The Burr XII autoregressive moving average model

Fernando José Monteiro de Araújo*[a], Renata Rojas Guerra[a], Fernando Arturo Peña-Ramírez[b]

[a]*Federal University of Santa Maria, Santa Maria, Brazil*
[b]*Universidad Nacional de Colombia, Bogota, Colombia*

Abstract

The present work introduces a new model class for random variables that have support in the positive real line. This model is designed to explain conditional quantiles and provides an alternative approach for modeling data with asymmetric behavior and heavy tails. Specifically, we present a novel autoregressive moving average model based on the τ–th quantile of the Burr XII distribution. The advantage of using the quantile is that it is less sensitive to heterogeneous populations and more robust in the presence of outliers than the average. Our proposed model enables the dynamic modeling of any quantile through a structured approach incorporating autoregressive terms, moving averages, time-varying regressors, and a link function. We adopt the conditional maximum likelihood method to estimate the model parameters and construct confidence intervals. Furthermore, we assess the performance of the proposed model's parameter estimators through Monte Carlo simulations. We also demonstrate the model's usefulness through diagnostic tools and empirical applications on two datasets related to the financial market and the environment. Furthermore, we also compare the new model's performance to that of competing models.

Keywords: Asymmetric data, Burr XII, conditional quantile, time series, forecast

1. Introduction

The Burr XII (BXII) distribution was initially introduced by Burr (Burr, 1942) as part of a system of distributions. The BXII is the twelfth model in this system and has found applications in various fields, including income studies (Bhatti et al., 2021; Guerra et al., 2021), poverty indicators (Dhongde and Minoiu, 2013; Thompson, 2013), anomaly detection (Sagrillo et al., 2023), and survival analysis (Low et al., 2021; Ramires et al., 2021). It is often used in economics to model income data and is sometimes referred to as the Singh-Maddala distribution (Singh and Maddala, 1976). Tadikamalla (1980), Zimmer et al. (1998), and Watkins (1997) have explored some of the properties of the BXII model in the methodological literature. In recent years, researchers have also investigated generalizations of the BXII distribution, including the Marshall-Olkin generalized Burr XII (Muse et al., 2021), Unit Burr XII (Ribeiro et al., 2022), Reflected Unit Burr XII (Ribeiro et al., 2021), and Generalized log-logistic Burr XII (Muse et al., 2021), among others.

In recent years, a growing number of researchers have focused on developing time series models under the assumption of non-Gaussian distributions. One such seed development is the Generalized Autoregressive Moving Average (GARMA) models, introduced by Benjamin

*Corresponding author

Email addresses: `nandomonteiro418@gmail.com` (Fernando José Monteiro de Araújo*), `renata.r.guerra@ufsm.br` (Renata Rojas Guerra), `fernando.p.ramirez@ufsm.br` (Fernando Arturo Peña-Ramírez)

et al. (2003). In the GARMA framework, the dependent variable is modeled as following an exponential family distribution, conditioned on its past history. One of the most widely used generalization models is for positive series with a Gamma distribution. After the proposal of the GARMA model by Benjamin et al. (2003), proposing new models that are not based on a Gaussian structure was left aside. However, it has recently had an ascendancy in the literature, Bayer et al. (2020) addressed an autoregressive moving average model based on the Rayleigh (RARMA) distribution for variables that assume values in positive reals. For variables that assume values in the range (0, 1), Cribari-Neto et al. (2023) they formulated a Beta autoregressive moving average (β-ARMA) model, and in the studies of Melchior et al. (2021), the Kumaraswamy autoregressive moving average (KARMA) model was introduced.

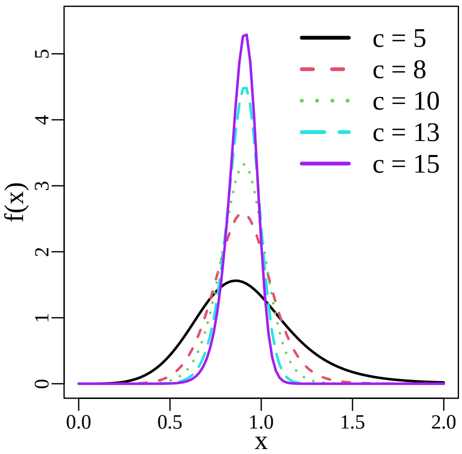
In the context of forecasting, usual models have already been used to model financial market series to forecast economic indices (Jeong and Lee, 2019), leasing value (Korbi and Lleshaj, 2020), and stock prices (Koh et al., 2020). Articles on trading volume forecasts have also been published recently, in the stock market in India Shah et al. (2022), Lee and Park (2022) in the real estate market. Therefore, time series analysis is a strategic tool used to predict future trends and moreover, it helps investors to make decisions that lead to good results and more profits. Wind speed data has also been exploited, mainly to help policymakers in the economic grid market to better utilize wind energy and to know the significant impact that forecasts can provide (Wang et al., 2018; Wu et al., 2022; Chen et al., 2022). Still considering applications in meteorological data, (Elsaraiti et al., 2019) analyzed the monthly average speed of Yenagoa, Nigeria, (Elsaraiti et al., 2019) predicted wind speed on historical data from the Chester region of Nova Scotia, Canada, (Huang and Gu, 2019) analyzed the wind speed series of Typhoon Chan-Hom and, finally, (Rasaki et al., 2018) proposed a model to predict the average monthly temperature in the city of Lagos, Nigeria.

