U.P.B. Sci. Bull., Series A, Vol. 85, Iss. 4, 2023 ISSN 1223-7027

CAUCHY SPLIT-BREAK PROCESS: ASYMPTOTIC PROPERTIES AND APPLICATION IN SECURITIES MARKET ANALYSIS

Eugen LJAJKO[1,a], Vladica STOJANOVIĆ[2], Marina TOŠIĆ[1,b], Ivan BOŽOVIĆ[3]

The paper presents a novel kind of nonlinear and non-stationary stochastic process, that can be applicable in the analysis of time series with accentuated and persistent fluctuations. Using Cauchy distributed innovations, the resulting model, named the Cauchy Split-BREAK (CSB) process, was examined in terms of its basic stochastic properties and asymptotic behaviour. To estimate the unknown parameters of the CSB process, an estimation procedure based on empirical characteristic functions is proposed, along with numerical simulations of thus obtained estimators. It is also shown that the CSB process can be a suitable stochastic model for analysing the dynamics of the securities market.

Keywords: nonlinear time series; pronounced and permanent fluctuations; non-stationarity; Cauchy distribution; parameters estimation; simulations.

1. Introduction

Stochastic modelling of time series with accentuated and persistent fluctuations is one of the important topics in contemporary research. To this end, various stochastic models are proposed, primarily devoted to application in social sciences and econometrics [1,16]. A particular problem arises when the observed time series have nonlinear and non-stationary dynamics, which usually reflects in increasing the complexity of their stochastic structure [7,13]. To solve this problem, Engle and Smith [6] proposed the so-called stochastic permanent break (STOPBREAK) process, later investigated by numerous authors, especially in the domain of structural and permanent changes in real-world data fluctuations [5,8]. Besides that, Stojanović et al. [21] introduce the so-called Split-BREAK process, also applied in modelling different time series with constant and pronounced fluctuations. Recently, some more general forms of the Split-BREAK process, the so-called General (that is, Gaussian) Split-BREAK (GSB) process, were introduced and discussed in Stojanović et al. [20,22,23], as well as Jovanović et al. [9].

Using a similar idea, the Split-BREAK model with Cauchy distributed innovations, named Cauchy Split-BREAK (CSB) process, is introduced here. The main motive is the fact that the Cauchy distribution is infinitely divisible and stable,

[1] Associate Professor, Department of Mathematics, Faculty of Sciences, University of Priština, Kosovska Mitrovica, Serbia, e-mail: a. eugen.ljajko@pr.ac.rs; b. marina.tosic@pr.ac.rs
[2] Full Professor, Department of Informatics & Computer Sciences, University of Criminal Investigation and Police Studies, Belgrade, Serbia, e-mail: vladica.stojanovic@kpu.edu.rs
[3] Associate Professor, Department of Macroeconomics, Faculty of Economics, University of Priština, Kosovska Mitrovica, Serbia, e-mail: ivan.bozovic@pr.ac.rs

which, as will be seen, will be important features for determining stochastic distributions of the basic series in the CSB model. In addition, since the Cauchy distribution does not have finite moments, it can be used in modelling the dynamics of non-stationary time series with large fluctuations and long tails, i.e., with a wide range of data. The following section presents the definition and key stochastic properties of the CSB process, applying the characteristic functions (CFs) method. The CSB process parameter estimation procedure, based on empirical characteristic functions (ECFs), is described in Section 3. Thereafter, Section 4 is devoted to Monte Carlo simulations of the proposed estimators, as well as the application of the CSB process in dynamic analysis of the total trading values of QUALCOMM Incorporated Common (QCOM) stocks. Finally, in Section 5 some concluding remarks are given.

2. CSB process. Definition and key properties

The main assumptions about the CSB process can be made based on its corresponding time series as given in the following:

Definition 2.1. *Let* $(\Omega, \mathcal{F}, P)$ *the probability space, expanded with filtration* $F = (\mathcal{F}_t)$, *where* $t = 0,1,\dots,T$ *is the set of time indices. The Cauchy Split-BREAK (CSB) process represents the following time series, defined on expanded basis* $(\Omega, \mathcal{F}, P, F)$:

i) (ε_t) *is an innovation series, that is, the independent identical distributed (IID) random variables (RVs) with zero-centred* $Cauchy(0, \lambda)$ *distribution, whose probability distribution function (PDF) is:*

$$f_\varepsilon(x) = \frac{\lambda}{\pi(x^2 + \lambda^2)}, \quad x \in \mathbb{R}, \tag{1}$$

and $\lambda > 0$ *is the scale parameter.*

ii) (m_t) *is a series of martingale means given by recurrence relation:*

$$m_t = m_{t-1} + q_{t-1}\,\varepsilon_{t-1} = m_0 + \sum_{j=0}^{t-1} q_j \varepsilon_j, \tag{2}$$

where is almost surely (as) $m_0 \overset{as}{=} \mu\ (const)$, $\varepsilon_{-1} = \varepsilon_0 \overset{as}{=} 0$, *and*

$$q_t = I(\varepsilon_{t-1}^2 > c) = \begin{cases} 1, & \varepsilon_{t-1}^2 > c \\ 0, & \varepsilon_{t-1}^2 \le c \end{cases} \tag{3}$$

is the Noise-Indicator with the parameter (critical value) $c > 0$.

iii) (y_t) *is a basic CSB series defined by the adaptive decomposition:*

$$y_t = m_t + \varepsilon_t. \tag{4}$$

Now, we give some practical interpretations of the concepts introduced in Definition 2.1. First, filtration $(\mathcal{F}_t)$ is a set of “information” about some (actual) time series at time t, so the RVs (ε_t) are $\mathcal{F}_t$-adaptive, for each $t = 0,1,\dots,T$. Moreover, according to the well-known fact about the Cauchy distribution, the cumulative distribution function (CDF) of the RVs (ε_t) is:

$$F_\varepsilon(x) \coloneqq P\{\varepsilon_t < x\} = \int_{-\infty}^{x} f_\varepsilon(z)dz = \frac{1}{2} + \frac{1}{\pi} arctan\left(\frac{x}{\lambda}\right), \quad x \in \mathbb{R}, \tag{5}$$

as well as the characteristic function $\varphi_\varepsilon(u) = \mathrm{e}^{-\lambda|u|}$, $u \in \mathbb{R}$. Thus, Cauchy distribution can be viewed as the Fourier transform of the Laplace distribution. Also, Cauchy distribution does not have moment generating function, that is, the finite moments of order greater than or equal to one. Then, assumption of the Cauchy distributed innovations (ε_t) is motivated by the fact that it, compared to other distributions, such as the Gaussian, can more adequately fit empirical distributions with long tails and pronounced peaks (see Section 4). Besides, (m_t) is $\mathcal{F}_{t-1}$ measurable series which represents the predictive and stability component, contrary to the innovations (ε_t) that make the deviation (noise) component of the of the CSB process. Finally, the parameter $c > 0$ is the critical value of reaction, which indicates significance of earlier realizations of series (ε_t), in order to include its current values in Eq. (2). More precisely, when $q_{t-1} = 0$, the martingale mean m_t is equal to its previous value m_{t-1}, and the main CSB series (y_t), given by Eq. (4), is then realized with 'low' fluctuation. Otherwise, the case $q_t = 1$ indicates a pronounced fluctuation of the series (y_t). Having in mind that the series(m_t) and (y_t) depend on the time moment $t \in T$ they are observed in, it is obviousely that they are non-stationary. Moreover, their basic distributional properties can be shown as follows:

