusion dynamics. In J. Klafter, S. Lim, and R. Metzler, editors, *Fractional Dynamics*, pages 331–352. World Scientific, Singapore, 2012.

[42] K. Burnecki, M. Muszkieta, G. Sikora, and A. Weron. Statistical modelling of subdiffusive dynamics in the cytoplasm of living cells: A FARIMA approach. *EPL*, 98:10004, 2012.

[43] K. Burnecki and G. Sikora. Estimation of FARIMA parameters in the case of negative memory and stable noise, 2012. Submitted to IEEE Trans. Signal Process.

[44] K. Burnecki and G. Sikora. FARIMA time series in modeling of UMTS data, 2012. In preparation.

[45] K. Burnecki, G. Sikora, and A. Weron. Fractional process as a unified model for subdiffusive dynamics in experimental data. *Phys. Rev. E*, 86:041912, 2012.

[46] K. Burnecki, A. Stanislavsky, and K. Weron. Statistical analysis of the maximum energy in solar X-ray flare activity. *Acta Phys. Polon. B*, 40:1303–1313, 2009.

[47] K. Burnecki and A. Weron. Lévy stable processes. From stationary to self-similar dynamics and back. *Acta Phys. Polon. B*, 35:1–16, 2004.

[48] K. Burnecki and A. Weron. Fractional Lévy stable motion can model subdiffusive dynamics. *Phys. Rev. E*, 82:021130, 2010.

[49] K. Burnecki, A. Wyłomańska, A. Beletskii, V. Gonchar, and A. Chechkin. Recognition of stable distribution with Lévy index α close to 2. *Phys. Rev. E*, 85:056711, 2012.

[50] D. Caccia, D. Percival, M. Cannon, G. Raymond, and J. Bassingthwaighte. Analyzing exact fractal time series: evaluating dispersional analysis and rescaled range methods. *Phys. A*, 246:609–632, 1997.

[51] S. Cambanis, C. Hardin Jr., and A. Weron. Ergodic properties of stationary stable processes. *Stochastic Process. Appl.*, 24:1–18, 1987.

[52] C. Canbay. The essential environmental cause of multiple sclerosis disease. *PIER*, 101:375–391, 2010.

[53] J. Chambers, C. Mallows, and B. Stuck. A method for simulating stable random variables. *J. Amer. Statist. Assoc.*, 71:340–344, 1976.

[54] T. Chan and J. Shen. *Image Processing. Variational, PDE, Wavelet and Stochastic Methods*. SIAM, Philadelphia, 2005.

[55] T. Chan and L. Vese. Active contours without edges. *IEEE Trans. Image Process.*, 10:266–277, 2001.

[56] T. Chang, T. Sauer, and S. Schiff. Tests for nonlinearlity in short stationary time series. *Chaos*, 5:118–126, 1995.

[57] A. Chechkin, V. Gonchar, and M. Szydlowsky. Fractional kinetics for relaxation and superdiffusion in a magnetic field. *Phys. Plasmas*, 9:78–88, 1987.

[58] A. Chernobai, K. Burnecki, S. Rachev, S. Trück, and R. Weron. Modelling catastrophe claims with left-truncated severity distributions. *Comput. Statist.*, 21:537–555, 2006.

[59] A. Choudhuri, P. Chatterjee, and J. Jiang. Predicting solar cycle 24 with a solar dynamo model. *Phys. Rev. Lett.*, 98:131103, 2007.

[60] S. Clark. The dark side of the Sun. *Nature*, 441:402–404, 2006.

[61] M. Coulon, M. Chabert, and A. Swami. Detection of multiple changes in fractional integrated ARMA processes. *IEEE Trans. Signal Process.*, 57:48–61, 2009.

[62] N. Crato and P. Rothman. Fractional integration analysis of long-run behaviour for US macroeconomic time series. *Econom. Lett.*, 45:287–291, 1994.

[63] J. D. Cryer and K.-S. Chan. *Time Series Analysis: With Applications in R, Second Ed.* Springer, New York, 2008.

[64] R. B. D'Agostino and M. A. Stephens. *Goodness-of-fit Techniques*, volume 68 of *Statistics: Textbooks and Monographs*. Dekker Inc., New York, 1986.