The BXII is a distribution for random variables with support in the positive real line and has the ability to represent asymmetric behaviors and heavy tails. These characteristics make the BXII a suitable alternative both in survival analysis applications and in economic, hydrological, and environmental indicators. It is in this context that the present work is inserted, which aims to propose a new autoregressive moving average model based on a reparametrization in terms of the quantiles of the BXII distribution. This new class of model can help in the analysis and prediction of variables with these asymmetry characteristics.

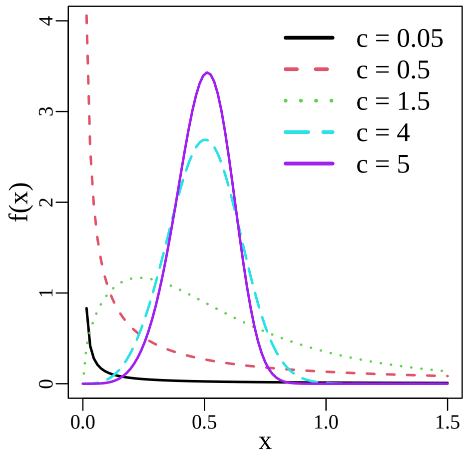
This paper is divided into seven sections which include the introduction. In Section 2, the BXII-ARMA model is proposed. Section 3 presents the estimation of the parameters via conditional maximum likelihood. In Section 4, attentiveness to the diagnostic analysis of the forecast. Section 5 presents the Monte Carlo simulations. Finally, Sections 6 and 7 contain the applications and the final remarks, respectively.

2. The Burr XII ARMA model

This section defines the BXII autoregressive moving average (BXII-ARMA) time series model. For that purpose, we consider the parametrization of the BXII distribution proposed by Araújo et al. (2022), which is based on the quantities μ and τ, where μ is the τ–th quantile with $\tau \in (0,1)$ a known constant. This parameterization satisfies the relation $d = -\log(1-\tau)\,/\,\log(1+\mu^c)$. Figure 1 shows the probability density function (pdf) shapes for reparameterized BXII with different parameter values. We provide a link to show the shapes that the BXII can take (`https://visionmt.shinyapps.io/RBXII/`).



(a) $\mu = 0.9$



(b) $\mu = 0.5$

Figure 1: Plots of the BXII pdf for different values of c, $\tau = 0.5$, and (a) $\mu = 0.9$, (b) $\mu = 0.5$.

Let $Y_1, Y_2, \ldots$ be a sequence of random variables, where each Y_t ($t \in \mathbb{Z}$) assumes values $y_t \in \mathbb{R}^+$. Besides, let $\mathcal{F}_t$ be the σ-field generated by past observations $\{\ldots, y_{t-2}, y_{t-1}, y_t\}$ (i.e., the smallest σ-field such that the variables $Y_1, \ldots, Y_t$ are measurable). Additionally, suppose that each Y_t conditional on previous information set $\mathcal{F}_{t-1}$ is distributed following a BXII law with parameters $c > 0$ and μ_t, where μ_t is the conditional τ–th quantile of Y_t. Thus, the conditional pdf of Y_t given $\mathcal{F}_{t-1}$ is

$$g\left(y_t|\mathcal{F}_{t-1}\right) = \log\left(\frac{1}{1-\tau}\right) \frac{c y_t^{c-1}}{\log\left(1+\mu_t^c\right)} \left(1 + y_t^c\right)^{\log(1-\tau)/\log\left(1+\mu_t^c\right)-1}, \tag{1}$$

and we denote as $Y_t|\mathcal{F}_{t-1} \sim \mathrm{BXII}(\mu_t, c)$.

The conditional cumulative distribution function (cdf) and conditional quantile function (cqf) of $Y_t|\mathcal{F}_{t-1}$ are

$$G\left(y_t|\mathcal{F}_{t-1}\right) = 1 - \left(1 + y_t^c\right)^{\log(1-\tau)/\log\left(1+\mu_t^c\right)} \tag{2}$$

and

$$Q\left(u|\mathcal{F}_{t-1}\right) = \left[(1-u)^{\log\left(1+\mu_t^c\right)/\log(1-\tau)} - 1\right]^{1/c},$$

respectively.