Theorem 2.1. *Let (m_t) and (y_t) be the CSB series defined by Eqs. (2) and (4), respectively, where $m_0 \overset{as}{=} \mu\ (const)$. Then, for any $x \in \mathbb{R}$ and $t = 0,\ 1,\dots,T$, the CDFs of (m_t) and (y_t) are:*

$$F_m(x,t) \coloneqq P\{m_t < x\} = \bigotimes_{j=1}^{t} \left[a_c F_\varepsilon^{(j)}(x) + (1-a_c)F_0(x)\right] \otimes F_\mu(x), \tag{6}$$

$$F_y(x,t) \coloneqq P\{y_t < x\} = \bigotimes_{j=1}^{t} \left[a_c F_\varepsilon^{(j)}(x) + (1-a_c)F_0(x)\right] \otimes F_C(x), \tag{7}$$

where "$\otimes$" is the convolution operator, $F_\varepsilon^{(j)}(x)$ and $F_0(x)$ are the CDFs of the RVs ε_t and $I_0 \overset{as}{=} 0$, respectively, $F_C(x) = F_\mu(x) \otimes F_\varepsilon(x)$ is the CDF of the RV $\mu + \varepsilon_t$ with Cauchy (μ,λ) distribution, and $a_c = P\{q_t = 1\} = P\{\varepsilon_t^2 > c\}$. Additionally, for $T = +\infty$, the following convergences (in distribution) hold:

$$\frac{1}{t}m_t \xrightarrow{d} Cauchy(0, a_c\lambda),$$
$$\frac{1}{t}y_t \xrightarrow{d} Cauchy(0, a_c\lambda), \qquad t \to +\infty. \tag{8}$$

Proof. Let us define the RVs $\zeta_t = q_t\varepsilon_t$, $t = 0, 1, \dots, T$, which are easily shown to represent a mutually uncorrelated RVs. Applying the conditional probabilities, the CDF of (ζ_t) is:

$$\begin{aligned} F_\zeta(x) &:= P\{\zeta_t < x\} \\ &= P\{\zeta_t < x | q_t = 1\} \cdot P\{q_t = 1\} + P\{\zeta_t < x | q_t = 0\} \cdot P\{q_t = 0\} \\ &= P\{\varepsilon_t < x\} \cdot P\{q_t = 1\} + P\{x > 0\} \cdot P\{q_t = 0\} \\ &= a_c F_\varepsilon(x) + (1 - a_c)F_0(x). \end{aligned}$$

where $\varphi_0(u) \equiv 1$ is the CF of the RV $I_0 \overset{as}{=} 0$. Based on that, the CF of the RVs (ζ_t) is as follows:

$$\begin{aligned} \varphi_\zeta(u) &= \int_{-\infty}^{+\infty} e^{iux} F_\zeta(dx) = \int_{-\infty}^{+\infty} e^{iux}[a_c F_\varepsilon + (1 - a_c)F_0](dx) \\ &= a_c\varphi_\varepsilon(u) + (1 - a_c)\varphi_0(u) \\ &= 1 + a_c\left(e^{-\lambda|u|} - 1\right). \end{aligned}$$

By applying Eq. (2), for the CFs of the series (m_t) one obtains:

$$\varphi_m(u, t) = \varphi_\mu(u)\prod_{j=0}^{t-1}\varphi_\eta(u) = e^{iu\mu}\left(1 + a_c\left(e^{-\lambda|u|} - 1\right)\right)^t, \tag{9}$$

where $\varphi_\mu(u) = \mathrm{e}^{iu\mu}$ is the CF of the RV $m_0 \overset{as}{=} \mu$. Thus, according to Eq. (9) and Lévy's correspondence theorem, Eq. (6) immediately follows. In a similar way, using Eq. (4), for the CFs of the series (y_t) one obtains:

$$\varphi_y(u, t) = \varphi_m(u)\varphi_\varepsilon(u) = e^{iu\mu - \lambda|u|}\left(1 + a_c\left(e^{-\lambda|u|} - 1\right)\right)^t. \tag{10}$$

Applying again Levy's correspondence theorem to the last expression, Eq. (7) immediately follows.

To prove the convergences in Eqs. (8), let us notice that according to Eqs. (9) and (10), the CFs of RVs m_t/t and y_t/t, when $t > 0$, are as follows:

$$\varphi_m\left(\frac{u}{t}, t\right) = e^{\frac{iu\mu}{t}} \cdot \left(1 + a_c\left(e^{-\frac{\lambda|u|}{t}} - 1\right)\right)^t,$$
$$\varphi_y\left(\frac{u}{t}, t\right) = e^{\frac{iu\mu - \lambda|u|}{t}} \cdot \left(1 + a_c\left(e^{-\frac{\lambda|u|}{t}} - 1\right)\right)^t.$$

Taking the limit values, when $t \to +\infty$ and $u \in \mathbb{R}$ is a fixed (but an arbitrary) value, it is obtained:

$$\lim_{t\to+\infty} \varphi_m\left(\frac{u}{t},t\right) = \lim_{t\to+\infty} \varphi_y\left(\frac{u}{t},t\right) = \lim_{t\to+\infty}\left(1 + a_c\left(e^{-\frac{\lambda|u|}{t}} - 1\right)\right)^t$$
$$= \lim_{t\to+\infty}\left(1 - \frac{a_c\lambda|u|}{t}\right)^t = e^{-a_c\lambda|u|}.$$

The last expression represents the CF of the Cauchy $(0, a_c\lambda)$ distribution, and both convergences in Eq. (8) are confirmed. ■