[65] R. Davies and D. Harte. Test for Hurst effect. *Biometrika*, 1:95–101, 1974.

[66] M. Denker and W. Woyczyński. *Introductory Statistics and Random Phenomena: Uncertainty, Complexity and Chaotic Behavior in Engineering and Science*. Birkhäuser, Boston, 1998.

[67] M. Deriche and A. Tewfik. Signal modeling with filtered discrete fractional noise processes. *IEEE Trans. Signal Process.*, 41:2839–2849, 1993.

[68] M. Dikpati and P. Gilman. Simulating and predicting solar cycles using a flux-transport dynamo. *Astrophys. J.*, 649:498–514, 2006.

[69] P. Ditlevsen. Anomalous jumping in a double-well potential. *Phys. Rev. E*, 60:172, 1999.

[70] P. Doukham, G. Oppenheim, and M. Taqqu, editors. *Theory and Applications of Long Range Dependence*. Birkhäuser, Boston, 2003.

[71] W. DuMouchel. On the asymptotic normality of the maximum–likelihood estimate when sampling from a stable distribution. *Ann. Statist.*, 1:948–957, 1973.

[72] P. Embrechts and M. Maejima. *Selfsimilar Processes*. Princeton Series in Applied Mathematics. Princeton Univ. Press, Princeton, NJ, 2002.

[73] R. Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50:987–1007, 1986.

[74] A. Erramilli, O. Narayan, and W. Willinger. Experimental queueing analysis with long-range dependent packet traffic. *IEEE/ACM Trans. Networing*, 4:209–223, 1996.

[75] W. Feller. *An Introduction to Probability Theory and its Applications*, volume 2. Wiley, New York, 1971.

[76] P. Flandrin. Wavelet analysis and synthesis of fractional Brownian motion. *IEEE Trans. Inform. Theory*, 38:910–917, 1992.

[77] G. Fouskitakis and S. Fassois. Pseudolinear estimation of fractionally integrated ARMA (ARFIMA) models with automotive application. *IEEE Trans. Signal Process.*, 47:3365–3380, 1999.

[78] R. Fox and M. S. Taqqu. Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *Ann. Statist.*, 14:517–532, 1986.

[79] M. Frigo and S. Johnson. A modified split-radix FFT with fewer arithmetic operations. *IEEE Trans. Signal Process.*, 55:111–119, 2007.

[80] J. Gajda, G. Sikora, and A. Wyłomańska. Variance change point test, 2012. Preprint.

[81] T. Gałkowski and L. Rutkowski. Nonparametric recovery of multivariate functions with applications to system identification. *Proc. IEEE*, 73:942–943, 1985.

[82] L. Gill-Alana. A fractionally integrated model for the Spanish real GDP. *Economics Bulletin*, 3:1–6, 2004.

[83] I. Golding and E. C. Cox. Physical nature of bacterial cytoplasm. *Phys. Rev. Lett.*, 96:098102, 2006.

[84] C. W. J. Granger and R. Joyeux. An introduction to long-memory time series models and fractional differencing. *J. Time Ser. Anal.*, 1:15–29, 1980.

[85] W. Greblicki, A. Krzyżak, and M. Pawlak. Distribution-free pointwise consistency of kernel regression estimate. *Ann. Statist.*, 12:1570–1575, 1985.

[86] W. Greblicki and M. Pawlak. Dynamic system identification with order statistics. *IEEE Trans. Inform. Theory*, 40:1474–1489, 1994.

[87] W. Greblicki and M. Pawlak. *Nonparametric System Identification*. Cambridge University Press, 2008.

[88] G. Guigas, C. Kalla, and M. Weiss. Probing the nanoscale viscoelasticity of intracellular fluids in living cells. *Biophys. J.*, 93:316–323, 2007.

[89] L. Györfi, M. Kohler, K. A., and H. Walk. *A Distribution-Free Theory of Nonparametric Regression*. Springer, New York, 2002.

[90] E. J. Hannan. The asymptotic theory of linear time-series models. *J. Appl. Probability*, 10:130–145, corrections, ibid. 10 (1973), 913, 1973.

[91] W. Härdle. *Applied Nonparametric Regression*. Cambridge University Press, Cambridge, 1990.