The h–th condition moment of $Y_t|\mathcal{F}_{t-1}$ can be expressed as

$$\mathrm{E}(Y_t^h|\mathcal{F}_{t-1}) = -\frac{\log\left(1-\tau\right)}{\log\left(1+\mu_t^c\right)} \mathrm{B}\left(-\frac{\log\left(1-\tau\right)}{\log\left(1+\mu_t^c\right)} - hc^{-1}, 1 + hc^{-1}\right),$$

where $h < c\ \left[-\log\left(1-\tau\right)/\log\left(1+\mu_t^c\right)\right]$ and $\mathrm{B}(\cdot)$ is the Beta function.

The dynamic general BXII-ARMA model has the following specification for the τ–th conditional quantile

$$\eta_t = g(\mu_t) = \alpha + x_t^{\top}\boldsymbol{\beta} + \sum_{i=1}^{p} \phi_i \left\{g(y_{t-i}) - x_{t-i}^{\top}\boldsymbol{\beta}\right\} + \sum_{j=1}^{q} \theta_j r_{t-j}, \tag{3}$$

where η_t is the linear predictor, $\alpha \in \mathbb{R}$ is a constant, $\boldsymbol{x_t}$ denotes the k-dimensional vector containing the covariates at time t, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_k)^\top$ is a k-dimensional vector of unknown coefficients associated to the covariates, $g(\cdot)$ is a strictly increasing and twice differentiable link function that relates the linear predictor to the τ–th quantile. This function has important purposes in interpreting the response variable. The $r_t = g(y_t) - g(\mu_t)$ term correspond to the random error; $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_p)^\top$, and $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_q)^\top$ are the autoregressive and moving average coefficients, respectively.

3. Parameter estimation

This section describes the estimation of the parameters of the model by conditional maximum likelihood method. Let $y_1, \ldots, y_n$ be the random sample satisfying the specification given by Equations (1) and (3) with $\boldsymbol{\delta} = (\alpha, \boldsymbol{\beta}^\top, \boldsymbol{\phi}^\top, \boldsymbol{\theta}^\top, c)^\top$ denoting the $(2 + k + p + q)$-dimensional parametric vector. The log-likelihood function of the BXII-ARMA model is expressed as

$$\ell = \ell(\boldsymbol{\delta}; y_t | \mathcal{F}_{t-1}) = \sum_{t=m+1}^{n} \ell_t(\mu_t, c), \tag{4}$$

where

$$\ell_t(\mu_t, c) = \log\left[\frac{c \log\left(\frac{1}{1-\tau}\right)}{\log(1 + \mu_t^c)}\right] + (c - 1)\log(y_t) + \left[\frac{\log(1 - \tau)}{\log(1 + \mu_t^c)} - 1\right]\log(1 + y_t^c),$$

$\mu_t = g^{-1}(\eta_t)$, and η_t is defined in Equation (3). Note that Equation (4) holds, once $\ell_t(\mu_t, c)$ is null for the first $m = \max(p, q)$ observations of y_t.

The conditional maximum likelihood estimators (CMLE), $\hat{\boldsymbol{\delta}}$ for $\boldsymbol{\delta}$ are obtained through the maximization of Equation (4). Furthermore, the CMLE are obtained by equalizing the score vector to zero and solving the resulting system of equations. The calculations for the conditionals score vector and Fisher's information matrix, and the construction of confidence intervals and hypothesis testings considering $\tau = 0.5$, the median for the BXII-ARMA model, are presented in the following sections.

3.1. Conditional score vector

The condicional score vector is given by $\boldsymbol{U}(\boldsymbol{\delta}) = [U_\alpha(\boldsymbol{\delta})^\top, \boldsymbol{U_\beta}(\boldsymbol{\delta})^\top, \boldsymbol{U_\phi}(\boldsymbol{\delta})^\top, \boldsymbol{U_\theta}(\boldsymbol{\delta})^\top, U_c(\boldsymbol{\delta})^\top]^\top$, where $\boldsymbol{U}_r(\boldsymbol{\delta}) = \partial\ell/\partial r$, for $r \in \boldsymbol{\delta}$. Therefore, using the chain rule, the $(2 + k + p + q)$ first components of the conditional score vector are obtained as

$$\frac{\partial \ell}{\partial \delta_r} = \sum_{t=m+1}^{n} \frac{\partial \ell_t(\mu_t, c)}{\partial \mu_t} \frac{\mathrm{d}\mu_t}{\mathrm{d}\eta_t} \frac{\partial \eta_t}{\partial \delta_r} = \sum_{t=m+1}^{n} \frac{z_t}{g'(\mu_1)} \frac{\partial \eta_t}{\partial \delta_r},$$

where $\mathrm{d}\mu_t/\mathrm{d}\eta = 1/g'(\mu_1)$, δ_r is the r–th element of $\boldsymbol{\delta}$ and

$$z_t = \frac{\partial \ell_t(\mu_t, c)}{\partial \mu_t} = -\frac{c\mu_t^{c-1}}{u_t \log(u_t)} \left[1 + h_t \log(1 + y_t^c)\right],$$

with $u_t = (1 + \mu_t^c)$ and $h_t = \log(1 - \tau)/\log(u_t)$.