Remark 2.1. The uncorrelated series (ζ_t) can be interpreted as a 'new' innovation series with 'occasional' zero values. Their corresponding CDF:

$$F_\zeta(x) = a_c F_\varepsilon(x) + (1 - a_c)F_0(x)$$

is obviously continuous almost everywhere, with the sole exception at $x = 0$, where the jump of size $1 - a_c$ occurs (see, e.g. Stojanović et al. [19]). This CDF is the mixture of Cauchy and discrete distribution concentrated at zero, which we call the Contaminated Cauchy Distribution (CCD). Also, the asymptotic relations in Eq. (8) indicate that series (m_t/t) and (y_t/t), generated by non-stationary time series (m_t) and (y_t), converge to the Cauchy distribution, when $t \to +\infty$. These can be easily observed by the convergence of CFs $\varphi_m(u/t, t)$ and $\varphi_y(u/t, t)$, as shown in Fig. 1. ■

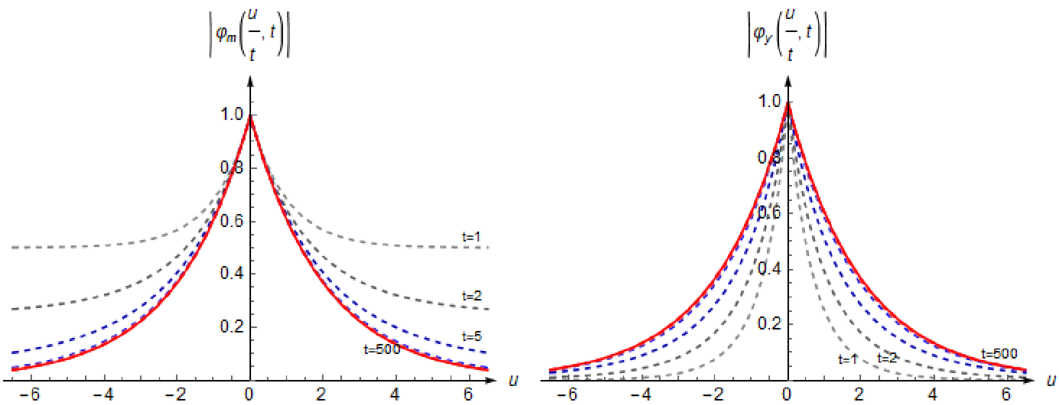

Fig. 1. Modulus convergences for CFs $\varphi_m(u/t, t)$ and $\varphi_y(u/t, t)$, $t = 1, 2, \dots, 500$.

In the following are presented the asymptotic properties of some other linear transformations of the non-stationary CSB series (m_t) and (y_t), also related to the Cauchy distribution.

Theorem 2.2. *Let us define, for an arbitrary* $\alpha \geq 1$, *the* α*-mean time series:*

$$\bar{M}_{t;\alpha} = \frac{1}{t^\alpha}\sum_{j=1}^{t} m_j, \qquad \bar{Y}_{t;\alpha} = \frac{1}{t^\alpha}\sum_{j=1}^{t} y_j,$$

where (m_t) *and* (y_t) *are the non-stationary time series given by Eqs. (2) and (4), respectively. Then, the following statements are valid:*

i) When $1 \leq \alpha \leq 2$*, the series* $\overline{M}_{t;\alpha}$ *and* $\overline{Y}_{t;\alpha}$ *are asymptotically Cauchy distributed, i.e., the following relations hold, when* $t \to +\infty$*:*

$$\overline{M}_{t;\alpha} \sim Cauchy\left(\mu t^{1-\alpha}, \frac{a_c \lambda t^{2-\alpha}}{2}\right),$$
$$\overline{Y}_{t;\alpha} \sim Cauchy\left(\mu t^{1-\alpha}, \lambda t^{1-\alpha} + \frac{a_c \lambda t^{2-\alpha}}{2}\right). \qquad (11)$$

ii) When $\alpha > 2$*, the series* $\overline{M}_{t;\alpha}$ *and* $\overline{Y}_{t;\alpha}$ *vanish asymptotically, i.e.,*

$$\overline{M}_{t;\alpha} \xrightarrow{d} I_0, \qquad \overline{Y}_{t;\alpha} \xrightarrow{d} I_0, \qquad t \to +\infty. \qquad (12)$$

Proof. Firstly we prove the convergence in Eq. (11) for the series $\overline{M}_{t;\alpha}$. Using the definition of series (m_t), given by Eq. (2), it is obtained:

$$\overline{M}_{t;\alpha} = \frac{1}{t^\alpha}\sum_{j=1}^{t} m_j = \frac{1}{t^\alpha}\sum_{j=1}^{t}\left(m_0 + \sum_{k=0}^{j-1} q_k \varepsilon_k\right) = \frac{1}{t^\alpha}\left(t m_0 + \sum_{j=0}^{t-1}(t-j) q_j \varepsilon_j\right)$$
$$= t^{1-\alpha} m_0 + \sum_{k=1}^{t} \frac{k}{t^\alpha} \zeta_{t-k}.$$

Therefore, $\overline{M}_{t;\alpha}$ is the sum of mutually uncorrelated RVs ζ_{t-k}, when $k = 1, \dots, t$. According to the well-known theoretical facts about the CFs, for the CFs of the RVs $\overline{M}_{t;\alpha}$ one obtains:

$$\varphi_{\overline{M};\alpha}(u,t) = \varphi_m\left(\frac{u}{t^{\alpha-1}}, 0\right)\prod_{k=1}^{t}\varphi_\zeta\left(\frac{ku}{t^\alpha}\right) = e^{iu\mu t^{1-\alpha}}\prod_{k=1}^{t}\left[1 + a_c\left(e^{-\frac{\lambda k|u|}{t^\alpha}} - 1\right)\right].$$

Now, let us denote the function:

$$\psi_M(u,t,\alpha) \coloneqq ln\, \varphi_{\overline{M};\alpha}(u,t) = iu\mu t^{1-\alpha} + \sum_{k=1}^{t} f_k(u,t,\alpha), \qquad (13)$$

wherein $f_k(u,t,\alpha) \coloneqq \ln\left[1 + a_c\left(\mathrm{e}^{-\frac{\lambda k|u|}{t^\alpha}} - 1\right)\right]$. Using the asymptotic relations:

$$ln\,(1+x) = x + o(x), \qquad e^x - 1 = x + o(x), \qquad x \to 0,$$

for $t \to +\infty$ and fixed (but an arbitrary) $u \in \mathbb{R}$, we get:

$$f_k(u,t,\alpha) = a_c\left(e^{-\frac{\lambda k|u|}{t^\alpha}} - 1\right) + o_k(t^{-\alpha}u) = -\frac{a_c \lambda k|u|}{t^\alpha} + o_k(t^{-\alpha}u),$$

where $o(z) \to 0$ and $o_k(z) \to 0$, when $z \to 0$. Substituting the last expression in Eq. (13) follows:

$$\psi_M(u,t,\alpha) = iu\mu t^{1-\alpha} - \frac{a_c\lambda}{t^\alpha}\sum_{k=1}^{t}\left(k|u| + o_k(t^{-\alpha}u)\right)$$
$$= iu\mu t^{1-\alpha} - \frac{a_c\lambda}{2t^\alpha}\cdot t(t+1)|u| + o(t^{1-\alpha}u),$$

and taking $t \to +\infty$, one obtains:

$$\psi_M(u,t,\alpha) \sim \begin{cases} iu\mu t^{1-\alpha} - a_c\lambda t^{2-\alpha}|u|/2, & 1 \le \alpha \le 2 \\ 0, & \alpha > 2. \end{cases}$$

Thus, replacing them into CFs $\varphi_{\bar{M};\alpha}(u,t)$, the first asymptotic relation in Eq. (11) is easily obtained.