[92] W. Härdle, G. Kerkyacharian, D. Picard, and A. Tsybakov. *Wavelets, Approximation, and Statistical Applications*. Springer, New York, 1998.

[93] Z. Hasiewicz. Hammerstein system identification by the Haar multiresolution approximation. *Int. J. Adapt. Control Signal Process.*, 13:691–717, 1999.

[94] Z. Hasiewicz. Modular neural networks for non-linearity recovering by the Haar approximation. *Neural Netw.*, 13:1107–1133, 2000.

[95] Z. Hasiewicz. Non-parametric estimation of non-linearity in a cascade time-series system by multiscale approximation. *Signal Process.*, 81:791–807, 2001.

[96] Z. Hasiewicz and G. Mzyk. Combined parametric-nonparametric identification of Hammerstein systems. *IEEE Trans. Automat. Control*, 49:1370–1375, 2004.

[97] Z. Hasiewicz, M. Pawlak, and P. Śliwiński. Nonparametric identification of nonlinearities in block-oriented systems by orthogonal wavelets with compact support. *IEEE Trans. Circuits Systems I Fund. Theory Appl.*, 52:427–442, 2005.

[98] Z. Hasiewicz and P. Śliwiński. Identification of non-linear characteristics of a class of block-oriented non-linear systems via Daubechies wavelet-based models. *Int. J. Systems Science*, 33:1121–1144, 2002.

[99] M. Hellmann, J. Klafter, D. Heermann, and M. Weiss. Challenges in determining anomalous diffusion in crowded fluids. *J. Phys.: Condens. Matter*, 23:234113, 2011.

[100] J. Hoffman-Jørgensen. Stable densities. *Theory Probab. Appl.*, 38:350–355, 1993.

[101] J. R. M. Hosking. Fractional differencing. *Biometrika*, 68:165–176, 1981.

[102] H. E. Hurst. Long-term storage capacity of reservoirs. *Trans. Am. Soc. Civ. Eng.*, 116:770–808, 1951.

[103] J. Ilow and H. Leung. Self-similar texture modeling using FARIMA processes with applications to satellite images. *IEEE Trans. Signal Process.*, 10:792–797, 2001.

[104] R. Jana and S. Dey. Change detection in teletraffic models. *IEEE Trans. Signal Process.*, 48:846–853, 2000.

[105] A. Janicki and A. Weron. *Simulation and Chaotic Behavior of α-stable Stochastic Processes*. Dekker Inc., New York, 1994.

[106] C. Jarque and A. Bera. A test for normality of observations and regression residuals. *Internat. Statist. Rev.*, 55:163–172, 1987.

[107] J.-H. Jeon and R. Metzler. Fractional Brownian motion and motion governed by the fractional Langevin equation in confined geometries. *Phys. Rev. E*, 81:021103, 2010.

[108] J.-H. Jeon, V. Tejedor, S. Burov, E. Barkai, C. Selhuber-Unkel, K. Berg-Sørensen, L. Oddershede, and R. Metzler. In vivo anomalous diffusion and weak ergodicity breaking of lipid granules. *Phys. Rev. Lett.*, 106:048103, 2011.

[109] R. Jha, P. Kaw, D. Kulkarni, and J. Parikh. Evidence of Lévy stable process in tokamak edge turbulence. *Phys. Plasmas*, 10:699, 2003.

[110] W. C. Kallenberg, T. Ledwina, and E. Rafajłowicz. Testing bivariate independence and normality. *Sankhya*, 59:42–59, 1997.

[111] M. Kanter. Stable densities under change of scale and total variation inequalities. *Ann. Probab.*, 3:697–707, 1975.

[112] T. Karagiannis, M. Molle, and M. Faloutsos. Long-range dependence ten years of internet traffic modeling. *IEEE Internet Computing*, 8:57–64, 2004.

[113] D. Karlis. An EM type algorithm for maximum likelihood estimation for the Normal Inverse Gaussian distribution. *Statist. Probab. Lett.*, 57:43–52, 2002.

[114] D. Karmeshu and A. Krishnamachari. Sequence variability and long-range dependence in DNA: An information theoretic perspective. In N. Pal et al., editors, *Neural Information Processing*, pages 1354–1361. Springer, Berlin, 2004.