The first order partial derivatives, $\partial\eta_t/\partial\delta_r$, can be obtained as

$$\frac{\partial\eta_t}{\partial\alpha} = 1 - \sum_{j=1}^{q}\theta_j\frac{\partial\eta_{t-j}}{\partial\alpha}, \quad \text{for } r = 1,$$

$$\frac{\partial\eta_t}{\partial\beta_l} = x_{tl} - \sum_{i=1}^{p}\phi_i x_{(t-i)l} - \sum_{j=1}^{q}\theta_j\frac{\partial\eta_{t-j}}{\partial\beta_l}, \quad \text{for } r = 2,\ldots,k+1, \text{ and } l = 1,\ldots,k,$$

$$\frac{\partial\eta_t}{\partial\phi_i} = g(y_{t-i}) - x_{t-i}^{\top}\boldsymbol{\beta} - \sum_{j=1}^{q}\theta_j\frac{\partial\eta_{t-j}}{\partial\phi_i}, \quad \text{for } r = k+2,\ldots,p+1, \text{ and } i = 1,\ldots,p,$$

and

$$\frac{\partial\eta_t}{\partial\theta_j} = r_{t-j} - \sum_{j=1}^{q}\theta_j\frac{\partial\eta_{t-j}}{\partial\theta_j}, \quad \text{for } r = p+2,\ldots,q+1, \text{ and } j = 1,\ldots,q.$$

The last component of the conditional score vector is given by

$$\frac{\partial\ell}{\partial c} = \sum_{t=m+1}^{n}\frac{\partial_t(\mu_t, c)}{\partial c} = \sum_{t=m+1}^{n} w_t,$$

where

$$w_t = \frac{1}{c} + \log(y_t) - \frac{(1-h_t)\,y_t^c\log(y_t)}{1+y_t^c} - \frac{\mu_t^c\log(\mu_t)}{u_t\log(u_t)}\left[1 + h_t\log(1+y_t^c)\right].$$

Let T a diagonal matrix defined by $T = \mathrm{diag}\{1/g'(\mu_{m+1}),\ldots,1/g'(\mu_n)\}$, $\boldsymbol{\nu} = \left(\frac{\partial\eta_{m+1}}{\partial\alpha},\ldots,\frac{\partial\eta_n}{\partial\alpha}\right)^{\top}$, $\mathbf{z} = (z_{m+1},\ldots,z_n)^{\top}$, and $\boldsymbol{M}, \boldsymbol{P}, \boldsymbol{R}$ are matrices with dimension $(n-m)\times k$, $(n-m)\times p$ and $(n-m)\times q$, respectively. The (i,j)–th element of these matrices are determined by

$$M_{i,j} = \frac{\partial\eta_{i+m}}{\partial\beta_j}, \quad P_{i,j} = \frac{\partial\eta_{i+m}}{\partial\phi_j}, \quad \text{and} \quad R_{i,j} = \frac{\partial\eta_{i+m}}{\partial\theta_j},$$

respectively. Thus, the elements of the conditional score vector reduce to

$$\begin{aligned} U_\alpha(\boldsymbol{\delta}) &= \boldsymbol{\nu}^{\top}\boldsymbol{T}\,\mathbf{z},\\ \boldsymbol{U}_{\boldsymbol{\beta}}(\boldsymbol{\delta}) &= \boldsymbol{M}^{\top}\boldsymbol{T}\,\mathbf{z},\\ \boldsymbol{U}_{\boldsymbol{\phi}}(\boldsymbol{\delta}) &= \boldsymbol{P}^{\top}\boldsymbol{T}\,\mathbf{z},\\ \boldsymbol{U}_{\boldsymbol{\theta}}(\boldsymbol{\delta}) &= \boldsymbol{R}^{\top}\boldsymbol{T}\,\mathbf{z} \end{aligned}$$

and

$$U_c(\boldsymbol{\delta}) = \boldsymbol{w}^{\top}\boldsymbol{T}\,\mathbf{1}.$$

The CMLE $\hat{\boldsymbol{\delta}}$ of $\boldsymbol{\delta}$ are obtained through the joint solution of the system of nonlinear equations $\boldsymbol{U}(\boldsymbol{\delta}) = \mathbf{0}$, where $\mathbf{0}$ is a null vector in $\mathbb{R}^{k+p+q+2}$. However, it is not possible to solve the derivatives analytically, and numerical optimization algorithms are indispensable. The *Broyden-Fletcher-Goldfarb-Shanno* (BFGS) algorithm (Ruszczynski, 2011), implemented in R software (Team, 2019), is used as the optimization method. The initial guesses for the moving average parameters, $\boldsymbol{\theta}$, are equal to zero and the initial guess for the shape parameter c is one, as suggested by Bayer et al. (2017). For α, $\boldsymbol{\beta}$ and $\boldsymbol{\phi}$, are estimated by the ordinary least squares method of a linear regression model. To adjust this model, we consider the procedure analogous to KARMA (Bayer et al., 2017).