A similar procedure can be conducted for the series $\bar{Y}_{t;\alpha}$. Using the previously proven facts and Eq. (4), we find that:

$$\bar{Y}_{t;\alpha} = \frac{1}{t^\alpha}\sum_{j=1}^{t}\left(m_j + \varepsilon_j\right) = \bar{M}_{t;\alpha} + \sum_{j=1}^{t}\frac{\varepsilon_j}{t^\alpha} = t^{1-\alpha}m_0 + \sum_{k=1}^{t}\frac{k}{t^\alpha}\zeta_{t-k} + \sum_{k=0}^{t-1}\frac{\varepsilon_{t-k}}{t^\alpha}$$
$$= t^{1-\alpha}m_0 + \frac{\varepsilon_t}{t^\alpha} + \sum_{k=1}^{t}(1 + kq_{t-k})\frac{\varepsilon_{t-k}}{t^\alpha}.$$

Since ε_{t-k}, $k = 0, 1, \dots, t$, are mutually independent RVs, the CFs of $\bar{Y}_{t;\alpha}$ are obtained as follows:

$$\varphi_{\bar{Y};\alpha}(u,t) = \varphi_m\left(\frac{u}{t^{\alpha-1}}, 0\right)\varphi_\varepsilon\left(\frac{u}{t^\alpha}\right)\prod_{k=1}^{t}\left[(1-a_c)\varphi_\varepsilon\left(\frac{u}{t^\alpha}\right) + a_c\,\varphi_\varepsilon\left(\frac{(k+1)u}{t^\alpha}\right)\right]$$
$$= e^{iu\mu t^{1-\alpha} - \frac{\lambda|u|}{t^\alpha}}\prod_{k=1}^{t}\left[(1-a_c)e^{-\frac{\lambda|u|}{t^\alpha}} + a_c\,e^{-\frac{\lambda(k+1)|u|}{t^\alpha}}\right]$$
$$= e^{iu\mu t^{1-\alpha}}\left(e^{-\frac{\lambda|u|}{t^\alpha}}\right)^{t+1}\prod_{k=1}^{t}\left[1 + a_c\left(e^{-\frac{\lambda k|u|}{t^\alpha}} - 1\right)\right].$$

Applying the same procedure as above, for the function $\psi_Y(u,t,\alpha) := \ln\varphi_{\bar{Y};\alpha}(u,t)$ we have:

$$\psi_Y(u,t,\alpha) = iu\mu t^{1-\alpha} - \frac{\lambda(t+1)|u|}{t^\alpha} + \sum_{k=1}^{t} ln\left[1 + a_c\left(e^{-\frac{\lambda k|u|}{t^\alpha}} - 1\right)\right]$$
$$= iu\mu t^{1-\alpha} - \frac{\lambda(t+1)|u|}{t^\alpha} - \frac{a_c\lambda}{t^\alpha}\sum_{k=1}^{t}\left(k|\,u| + o_k(t^{-\alpha}u)\right)$$
$$= iu\mu t^{1-\alpha} - \frac{\lambda(t+1)|u|}{t^\alpha} - \frac{a_c\lambda}{2t^\alpha}\cdot t(t+1)|u| + o(t^{1-\alpha}u).$$

Taking $t \to +\infty$ in the last expression, one obtains:

$$\psi_Y(u,t,\alpha) \sim \begin{cases} iu\mu t^{1-\alpha} - \lambda t^{1-\alpha}|u| - a_c\lambda t^{2-\alpha}|u|/2, & 1 \le \alpha \le 2 \\ 0, & \alpha > 2, \end{cases}$$

and replacing this expression into CFs $\varphi_{\bar{Y};\alpha}(u,t)$ the theorem is completely proven. ■

Remark 2.2. The previous theorem gives important features of the non-stationary CSB series, that is, shows that series (m_t) and (y_t) are asymptotically closed for the Cauchy distribution under some linear transformations. The case of $\alpha = 2$ should be especially emphasized as an interesting one, since relations in Eq. (11) in that case give:

$$\frac{1}{t^2}\sum_{j=1}^{t} m_j \xrightarrow{d} Cauchy\left(0, \frac{a_c\lambda}{2}\right), \quad \frac{1}{t^2}\sum_{j=1}^{t} y_j \xrightarrow{d} Cauchy\left(0, \frac{a_c\lambda}{2}\right), \quad t \to +\infty. \tag{14}$$

This is a generalized version of the central limit theorem for the so-called stable distributions (see, e.g. Campbell et al. [2, pp. 778]). ■

At last part of this section, let us define another CSB series, the so-called increments:

$$X_t = y_t - y_{t-1}, \quad t = 1, \dots, T. \tag{15}$$

In accordance with Eqs. (1), (2) and (6), we can represent this series as:

$$X_t = \varepsilon_t - \theta_{t-1}\varepsilon_{t-1}, \tag{16}$$

where $\theta_t = 1 - q_t = I(\varepsilon_{t-1}^2 \le c)$. Obviously, the series (X_t) is a stationary stochastic process with a random coefficient θ_t, and it operates in two modes:

a) Emphasized fluctuations of the series (ε_t) in the previous time moment implicate $\theta_{t-1} = 0$, and Eq. (16) becomes $X_t = \varepsilon_t$.
b) If ε_{t-1}^2 do not overdraw the critical value c, it follows $\theta_{t-1} = 1$ and X_t is given as a linear, integrated MA(1) process $X_t = \varepsilon_t - \varepsilon_{t-1}$.

For these reasons, one can consider the series (X_t) to be an 'optional' moving average (MA) stochastic process and, therefore, express its key stochastic properties in the following way:

Theorem 2.3. *Let (X_t) be the CSB series defined by Eqs. (15) and (16). Then, for any $x \in \mathbb{R}$ and $t = 0, 1, \dots, T$, the CDF of the RVs (X_t) is given by*:

$$F_X(x) := P\{X_t < x\} = (1 - b_c)F_\varepsilon(x) + b_c F_{2\varepsilon}(x), \tag{17}$$

where $b_c = E(\theta_t) = P\{\varepsilon_{t-1}^2 \le c\} = 1 - a_c$, and $F_\varepsilon(x)$, $F_{2\varepsilon}(x)$ are, respectively, the CDFs of the Cauchy $(0, \lambda)$ and Cauchy $(0, 2\lambda)$ distributions.