[115] E. Kepten, I. Bronshtein, and Y. Garini. Ergodicity convergence test suggests telomere motion obeys fractional dynamics. *Phys. Rev. E*, 83:041919, 2011.

[116] J. Klafter, S. Lim, and R. Metzler, editors. *Identification and validation of fractional subdiffusion dynamics*. World Scientific, Singapore, 2012.

[117] S. Kogon and D. Manolakis. Signal modeling with self-similar α-stable processes: The fractional Lévy stable motion model. *IEEE Trans. Signal Process.*, 44:1006–1010, 1996.

[118] P. S. Kokoszka. Prediction of infinite variance fractional ARIMA. *Probab. Math. Statist.*, 16:65–83, 1996.

[119] P. S. Kokoszka and M. S. Taqqu. Fractional ARIMA with stable innovations. *Stochastic Process. Appl.*, 60:19–47, 1995.

[120] P. S. Kokoszka and M. S. Taqqu. Parameter estimation for infinite variance fractional ARIMA. *Ann. Statist.*, 24:1880–1913, 1996.

[121] A. Kolmogorov. Wienersche Spiralen und einige andere interessante Kurven in Hilbertschen Raum. *C.R. (Doklady) Acad. Sci. URSS (N.S.)*, 26:115–118, 1940.

[122] I. A. Koutrouvelis. Regression-type estimation of the parameters of stable laws. *J. Amer. Statist. Assoc.*, 75:918–928, 1980.

[123] A. Krzyżak. Identification of discrete Hammerstein systems by the Fourier series regression estimate. *Internat. J. Systems Sci.*, 20:1729–1744, 1989.

[124] A. Krzyżak. On estimation of a class of nonlinear systems by the kernel regression estimate. *IEEE Trans. Inform. Theory*, 36:141–152, 1990.

[125] K. Kyungduk. Bayesian wavelet-based methods for the detection of multiple changes of the long memory parameter. *IEEE Trans. Signal Process.*, 54:4461–4470, 2006.

[126] J. Lamperti. Semi-stable stochastic processes. *Trans. Amer. Math. Soc.*, 104:62–78, 1962.

[127] A. Lasota and M. Mackey. *Chaos, Fractals and Noise. Stochastic Aspects of Dynamics*. Springer-Verlag, New York, 1994.

[128] F. Lepreti, C. Carbone, and P. Veltri. Solar flare waiting time distribution: Varying-rate Poisson or Lévy function? *Astrophys. J. Lett.*, 555:L133, 2001.

[129] K.-J. Li, B. Schmieder, and Q.-S. Li. Statistical analysis of the X-ray flares ($m >= 1$) during the maximum period of solar cycle 22. *Astronom. Astrophys. Suppl.*, 131:99–104, 1998.

[130] J. Liu, Y. Shu, L. Zhang, and F. Xue. Traffic modeling based on FARIMA models. In *Electrical and Computer Engineering, 1999 IEEE Canadian Conference on*, volume 1, pages 162–167, 1999.

[131] G. M. Ljung and G. E. P. Box. On a measure of lack of fit in time series models. *Biometrika*, 65:297–303, 1978.

[132] A. W. Lo. Fat tails, long memory, and the stock market since the 1960s. *Economic Notes*, 26:219–252, 2001.

[133] A. W. Lou. Long-term memory in stock market prices. *Econometrica*, 59:1279–1313, 1991.

[134] M. Maejima. Self-similar processes and limit theorems. *Sugaku Expositions*, 2:102–123, 1989.

[135] M. Magdziarz and A. Weron. Fractional Langevin equation with alpha-stable noise. A link to fractional ARIMA time series. *Studia Math.*, 181:47–69, 2007.

[136] M. Magdziarz and A. Weron. Anomalous diffusion: Testing ergodicity breaking in experimental data. *Phys. Rev. E*, 84:051138, 2011.

[137] M. Magdziarz and A. Weron. Ergodic properties of anomalous diffusion processes. *Ann. Phys.*, 326:2431–2443, 2011.

[138] M. Magdziarz, A. Weron, K. Burnecki, and J. Klafter. Fractional Brownian motion versus the continuous-time random walk: A simple test for subdiffusive dynamics. *Phys Rev. Lett.*, 103:180602, 2009.

[139] B. B. Mandelbrot. *The Fractal Geometry of Nature.* W. H. Freeman and Co., San Francisco, Calif., 1982.