3.2. Conditional Fisher's information matrix

This section presents the conditional Fisher's information matrix of the BXII-ARMA model. The second-order derivatives of ℓ, with respect to $\boldsymbol{\delta}$ are obtained using the chain rule as

$$\begin{aligned}\frac{\partial^2 \ell_t(\mu_t, c)}{\partial \delta_i \partial \delta_j} &= \sum_{t=m+1}^{n} \frac{\partial}{\partial \mu_t}\left(\frac{\partial \ell_t(\mu_t, c)}{\partial \mu_t}\frac{\mathrm{d}\mu_t}{\mathrm{d}\eta_t}\frac{\partial \eta_t}{\partial \delta_j}\right)\frac{\mathrm{d}\mu_t}{\mathrm{d}\eta_t}\frac{\partial \eta_t}{\partial \delta_i} \\ &= \sum_{t=m+1}^{n}\left[\frac{\partial^2 \ell_t(\mu_t, c)}{\partial \mu_t^2}\frac{\mathrm{d}\mu_t}{\mathrm{d}\eta_t}\frac{\partial \eta_t}{\partial \delta_j} + \frac{\partial \ell_t(\mu_t, c)}{\partial \mu_t}\frac{\partial}{\partial \mu_t}\left(\frac{\mathrm{d}\mu_t}{\mathrm{d}\eta_t}\frac{\partial \eta_t}{\partial \delta_j}\right)\right]\left(\frac{\mathrm{d}\mu_t}{\mathrm{d}\eta_t}\frac{\partial \eta_t}{\partial \delta_i}\right),\end{aligned}$$

where $i, j \in \{1, \ldots, k+p+q+2\}$, and

$$\frac{\partial^2 \ell_t(\mu_t)}{\partial \mu_t^2} = \frac{c\mu_t^{c-2}}{u_t^2 \log(u_t)}\left[u_t - c + \frac{c\mu_t^c}{\log(u_t)}\right] + \frac{c\mu_t^{c-2} h_t \log(1+y_t^c)}{u_t^2 \log(u_t)}\left[u_t - c + \frac{2c\mu_t^c}{\log(u_t)}\right]. \tag{5}$$

Since $\mathbb{E}(\partial \ell_t(\mu_t)/\partial \mu_t \big| \mathcal{F}_{t-1}) = 0$, it follows that

$$\mathbb{E}\left(\frac{\partial^2 \ell_t(\mu_t)}{\partial \delta_i \partial \delta_j}\middle|\mathcal{F}_{t-1}\right) = \sum_{t=m+1}^{n} \mathbb{E}\left(\frac{\partial^2 \ell_t(\mu_t)}{\partial \mu_t^2}\middle|\mathcal{F}_{t-1}\right)\left(\frac{d\mu_t}{d\eta_t}\right)^2 \frac{\partial \eta_t}{\partial \delta_i}\frac{\partial \eta_t}{\partial \delta_j}. \tag{6}$$

By combining (5) and (A.1) from Appendix A and replacing in (6) we obtain

$$\mathbb{E}\left(\frac{\partial^2 \ell}{\partial \delta_i \partial \delta_j}\middle|\mathcal{F}_{t-1}\right) = \sum_{t=m+1}^{n} \frac{a_t}{g'(\mu_t)^2}\frac{\partial \eta_t}{\partial \delta_i}\frac{\partial \eta_t}{\partial \delta_j},$$

where

$$a_t = -\frac{c^2 \mu_t^{2(c-1)}}{u_t^2 \log^2(u_t)}\frac{1}{[\mathrm{g}'(\mu_t)]^2}$$

The derivated of $U_c(\boldsymbol{\delta})$ whith respect of $\boldsymbol{\delta}^* = (\alpha, \boldsymbol{\beta}^\top, \boldsymbol{\phi}^\top, \boldsymbol{\theta}^\top)$, is given by

$$\frac{\partial^2 \ell_t(\mu_t, c)}{\partial \delta_i^* \partial c} = \sum_{t=m+1}^{n} \frac{\partial}{\partial c}\left(\frac{\partial^2 \ell_t(\mu_t, c)}{\partial \mu_t \partial c}\right)\frac{\mu_t}{\eta_t}\frac{\eta_t}{\mu_t} = \sum_{t=m+1}^{n} \frac{\partial z_t}{\partial c}\frac{\mu_t}{\eta_t}\frac{\eta_t}{\mu_t},$$

where δ_i^* is the i–th element of $\boldsymbol{\delta}^*$.