Proof. Similarly as above, let us define the series of uncorrelated RVs $\eta_t = \theta_t\varepsilon_t$, $t = 0, 1, \dots, T$. Using the conditional probabilities, for the CDF of the RVs (ξ_t) one obtains:

$$F_\eta(x) := P\{\eta_t < x\}$$
$$= P\{\eta_t < x|\theta_t = 1\} \cdot P\{\theta_t = 1\} + P\{\eta_t < x|\theta_t = 0\} \cdot P\{\theta_t = 0\}$$
$$= P\{\varepsilon_t < x\} \cdot P\{\theta_t = 1\} + P\{x > 0\} \cdot P\{\theta_t = 0\}$$
$$= b_c F_\varepsilon(x) + (1 - b_c) F_0(x).$$

According to this, the CF of the RVs (η_t) is obtained as follows:

$$\varphi_\eta(u) := \int_{-\infty}^{+\infty} e^{iux} F_\eta(dx) = \int_{-\infty}^{+\infty} e^{iux} [b_c F_\varepsilon + (1 - b_c) F_0](dx)$$
$$= b_c \varphi_\varepsilon(u) + (1 - b_c)\varphi_0(u)$$
$$= 1 + b_c\left(e^{-\lambda|u|} - 1\right). \tag{18}$$

According to Eq. (18), for the CF of the series (X_t), given by Eq. (16), we get:

$$\varphi_X(u) = \varphi_\varepsilon(u) \cdot \varphi_\eta(-u) = (1 - b_c)e^{-\lambda|u|} + b_c e^{-2\lambda|u|}. \tag{19}$$

Thus, Eq. (17) immediately follows by applying the Levy correspondence theorem to Eq. (19). ■

Remark 2.3. Note that by differentiating Eq. (17), we obtain the PDF of the series(X_t) as follows:

$$f_X(x) = (1 - b_c)\frac{dF_\varepsilon(x)}{dx} + b_c \frac{dF_{2\varepsilon}(x)}{dx} = \frac{\lambda}{\pi}\left(\frac{1 - b_c}{x^2 + \lambda^2} + \frac{2b_c}{x^2 + 4\lambda^2}\right). ■$$

3. Parameters estimation. Empirical characteristic function method

In this part, we estimate the unknown parameters of the CSB process, that is, the critical value (c) and the scale parameter (λ). For that cause, we denote further $\theta = (b_c, \lambda)'$ and use the increments (X_t), which, as already mentioned, are the only observable and stationary series of the CSB processes. Also, this series has a similar structure to linear MA processes, but with Cauchy distributed innovations (ε_t). The Cauchy distribution is peculiar purpose of its heavy tail and the difficulty in estimating its parameters (see, e.g. [12]). For instance, the moment-based estimation procedures cannot be applied because the mean and variance of the Cauchy distribution do not exist, while the maximum likelihood estimators (MLEs) require complex calculations. Therefore, here we propose the empirical characteristic function (ECF) method, based on matching the ECF with the theoretical CF of the stationary series (X_t).

The ECF method was implemented in time series analysis by the pioneering work of Knight and Satchell [10], and later examined in detail by Knight and Yu [11] and Yu [24]. Thereafter, some extensions of CF-based estimators are discussed, e.g., in Meintanis [14] or Carrasco and Kotchoni [3]. Following these ideas, an ECF procedure similar to those in Stojanović et al. [18,20] is described here. It is worth pointing out that the main preeminence of the ECF method is the fact that the theoretical CFs are uniformly bounded, which implies the numerical stability of the estimators obtained in this manner. Furthermore, in accordance with the bijective correspondence between the CFs and their corresponding CDFs, the ECFs retain all the 'information' present in the sample. In that sense, the general definition of the CF of order $r \geq 1$ can be given as follows:

Definition 3.1. *Let* $\mathbf{u} = (u_1, \dots, u_r) \in \mathbb{R}^r$ *and* $\boldsymbol{X}_t^{(r)} := (X_t, \dots, X_{t+r-1})'$, $t = 0, 1, \dots, T - \mathrm{r} + 1$, *be the overlapping blocks of the series* (X_t). *The* r*-dimensional CF of vector* $\boldsymbol{X}_t^{(r)}$ *is given as follows:*

$$\varphi_X^{(r)}(\boldsymbol{u};\theta) := E\left[exp\left(i\boldsymbol{u}'\boldsymbol{X}_t^{(r)}\right)\right] = E\left[exp\left(i\textstyle\sum_{j=1}^{r} u_j X_{t+j-1}\right)\right]. \quad (20)$$

An explicit expression for the CFs of increments (X_t) is given by the following statement:

Theorem 3.1. *Let* (X_t) *be the series introduced by Eqs. (15) and (16). Then, the CFs of order* $r \in \mathbb{N}$ *of the* r*-dimensional stochastics process* $\left(\boldsymbol{X}_t^{(r)}\right)$ *are given by:*

$$\varphi_X^{(r)}(\boldsymbol{u};\theta) = e^{-\lambda|u_r|}\prod_{j=0}^{r-1}\left[(1-b_c)e^{-\lambda|u_j|} + b_c e^{-\lambda|u_j - u_{j+1}|}\right], \quad (21)$$

where $u_0 = 0$ *and* $b_c = P\{\varepsilon_{t-1}^2 \leq c\}$.

Proof. Note first that, according to Eq. (19), the statement is obviously valid in the case when $r = 1$ and $u_1 = u$. Now, suppose that $r > 1$ and denote:

$$\begin{aligned}\mathcal{L}(\boldsymbol{u};\theta) &\coloneqq exp\left(i\boldsymbol{u}'\boldsymbol{X}_t^{(r)}\right) = exp\left(i\sum_{j=0}^{r-1} u_{j+1}X_{t+j}\right) \\ &= exp\left[i\sum_{j=0}^{r-1} u_{j+1}\left(\varepsilon_{t+j} - \theta_{t+j-1}\varepsilon_{t+j-1}\right)\right] \\ &= exp\left[i\left(u_r\varepsilon_{t+r-1} + \sum_{j=0}^{r-1}\left(u_j - \theta_{t+j-1}u_{j+1}\right)\varepsilon_{t+j-1}\right)\right].\end{aligned}$$