[140] B. B. Mandelbrot and J. W. Van Ness. Fractional Brownian motions, fractional noises and applications. *SIAM Rev.*, 10:422–437, 1968.

[141] B. B. Mandelbrot and J. R. Wallis. Noah, Joseph and operational hydrology. *Water Resour. Res.*, 4:909–918, 1968.

[142] M. Mann, R. Bradley, and M. Hughes. Global-scale temperature patterns and climate forcing over the past six centuries. *Nature*, 392:779–787, 1998.

[143] R. N. Mantegna and H. E. Stanley. *An Introduction to Econophysics.* Cambridge Univ. Press, Cambridge, 2000.

[144] G. Marsaglia and T. Bray. A convenient method for generating normal variables. *SIAM Rev.*, 29:260–264, 1964.

[145] G. Marsaglia and W. Tsang. The ziggurat method for generating random variables. *J. Stat. Soft.*, 5, 2000.

[146] J. McCarthy, R. DiSario, and H. Saraoglu. A recursive algorithm for fractionally differencing long data series. *Journal of Modern Applied Statistical Methods*, 2:272–278, 2003.

[147] J. McCulloch. Simple consistent estimators of stable distribution. *Commun. Statist. — Simul. Comput.*, 15:1109–1136, 1986.

[148] J. McCulloch. Financial applications of stable distributions. In *Statistical Methods in Finance*, volume 14, pages 393–425. North-Holland, Amsterdam, 1996.

[149] J. Mende. Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications. *Proc. IEEE*, 79:278–305, 1991.

[150] S. Mercik, K. Weron, K. Burnecki, and A. Weron. Enigma of self-similarity of fractional Lévy stable motions. *Acta Phys. Polon. B*, 34:3773–3793, 2003.

[151] R. Metzler, A. V. Chechkin, and J. Klafter. Lévy statistics and anomalous transport: Lévy flights and subdiffusion. In R. A. Meyers, editor, *Encyclopedia of Complexity and Systems Science*, pages 5218–5239. Springer, Berlin, 2009.

[152] R. Metzler and J. Klafter. The random walk's guide to anomalous diffusion: A fractional dynamics approach. *Phys. Rep.*, 339:1–77, 2000.

[153] Z. Michna. Self-similar processes in collective risk theory. *J. Appl. Math. Stochastic Anal.*, 11:429–448, 1998.

[154] D. Middleton. Non-Gaussian noise models in signal processing for telecommunications: new methods an results for class a and class b noise models. *IEEE Trans. Inform. Theory*, 45:1129–1149, 1999.

[155] J. Mielniczuk and P. Wojdyllo. Estimation of Hurst exponent revisited. *Computational Statistics & Data Analysis*, 51:4510–4525, 2007.

[156] T. Mikosch, T. Gadrich, C. Klüppelberg, and R. J. Adler. Parameter estimation for ARMA models with infinite variance innovations. *Ann. Statist.*, 23:305–326, 1995.

[157] T. Mizuuchi et al. Edge fluctuation studies in Heliotron J. *Journ. Nuclear Materials*, 337-339:332–336, 2005.

[158] M. Mossberg. Analysis of moments based methods for fractional Gaussian noise estimation. *IEEE Trans. Signal Process.*, 60:3823–3827, 2012.

[159] Z. Nahorski and J. Studziński. Small sample bias reduction of the first-order autoregressive parameter least-squares estimator. *IEEE Trans. Automat. Control*, AC-30:893–895, 1985.

[160] D. Nandy and A. Choudhuri. Explaining the latitudinal distribution of sunspots with deep meridional flow. *Science*, 296:1671–1673, 2002.

[161] C. Nikias and M. Shao. *Signal Processing with Alpha-Stable Distributions and Applications*. Wiley, New York, 1995.

[162] J. Nolan. Maximum likelihood estimation and diagnostics for stable distributions. In O. Barndorff-Nielsen, T. Mikosch, and S. Resnick, editors, *Lévy Processes: Theory and Applications*, pages 379–400. Birkhäuser, Boston, MA, 2001.

[163] J. Nolan. Numerical calculation of stable densities and distribution functions. *Commun. Statist. -Stochastic Models*, 13:759–774, 1997.