$$\begin{aligned}\frac{\partial z_t}{\partial c} = &\frac{c\mu_t^{c-1}}{u_t \log(u_t)}\left[\frac{\mu_t^c \log(\mu_t)}{u_t \log(u_t)} - \frac{h_t y_t^c \log(y_t)}{1+y_t^c} - \frac{\log(\mu_t)}{u_t} - \frac{1}{c}\right] \\ &+ \frac{\mu_t^{c-1} h_t \log(1+y_t^c)}{u_t^2 \log(u_t)}\left[\frac{2c\mu_t^c \log(\mu_t)}{\log(u_t)} - c\log(\mu_t) - u_t\right].\end{aligned}$$

Therefore, we have (A.1) from Appendix A that

$$\mathbb{E}\left(\frac{\partial^2 \ell}{\partial \delta_i^* \partial c}\middle|\mathcal{F}_{t-1}\right) = \sum_{t=m+1}^{n} \frac{b_t}{g'(\mu_t)}\frac{\partial \eta_t}{\partial \delta_i^*},$$

and b_t, is expressed as

$$b_t = -\frac{\mu_t^{c-1}}{u_t \log(u_t)}\left\{\frac{c\mu_t^c \log(\mu_t)}{u_t \log(u_t)} - \frac{\log(1-\tau)\,[\gamma - 1 + \psi(-h_t)]}{\log(u_t) - \log(1-\tau)}\right\},$$

where γ denotes the *Euler-Mascheroni* constant and $\psi(\cdot)$ is the digamma function.

Finally, we calculate the darivate of $U_c(\boldsymbol{\delta})$ whith respect of c as

$$\frac{\partial^2 \ell_t(\mu_t,\, c)}{\partial c^2} = \frac{h_t \mu_t^c \log^2(\mu_t)}{u_t^2 \log(u_t)} \left[\frac{2\mu_t^c \sum_{i=1}^n \log(1+y_t^c)}{\log(u_t)} - \frac{2u_t}{\log(\mu_t)} \sum_{i=1}^n \left[\frac{y_t^c \log(y_t)}{1+y_t^c} \right] - \sum_{i=1}^n \log(1+y_t^c) \right]$$
$$+ (h_t - 1) \sum_{i=1}^n \left[\frac{y_t^c \log^2(y_t)}{(1+y_t^c)^2} \right] + \frac{n\mu_t^c \log^2(\mu_t)}{u_t^2 \log(u_t)} \left[\frac{\mu_t^c}{\log(u_t)} - 1 \right] - \frac{n}{c^2}, \tag{7}$$

with that using the result in A.1 from Appendix A, we obtain

$$\mathbb{E}\left(\frac{\partial^2 \ell}{\partial c^2} \middle| \mathcal{F}_{t-1} \right) = \sum_{t=m+1}^n d_t,$$

where

$$d_t = -\frac{\mu_t^c \log^2(\mu_t)}{u_t^2 \log^2(u_t)} \left\{ \mu_t^c - \frac{2\log(1-\tau)\left[(\psi(-h_t) + \gamma - 1\right] u_t \log(u_t))}{\log(\mu_t) c [\log(u_t) - \log(1-\tau)]} \right\} - \frac{n}{c^2}$$
$$+ \frac{\log(1-\tau)}{6c^2[2\log(u_i) - \log(1-\tau)]} \left\{ \pi^2 + 6\left[\psi'(1-h_t) + 2\gamma \left[\psi(1-h_t) - 1 \right] \right. \right.$$
$$\left. \left. + \; \psi(1-h_t)\left[\psi(1-h_t) - 2 \right] + \gamma^2 \right] \right\},$$

for $t = m+1, \ldots, n$, and $\psi'(\cdot)$ is the trigamma function.

Let $A = \operatorname{diag}\{a_{m+1}, \ldots, a_n\}$, $B = \operatorname{diag}\{b_{m+1}, \ldots, b_n\}$, $D = \operatorname{diag}\{d_{m+1}, \ldots, d_n\}$, the joint conditional Fisher information matrix for $\boldsymbol{\delta}$ is