According to the last expression and Eqs. (18) and (20), the r-dimensional CF of the series $\left(\boldsymbol{X}_t^{(r)}\right)$ is obtained as $\varphi_X^{(r)}(\mathbf{u};\theta) := E[\mathcal{L}(\mathbf{u};\theta)]$, and Eq. (21) immediately follows. ■

Further, let us denote $\boldsymbol{X}_T := \{X_1, \dots, X_T\}$ as some realization of length $T \in \mathbb{N}$ of the increments (X_t), as well as r-dimensional ECF matching them as:

$$\tilde{\varphi}_T^{(r)}(\boldsymbol{u}) := \frac{1}{T-r+1}\sum_{t=1}^{T-r+1} exp\left(i\boldsymbol{u}'\boldsymbol{X}_t^{(r)}\right).$$

As previously stated, the main objective of the ECF method is to minimize the 'distance' between the theoretical CF and its corresponding ECF. The appropriate ECF estimators are then obtained by a minimization the following objective function:

$$S_T^{(r)}(\theta) := \int_{\mathbb{R}^r} g(\boldsymbol{u})\left|\varphi_X^{(r)}(\boldsymbol{u};\theta) - \tilde{\varphi}_T^{(r)}(\boldsymbol{u})\right|^2 \boldsymbol{du} \quad (22)$$

with respect to the parameter $\theta = (b_c, \lambda)'$. Here, $\varphi_X^{(r)}(\mathbf{u};\theta)$ is the CF of the order $r \geq 1$, defined by Eq. (20), $\mathbf{du} := \mathrm{d}u_1 \cdots \mathrm{d}u_r$, and $g: \mathbb{R}^r \to \mathbb{R}^+$ is some weight function. Therefore, the ECF estimates are solutions to the following minimization equation:

$$\hat{\theta}_T^{(r)} = \arg\min_{\theta\in\Theta} S_T^{(r)}(\theta),$$

where $\Theta = (0,1) \times (0, +\infty)$ is a non-trivial parameter space. According to some general results of ECF-asymptotic theory (see, e.g., Knight and Yu [11] or Stojanović et al. [18]), strong consistency and asymptotic normality (AN) of the ECF estimators, under some regulatory conditions, can be proved. Moreover, the above procedure holds if CF is of order $r \geq 1$ at least equal to the number of its parameters. For that purpose, we base the estimation procedure on the two-dimensional CF of the vector series $\boldsymbol{X}_t^{(2)} := (X_t, X_{t+1})'$. The objective function $S_T^{(2)}$ then represents a double integral with weight $g: \mathbb{R}^2 \to \mathbb{R}^+$, and it can be numerically calculated by using some cubature formulas. Let us notice that, according to Eq. (21), the two-dimensional CF of the series (X_t) can be expressed in an explicit form as follows:

$$\varphi_X^{(2)}(u_1, u_2; \theta) = e^{-\lambda|u_2|}\left[1 + b_c\left(e^{-\lambda|u_1|} - 1\right)\right]\left[(1-b_c)e^{-\lambda|u_1|} + b_c e^{-\lambda|u_1-u_2|}\right], \quad (23)$$

and the appropriate ECF is the real-valued function:

$$\tilde{\varphi}_T^{(2)}(u_1, u_2; \theta) = \frac{1}{T-1}\sum_{t=1}^{T-1} cos(u_1 X_t + u_2 X_{t+1})]. \quad (23)$$

As an illustration, in Fig. 2 are shown 3D plots of the two-dimensional CF and the corresponding ECF of increments (X_t), when $T = 1500$, $b_c = 0.5$, and $\lambda = 1$.

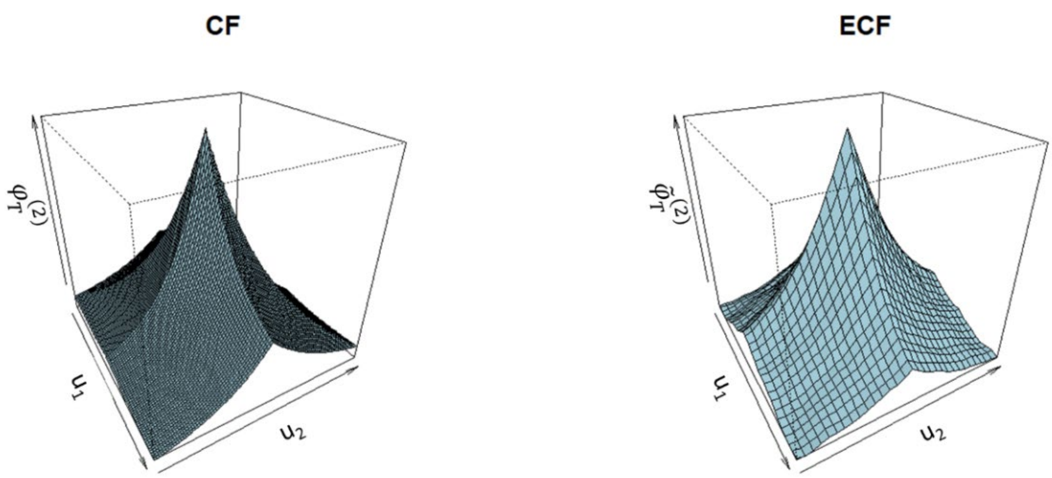

Fig. 2. 3D plots of the two-dimensional CF (left) and the corresponding ECF (right) of the series $\boldsymbol{X}_t^{(2)} = (X_t, X_{t+1})'$.

4. Numerical simulations & application

In this part, we present the implementation of the previously mentioned ECF procedure for estimating the parameters $\theta = (b_c, \lambda)'$ of the increment series (X_t). Thereby, according to Eq. (5), estimates $\hat{c}$ of the critical value can be easily obtained by solving the following equation (with respect to $c > 0$):

$$P\{\varepsilon_t^2 \leq c\} = \hat{b}_c \quad \Leftrightarrow \quad \hat{c} = \left[F_\varepsilon^{-1}\left(\frac{\hat{b}_c + 1}{2}\right)\right]^2 = \lambda\, tg\left(\frac{\pi \hat{b}_c}{2}\right). \tag{24}$$

Using the aforementioned results, primarily Eqs. (22) and (23), ECF estimates can be calculated by minimizing the following double integral:

$$S_T^{(2)}(\theta) = \int_{\mathbb{R}^2} g(\boldsymbol{u}) \left|\varphi_X^{(2)}(\boldsymbol{u};\theta) - \tilde{\varphi}_T^{(2)}(\boldsymbol{u})\right|^2 \boldsymbol{du}. \tag{25}$$

Here, $g(\mathbf{u}) = g(u_1, u_2) = \exp\left(-(u_1^2 + u_2^2)\right)$ is the exponential weight, which places more weights around the origin, in accordance with the fact that the CF at this point contains the most information about the PDF of the estimated model. In order to solve integral in Eq. (25), the following numerical approximation is used, based on the N-point Gauss-Hermitian cubature formula:

$$I(f, g) := \iint_{\mathbb{R}^2} g(u_1, u_2) f(u_1, u_2)\, du_1 du_2 \approx \sum_{j=1}^{N} \omega_j f\left(v_{1j}, v_{2j}\right). \tag{26}$$

Here, $\left(v_{1j}, v_{2j}\right)$ are the cubature nodes, and ω_j are the appropriate weight coefficients, obtained using the package "Orthogonal polynomials" [4] in WOLFRAM MATHEMATICA software. In our case, the objective function in Eq. (25)

is minimized using the cubature formulas in Eq. (26), with $N = 81$ nodes and using the R-function "nlminb".