[164] I. Norros. On the use of fractional Brownian motion in the theory of connectionless networks. *IEEE Journal on Selected Areas in Communications*, 13:953–962, 1995.

[165] A. Oppenheimer and R. Schafer. *Digital Signal Processing*. Prentice-Hall, Englewood Cliffs, NJ, 1975.

[166] S. Painter. Long-range dependence in the subsurface: Empirical evidence and simulation methods, 1998. Invited paper at the American Geophysical Union 1998 Fall Meeting.

[167] W. Pan et al. Viscoelasticity in homogeneous protein solutions. *Phys. Rev. Lett.*, 102:058101, 2009.

[168] M.S. Paolella. Testing the stable Paretian assumption. *Math. Comput. Modelling*, 34:1095–1112, 2001.

[169] M. Pawlak and Z. Hasiewicz. Nonlinear system identification by the Haar multiresolution analysis. *IEEE Trans. Circuits Systems I Fund. Theory Appl.*, 45:945–961, 1998.

[170] M. Pawlak, Z. Hasiewicz, and P. Wachel. On nonparametric identification of Wiener systems. *IEEE Trans. Signal Process.*, 55:482–492, 2007.

[171] M. Pawlak, E. Rafajłowicz, and A. Krzyżak. Postfiltering versus prefiltering for signal recovery from noisy samples. *IEEE Trans. Inf. Theory*, 49:3195–3212, 2003.

[172] C.-K. Peng, J. Mietus, J. Hausdorff, S. Havlin, H. Stanley, and A. Goldberger. Long-range anticorrelations and non-Gaussian behavior of the heartbeat. *Phys. Rev. Lett.*, 70:1343–1346, 1993.

[173] E. Perrin, R. Harba, R. Jennane, and I. Iribarren. Fast and exact synthesis for 1-D fractional Brownian motion and fractional Gaussian noises. *IEEE Signal Process. Lett.*, 9:382–384, 2002.

[174] B. Pittel, W. A. Woyczyński, and J. A. Mann. Random tree type partitions as a model for acyclic polymerization: Holtsmark (3/2-stable) distribution of the supercritical gel. *Ann. Probab.*, 18:319–341, 1990.

[175] L. Rabiner. A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77:257–286, 1989.

[176] E. Rafajłowicz. Testing (non-)existence of input-output relationships by estimating fractal dimensions. *IEEE Trans. Signal Process.*, 52:3151–3159, 2004.

[177] E. Rafajłowicz, M. Pawlak, and A. Steland. Nonlinear image processing and filtering: A unified approach based on vertically weighted regression. *Int. J. Appl. Math. Comput. Sci.*, 18:49–61, 2008.

[178] P. M. Robinson, editor. *Time Series with Long Memory.* Oxford University Press, 2003.

[179] S. Ross. *Simulation.* Academic Press, San Diego, 2002.

[180] L. Rutkowski. On system identification by nonparametric function fitting. *IEEE Trans. Automat. Control*, 28:225–227, 1982.

[181] L. Rutkowski. On nonparametric identification with prediction of time-varying systems. *IEEE Trans. Automat. Control*, 29:58–60, 1984.

[182] L. Rutkowski. Nonparametric identification of quasi-stationary systems. *Systems Control Lett.*, 6:33–35, 1985.

[183] L. Rutkowski. *New Soft Computing Techniques for System Modeling, Pattern Classification and Image Processing (Studies in Fuzziness and Soft Computing).* Springer, Berlin, 2004.

[184] A. M. Sabatini. A statistical mechanical analysis of postural sway using non-Gaussian FARIMA stochastic models. *IEEE Trans. Biomed. Eng.*, 47:1219–1227, 2000.

[185] S. Samko, A. Kilbas, and D. Maritchev. *Integrals and Derivatives of the Fractional Order and Some of Their Applications*. Gordon and Breach Science Publishers, Amsterdam, 1993.

[186] G. Samorodnitsky. Long memory and self-similar processes. *Ann. Fac. Sci. Toulouse Math.*, 15:107–123, 2006.

[187] G. Samorodnitsky. Long range dependence. *Found. Trends Stoch. Syst.*, 1:163–257, 2006.