$$\boldsymbol{K} = \boldsymbol{K}(\boldsymbol{\delta}) = \begin{pmatrix} K_{(\alpha,\alpha)} & \boldsymbol{K}_{(\alpha,\boldsymbol{\beta})} & \boldsymbol{K}_{(\alpha,\boldsymbol{\phi})} & \boldsymbol{K}_{(\alpha,\boldsymbol{\theta})} & K_{(\alpha,c)} \\ \boldsymbol{K}_{(\boldsymbol{\beta},\alpha)} & \boldsymbol{K}_{(\boldsymbol{\beta},\boldsymbol{\beta})} & \boldsymbol{K}_{(\boldsymbol{\beta},\boldsymbol{\phi})} & \boldsymbol{K}_{(\boldsymbol{\beta},\boldsymbol{\theta})} & \boldsymbol{K}_{(\boldsymbol{\beta},c)} \\ \boldsymbol{K}_{(\boldsymbol{\phi},\alpha)} & \boldsymbol{K}_{(\boldsymbol{\phi},\boldsymbol{\beta})} & \boldsymbol{K}_{(\boldsymbol{\phi},\boldsymbol{\phi})} & \boldsymbol{K}_{(\boldsymbol{\phi},\boldsymbol{\theta})} & \boldsymbol{K}_{(\boldsymbol{\phi},c)} \\ \boldsymbol{K}_{(\boldsymbol{\theta},\alpha)} & \boldsymbol{K}_{(\boldsymbol{\theta},\boldsymbol{\beta})} & \boldsymbol{K}_{(\boldsymbol{\theta},\boldsymbol{\phi})} & \boldsymbol{K}_{(\boldsymbol{\theta},\boldsymbol{\theta})} & \boldsymbol{K}_{(\boldsymbol{\theta},c)} \\ K_{(c,\alpha)} & \boldsymbol{K}_{(c,\boldsymbol{\beta})} & \boldsymbol{K}_{(c,\boldsymbol{\phi})} & \boldsymbol{K}_{(c,\boldsymbol{\theta})} & K_{(c,c)} \end{pmatrix}, \tag{8}$$

where $K_{(\alpha,\alpha)} = \boldsymbol{v}^\top \boldsymbol{A}\boldsymbol{T}^2 \boldsymbol{v}$, $\boldsymbol{K}_{(\alpha,\boldsymbol{\beta})} = \boldsymbol{K}^\top_{(\boldsymbol{\beta},\alpha)} = \boldsymbol{v}^\top \boldsymbol{A}\boldsymbol{T}^2 \boldsymbol{M}$, $\boldsymbol{K}_{(\alpha,\boldsymbol{\phi})} = \boldsymbol{K}^\top_{(\boldsymbol{\phi},\alpha)} = \boldsymbol{v}^\top \boldsymbol{A}\boldsymbol{T}^2 \boldsymbol{P}$, $\boldsymbol{K}_{(\alpha,\boldsymbol{\theta})} = \boldsymbol{K}^\top_{(\boldsymbol{\theta},\alpha)} = \boldsymbol{v}^\top \boldsymbol{A}\boldsymbol{T}^2 \boldsymbol{R}$, $K^\top_{(\alpha,\,c)} = K^\top_{(c,\alpha)} = \boldsymbol{v}^\top \boldsymbol{B}\boldsymbol{T}\boldsymbol{1}$, $\boldsymbol{K}_{(\boldsymbol{\beta},\boldsymbol{\beta})} = \boldsymbol{M}^\top \boldsymbol{A}\boldsymbol{T}^2 \boldsymbol{M}$, $\boldsymbol{K}_{(\boldsymbol{\beta},\boldsymbol{\phi})} = \boldsymbol{K}^\top_{(\boldsymbol{\phi},\boldsymbol{\beta})} = \boldsymbol{M}^\top \boldsymbol{A}\boldsymbol{T}^2 \boldsymbol{P}$, $\boldsymbol{K}_{(\boldsymbol{\beta},\boldsymbol{\theta})} = \boldsymbol{K}^\top_{(\boldsymbol{\theta},\boldsymbol{\beta})} = \boldsymbol{M}^\top \boldsymbol{A}\boldsymbol{T}^2 \boldsymbol{R}$, $\boldsymbol{K}_{(\boldsymbol{\beta},c)} = \boldsymbol{K}^\top_{(c,\boldsymbol{\beta})} = \boldsymbol{M}^\top \boldsymbol{B}\boldsymbol{T}\boldsymbol{1}$, $\boldsymbol{K}_{(\boldsymbol{\phi},\boldsymbol{\phi})} = \boldsymbol{P}^\top \boldsymbol{A}\boldsymbol{T}^2 \boldsymbol{P}$, $\boldsymbol{K}_{(\boldsymbol{\phi},\boldsymbol{\theta})} = \boldsymbol{K}^\top_{(\boldsymbol{\theta},\boldsymbol{\phi})} = \boldsymbol{P}^\top \boldsymbol{A}\boldsymbol{T}^2 \boldsymbol{R}$, $\boldsymbol{K}_{(\boldsymbol{\phi},c)} = \boldsymbol{K}^\top_{(c,\boldsymbol{\phi})} = \boldsymbol{P}^\top \boldsymbol{B}\boldsymbol{T}\boldsymbol{1}$ $\boldsymbol{K}_{(\boldsymbol{\theta},\boldsymbol{\theta})} = \boldsymbol{R}^\top \boldsymbol{A}\boldsymbol{T}^2 \boldsymbol{R}$, $\boldsymbol{K}_{(\boldsymbol{\theta},c)} = \boldsymbol{K}^\top_{(c,\boldsymbol{\theta})} = \boldsymbol{R}^\top \boldsymbol{B}\boldsymbol{T}\boldsymbol{1}$, $K_{(c,c)} = \operatorname{tr}(\boldsymbol{S})$.