Further, two sample sizes $T = 250$ and $T = 1500$ have been considered, and $N = 500$ independent realizations $\{X_1, \dots, X_T\}$ of the series (X_t) with Cauchy innovations (ε_t) were generated for both of them. True parameters values are $b_c = 0.5$ and $c = \lambda = 1$, and their initial estimates were taken randomly from the uniform distributions $\mathcal{U}(0,1)$ and $\mathcal{U}(0,2)$, respectively. Table 1 shows the obtained numerical results, that is, the mean values (Mean), minimums (Min.), maximums (Max.), the mean-squared estimation error (MSEE), along with the values of the objective function $S_T^{(2)}(\theta)$. It is evident that ECF estimates converge, because the MSEE and $S_T^{(2)}(\theta)$ values decrease as the sample size increases. In addition, notice that the estimates of the critical value $c > 0$ have a slightly higher MSEE, as a consequence of the two-step estimation procedure, based on Eq. (24).

Table 1.

Estimated parameters obtained from Monte Carlo simulations of the CSB process. (True parameters are: $b_c = 0.5, c = \lambda = 1$)

Sample size	$T = 250$				$T = 1500$			
Parameters	b_c	c	λ	$S_T^{(2)}$	b_c	c	λ	$S_T^{(2)}$
Min.	0.3006	0.6956	0.7019	2.63E-06	0.3255	0.7274	0.7257	1.84E-06
Mean	0.4974	1.0710	1.0059	3.21E-05	0.4989	1.0582	1.0054	2.67E-05
Max.	0.6995	1.3652	1.3140	1.39E-04	0.6743	1.2987	1.2744	9.00E-05
MSEE	0.0136	0.0403	0.0290	-	0.0100	0.0303	0.0248	-

Thereafter, in order to display the practical application of the CSB process, the fitting of the dynamics of the total trading values of QUALCOMM Incorporated Common (QCOM) stocks is described. The sample data set is taken on the basis of official stock market quotations from the National Association of Securities Dealers Automated Quotations (NASDAQ) [15]. In this way, five-year historical data, from May 29, 2018. until May 24, 2023. are considered as univariate time series of the length $T = 1257$. In addition, the so-called log-volumes, obtained as the natural logarithm of the total monetary value of the trading volume, are observed as the basic time series:

$$y_t := ln(P_t \cdot V_t), \quad t = 0, 1, \dots, T,$$

where (P_t) and (V_t) are, respectively, price and trading volumes of QCOM stocks. The use of log-volume, as pointed out in [17], changes the interpretation of activity shocks, because the growth trend does not affect unexpected values in their dynamics. Additionally, the increments of CSB process can be expressed as follows:

$$X_t := y_t - y_{t-1} = ln\frac{P_t}{P_{t-1}} + ln\frac{V_t}{V_{t-1}}, \qquad t = 1, \dots, T,$$

that is, they represent the sum of the log-returns of stock prices and trading volumes. Thereafter, by using Eqs. (2)−(4), the series (m_t) and (ε_t) can be obtained by the following recurrence procedure:

$$\begin{cases} \varepsilon_t = y_t - m_t, \\ m_t = m_{t-1} + \varepsilon_{t-1} I\{\varepsilon_{t-2}^2 \geq \hat{c}\}. \end{cases} \tag{27}$$

Here, $\hat{c}$ is the estimated critical value, obtained according to Eq. (24), and starting values for the iterative procedure in Eq. (24) are $\varepsilon_0 = \varepsilon_{-1} = 0$. Using the well-known facts about the Cauchy distribution (see, e.g. [12]), the median $\hat{\mu}$ of the series (y_t) is used as an estimate of the parameter μ. At the same time, using the modelled values (ε_t), given by Eq. (27), the mean absolute deviation (MAD):

$$\tilde{\lambda} = \frac{1}{T} \sum_{t=1}^{T} |\varepsilon_t - \hat{\mu}|$$

is taken as the initial estimate of the scale parameter λ. The ECF procedure mentioned above is then applied and thus obtained estimated parameters values, along with the key statistical indicators of the CSB series, are shown in Table 2.

Table 2.

Estimated parameters and key statistical indicators of the QCOM stocks data

Parameters	Estimates	Statistical indicators	CSB series			
			(y_t)	(m_t)	(X_t)	(ε_t)
b_c	1.31E−03	Min.	19.115	19.115	-1.5517	-1.5517
c	6.79E−04	Max.	23.236	23.235	2.2832	3.0026
λ	0.3033	Median	20.632	20.231	-0.0212	-0.0186
$S_T^{(2)}$	9.89E−05	MAD	0.4169	0.3327	0.3306	0.4067

Based on these results, it can be noted that the estimates of the log-volumes (y_t) and the martingale means (m_t) are quite ‘close to each other’. Also, the increments (X_t) and the innovation series (ε_t) have similar estimated values. Note that this is a consequence of previous theoretical results given in Theorems 2.1. and 2.3. Finally, a ‘small’ estimated values of parameters b_c (and c) indicate that their true values are $b_c = c = 0$. Therefore, the series (X_t) and (ε_t) become equal, which means that (X_t) have a Cauchy (μ, λ) distribution. This implies:

$$X_t = y_t - y_{t-1} = \varepsilon_t \iff y_t = y_{t-1} + \varepsilon_t,$$

that is, the series (y_t) has independent increments. Thus, according to Eq. (1) it follows $y_{t-1} = m_t$, and all of the 'past information' is contained into previous realization of (y_t). It makes the overall statistical analysis simpler, because according to Theorem 2.2, RVs (y_t) then have a Cauchy $(\mu, \lambda t)$ distribution. As an illustration, the empirical PDFs (given by histograms) and the theoretical PDFs (given by lines) of the CSB series (y_t) and (X_t), respectively, are shown in Fig. 3.