[188] G. Samorodnitsky and M. S. Taqqu. *Stable Non-Gaussian Random Processes*. Stochastic Modeling. Chapman & Hall, New York, 1994.

[189] W. Schneider. Stable distributions: Fox function representation and generalization. In S. Albeverio, G. Casati, and D. Merlini, editors, *Stochastic Processes in Classical and Quantum Systems*, volume 262 of *Lecture Notes in Physics*, pages 497–511. Springer Berlin / Heidelberg, 1986.

[190] D. Sims et al. Scaling laws of marine predator search behaviour. *Nature*, 451:1098–1102, 2008.

[191] A. Stanislavsky. Memory effects and macroscopic manifestation of randomness. *Phys. Rev. E*, 61:4752–4759, 2000.

[192] A. Stanislavsky, K. Burnecki, M. Magdziarz, A. Weron, and K. Weron. FARIMA modeling of solar flare activity from empirical time series of soft X-ray solar emission. *Astrophys. J.*, 693:1877–1882, 2009.

[193] W. Staszewski. Identification of non-linear systems using multi-scale ridges and skeletons of the wavelet transform. *J. Sound Vibration*, 214:639–658, 1998.

[194] S. Stoev, V. Pipiras, and M. Taqqu. Estimation of the self-similarity parameter in linear fractional stable motion. *Signal Processing*, 82:1873–1901, 2002.

[195] S. Stoev and M. S. Taqqu. Simulation methods for linear fractional stable motion and FARIMA using the fast Fourier transform. *Fractals*, 12:95–121, 2004.

[196] S. A. Stoev, G. Michailidis, and M. S. Taqqu. Estimating heavy-tail exponents through max self–similarity. *IEEE Trans. Inf. Theory*, 57:1651–1636, 2011.

[197] H. Svensmark and E. Friis-Christensen. Variation of cosmic ray flux and global cloud coverage — a missing link in solar-climate relationship. *Solar -Terr. Phys.*, 59:1225–1232, 1997.

[198] J. Szostak and E. Blackburn. Cloning yeast telomeres on linear plasmid vectors. *Cell*, 29:245–255, 1982.

[199] J. Szostak and E. Blackburn. Identification of a specific telomere terminal transferase activity in Tetrahymena extracts. *Cell*, 43:405–413, 1985.

[200] J. Szostak and E. Blackburn. A telomeric sequence in the RNA of Tetrahymena telomerase required for telomere repeat synthesis. *Nature*, 337:331–337, 1989.

[201] J. Szymanski and M. Weiss. Elucidating the origin of anomalous diffusion in crowded fluids. *Phys. Rev. Lett.*, 103:038102, 2009.

[202] P. Śliwiński. *Nonlinear System Identification by Haar Wavelets*. Springer, Berlin, 2012.

[203] P. Śliwiński and Z. Hasiewicz. Computational algorithms for wavelet identification of nonlinearities in Hammerstein systems with random inputs. *IEEE Trans. Signal Process.*, 56:846–851, 2008.

[204] P. Śliwiński, J. W. Rozenblit, M. W. Marcellin, and R. Klempous. Wavelet amendment of polynomial models in Hammerstein systems identification. *IEEE Trans. Automat. Contr.*, 54:820–825, 2009.

[205] M. S. Taqqu and V. Teverovsky. On estimating the intensity of long-range dependence in finite and infinite variance time series. In *A Practical Guide to Heavy Tails*, pages 177–217. Birkhäuser, Boston, MA, 1998.

[206] V. Teverovsky, M. S. Taqqu, and W. Willinger. A critical look at Lo's modified R/S statistic. *Journal of Statistical Planning and Inference*, 80:211–227, 1999.

[207] H. Tianyun. On use of the alpha stable self-similar stochastic process to model aggregated VBR video traffic. *J. Syst. Eng. Electron.*, 17:677–684, 2006.

[208] I. Tolić-Norrelykke et al. Anomalous diffusion in living yeast cells. *Phys. Rev. Lett.*, 93:078102, 2004.

[209] V. Uchaikin and V. Zolotarev. *Chance and Stability*. VSP Press, Utrecht, 1999.

[210] C. Varotsos and D. Kirk-Davidoff. Long-memory processes in ozone and temperature variations at the region 60° S—60° N. *Atmos. Chem. Phys.*, 6:4096–4100, 2006.