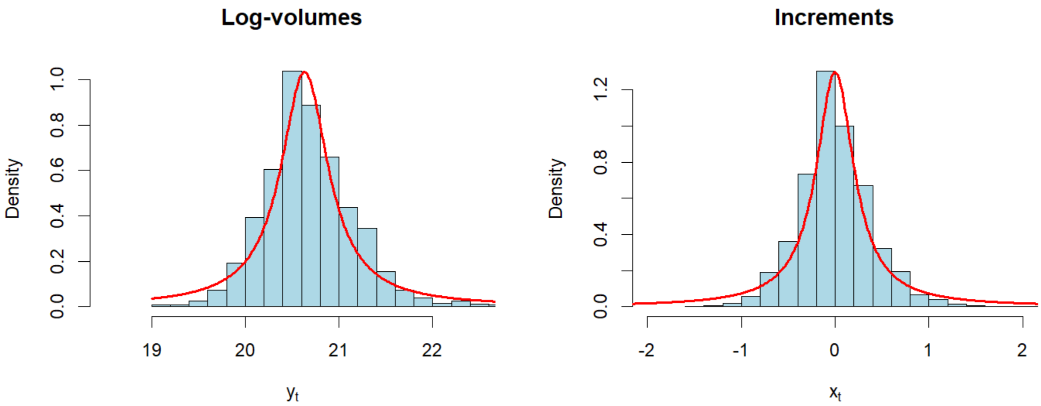

Fig. 3. Empirical distributions of real-world data (histograms) and their corresponding PDFs fitted with Cauchy distributions (lines).

5. Conclusions

The manuscript presents a new nonlinear stochastic model, named the Cauchy Split-BREAK (CSB) process, convenient for empirical analysis of time series with persistent and pronounced fluctuations. Stochastic characteristics of the CSB process are investigated, with special emphasis on its asymptotic properties. We implemented a procedure based on the ECF method for the CSB model parameters estimation. Thus obtained results were applied in modelling the dynamics of the total value of the trading volume of QCOM shares. It is worth noting that, with certain modifications, similar stochastic models as well as estimation techniques can be used to fit some related non-linear time series.

R E F E R E N C E S

[1]. *O. Blanchard, J. L'Huillier, G. Lorenzoni,* News, Noise, and Fluctuations: An Empirical Exploration, Am. Econ. Rev. **103:7**(2013), 3045–3070.

[2]. *B. Campbell, N. Balakrishnan; B. Vidakovic; S. Kotz,* Encyclopedia of Statistical Sciences, John Wiley & Sons, 2006.

[3]. M. *Carrasco, R. Kotchoni,* Efficient Estimation Using the Characteristic Function, Econometric Theory **33:2**(2017), 479–526.

[4]. *A.S. Cvetković, G.V. Milovanović*, The Mathematica Package "OrthogonalPolynomials", Facta Univ. Ser. Math. Inform. **19**(2004), 17–36.

[5]. *Y. Dendramis, G. Kapetanios, E. Tzavalis,* Shifts in Volatility Driven by Large Stock Market Shocks, J. Econom. Dynam. Control **55**(2015), 130–147.

[6]. *R.F. Engle, A.D. Smith,* Stochastic Permanent Breaks, Rev. Econ. Stat. **81:4**(1999), 553–574.

[7]. *C. Granero-Belinchón, S. Roux, N. Garnier,* Information Theory for Non-Stationary Processes with Stationary Increments, Entropy **21**(2019), Article No. 1223.

[8]. *U. Hassler, P.M.M. Rodrigues, A. Rubia,* Persistence in the Banking Industry: Fractional Integration and Breaks in Memory, J. Emp. Finan. **29**(2014), 95–112.

[9]. *M. Jovanović, V. Stojanović, K. Kuk, B. Popović, P. Čisar,* Asymptotic Properties and Application of GSB Process: A Case Study of the COVID-19 Dynamics in Serbia, Mathematics **10**(2022), Article No. 3849.

[10]. *J.L. Knight, S.E. Satchell,* Estimation of Stationary Stochastic Processes via the Empirical Characteristic Function, Department of Economics, University of Western Ontario, 1996.

[11]. *J.L. Knight, J. Yu,* Empirical Characteristic Function in Time Series Estimation, Econometric Theory **18**(2002), 691–721.
[12]. *W.B. Lindquista, S.T. Racheva,* Taylor's Law and Heavy-tailed Distributions, PNAS, **118:50**(2021), Article No. e2118893118.
[13]. *E. Ljajko, M. Tošić, T. Kevkić, V. Stojanović*, Application of the Homotopy Perturbations Method in Approximation Probability Distributions of Non-Linear Time Series, U.P.B. Sci. Bull., Series A **83:2**(2021), 177–186.
[14]. *S.G. Meintanis,* A Review of Testing Procedures Based on the Empirical Characteristic Function, South African Statist. J. **50:1**(2016), 1–14.
[15]. *Nasdaq,* https://nasdaq.com (accessed on May 26, 2023).
[16]. *N. Rebei, R. Sbia,* Transitory and Permanent Shocks in the Global Market for Crude Oil, J. Appl. Econometrics, **36:7**(2021), 1047–1064.
[17]. *M.K. So, C.W. Chen, T.C. Chiang, D.S. Lin,* Modelling Financial Time Series with Threshold Nonlinearity in Returns and Trading Volume, Appl. Stoch. Models Bus. Ind. **23:4**(2007), 319-338.
[18]. *V. Stojanović, T. Kevkić, E. Ljajko, G. Jelić,* Noise-Indicator ARMA Model with Application in Fitting Physically-Based Time Series, U.P.B. Sci. Bull., Series A, **81:2**(2019), 257–264.
[19]. *V. Stojanović, T. Kevkić, G. Jelić,* Application of the Homotopy Analysis Method in Approximation of Convolutions Stochastic Distributions, U.P.B. Sci. Bull., Series A **79:4**(2017), 103-112.
[20]. *V. Stojanović, G.V. Milovanović, G. Jelić,* Distributional Properties and Parameters Estimation of GSB Process: An Approach Based on Characteristic Functions, ALEA–Lat. Am. J. Prob. Math. Stat. **13:2**(2016), 835–861.
[21]. *V. Stojanović, B.Č. Popović, P. Popović*, The Split-BREAK Model, Braz. J. Prob. Stat. **25:1**(2011), 44–63.
[22]. *V. Stojanović, B.Č. Popović, P. Popović*, Stochastic Analysis of GSB Process, Publ. Inst. Math. **95:109**(2014), 149–159.
[23]. *V. Stojanović, B.Č. Popović, P. Popović*, Model of General Split-BREAK Process, REVSTAT– Stat. J. **13:2**(2015), 145–168.
[24]. *J. Yu*, Empirical characteristic function estimation and its applications, Econometric Rev. **23:2**(2004), 93–123.