[211] J. Venter and P. de Jongh. Risk estimation using the Normal Inverse Gaussian distribution. *J. Risk*, 4:1–23, 2002.

[212] M. Weiss, M. Elsner, F. Kartberg, and T. Nilsson. Anomalous subdiffusion is a measure for cytoplasmic crowding in living cells. *Biophys. J.*, 87:3518–3524, 2004.

[213] A. Weron, K. Burnecki, S. Mercik, and K. Weron. Complete description of all self-similar models driven by Lévy stable noise. *Phys. Rev. E*, 71:016113, 2005.

[214] A. Weron and M. Magdziarz. Anomalous diffusion and semimartingales. *EPL*, 86:60010, 2009.

[215] R. Weron. Lévy-stable distributions revisited: tail index > 2 does not exclude the Lévy-stable regime. *Int. Journ. Mod. Phys. C*, 12:209–223, 2001.

[216] R. Weron. Computationally intensive value at risk calculations. In J. Gentle, W. Härdle, and Y. Mori, editors, *Handbook of Computational Statistics*, pages 911–950. Springer, Berlin, 2004.

[217] R. Weron. *Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach*. Wiley, Chichester, 2006.

[218] R. Weron and B. Przybylowicz. Hurst analysis of electricity price dynamics. *Phys. A*, 283:462–468, 2000.

[219] R. Weron and A. Weron. Computer simulation of Lévy stable variables and processes. In P. Garbaczewski, M. Wolf, and A. Weron, editors, *Chaos – The Interplay Between Stochastic and Deterministic Behaviour*, volume 457 of *Lecture Notes in Physics*, pages 379–392. Springer Berlin / Heidelberg, 1995.

[220] M. Wheatland. Distribution of flare energies based on independent reconnecting structures. *Sol. Phys.*, 208:33–42, 2002.

[221] M. Wheatland. The energetics of a flaring solar active region and observed flare statistics. *Astrophys. J.*, 679:1621–1628, 2008.

[222] W. Willinger, M. Taqqu, R. Sherman, and D. Wilson. Self-similarity through high-variability: statistical analysis of Ethernet LAN traffic at the source level. *IEEE/ACM Trans. Net.*, 5:71–96, 1997.

[223] W. Woyczyński. *A First Course in Statistics for Signal Analysis. 2nd Ed.* Birkhäuser, Boston, 2011.

[224] W. Woyczyński and A. Saichev. *Distributions in the Physical and Engineering Sciences. Volume 1: Distributional and Fractal Calculus, Integral Transforms and Wavelets.* Birkhäuser, Boston, 1997.

[225] W. Wu, G. Michailidis, and D. Zhang. Simulating sample paths of linear fractional stable motion. *IEEE Trans. Inform. Theory*, 50:1086–1096, 2004.

[226] J. Xiu and Y. Jin. Empirical study of ARFIMA model based on fractional differencing. *Phys. A*, 377:138–152, 2007.

[227] J. Zarzycki. Multidimensional nonlinear Schur parametrization of nonGaussian stochastic signals, part one: Statement of the problem. *Multidimens. Systems Signal Process.*, 15:217–241, 2004.

[228] J. Zarzycki. Multidimensional nonlinear Schur parametrization of nonGaussian stochastic signals, part three: Low-complexity staircase solution. *Multidimens. Systems Signal Process.*, 15:313–340, 2004.

[229] J. Zarzycki. Multidimensional nonlinear Schur parametrization of nonGaussian stochastic signals, part two: Generalized Schur algorithm. *Multidimens. Systems Signal Process.*, 15:243–275, 2004.

[230] A. Zygmund. *Trigonometric Series.* Cambridge University Press, London, 1959.

Wydawnictwa Politechniki Wrocławskiej
są do nabycia w księgarni „Tech”
plac Grunwaldzki 13, 50-377 Wrocław
budynek D-1 PWr., tel. 71 320 29 35
Prowadzimy sprzedaż wysyłkową
zamawianie.ksiazek@pwr.wroc.pl

ISBN 978-83-7493-734-4

INNOVATIVE
ECONOMY
NATIONAL COHESION STRATEGY

Wrocław University of Technology

EUROPEAN UNION
EUROPEAN REGIONAL
DEVELOPMENT FUND