3.3. Confidence Intervals and hypothesis testing

Under usual regularity conditions, the asymptotic normality property of CMLE ensures that when the sample size increases

$$\hat{\boldsymbol{\delta}} \sim N_{(2+k+p+q)}\left(\boldsymbol{\delta},\, \boldsymbol{K}^{-1} \right),$$

approximately, where $\hat{\boldsymbol{\delta}} = (\hat{\alpha}, \hat{\boldsymbol{\beta}}^\top, \hat{\boldsymbol{\phi}}^\top, \hat{\boldsymbol{\theta}}^\top,\, \hat{c})^\top$ is CMLE of the parametric vector $\boldsymbol{\delta}$, $N_{(2+k+p+q)}$ denotes a $(2+k+p+q)$–multivariate normal distribution and $\boldsymbol{K}^{-1}$ is the inverse of conditional Fisher's information matrix. Based on the asymptotic normality property, we can build confidence intervals for the model parameters. Let $\hat{\delta}_r$ be the r–th component of $\hat{\boldsymbol{\delta}}$, with $r = 1, \ldots, (2+k+p+q)$. Follow that $(\hat{\delta}_r - \delta_r)\widehat{\text{ep}}(\hat{\delta}_r)^{-1} \sim N(0,\, 1)$, where the $\widehat{\text{ep}}(\hat{\delta}_r)$ is an estimate for the standard error of $\hat{\delta}_r$ equal to $\sqrt{\boldsymbol{K}(\hat{\boldsymbol{\delta}})^{rr}}$ and $\boldsymbol{K}(\hat{\boldsymbol{\delta}})^{rr}$ is the r–th diagonal element of $\boldsymbol{K}^{-1}(\hat{\boldsymbol{\delta}})$. Therefore, the asymptotic confidence interval for δ_r considering $(1-\alpha)\times 100\%$ confidential, is expressed by

$$[\hat{\delta}_r - z_{1-\alpha/2}\,\widehat{\text{ep}}(\hat{\delta}_r);\ \hat{\delta}_r + z_{1-\alpha/2}\,\widehat{\text{ep}}(\hat{\delta}_r)],$$

where $0 < \alpha < 1/2$ is the significance level and z_ζ indicates the quantile ζ of the normal distribution $N(0, 1)$ (Davison and Hinkley, 1997).

To test if a given value, say δ_r^0, is equivalent to the true value of the parameter δ_r, that is, $H_0 : \delta_r = \delta_r^0$ vs $H_1 : \delta_r \neq \delta_r^0$, we can consider Wald's statistic (Wald, 1943), expressed by

$$\text{Z} = \frac{\delta_r - \delta_r^0}{\widehat{\text{ep}}(\hat{\delta}_r)}.$$

The Z-test statistic follows an asymptotically standard normal distribution and is compared to the quantiles of the standard normal distribution (Pawitan, 2001).

4. Diagnostic analysis forecasting

In this section, we present a diagnostic analysis forecasting to identify whether a particular model is suitable and fully captures the dynamics of a data set. The diagnostic analysis is fundamental to ensure more accurate forecasts in time-series studies. In what follows, we present and discuss some forecast diagnostic measures.

We use some comparison and model selection measures, which are obtained through the maximized conditional log-likelihood function of the BXII-ARMA model. They are known as Akaike Information Criterion (AIC) (Akaike, 1973), Bayesian Information Criterion (BIC) (Schwarz, 1978), and Hannan-Quinn information criterion (HQ) (Hannan and Quinn, 1979) given by

$$\begin{aligned}
\text{AIC} &= -2\ell(\hat{\boldsymbol{\theta}}) + 2p,\\
\text{BIC} &= -2\ell(\hat{\boldsymbol{\theta}}) + p\log(n), \text{ and}\\
\text{HQ} &= 2\ell(\hat{\boldsymbol{\theta}})\left(\frac{n}{n-m}\right) + p\log[\log(n)].
\end{aligned}$$

We consider the quantile residuals (Dunn and Smyth, 1996) to verify that the model provides a good fit to the data. They are defined as

$$r_i^q = \Phi^{-1}\{\text{F}(y_t|\mathcal{F}_{t-1})\},$$

where $\Phi^{-1}(\cdot)$ denotes the standard normal quantile function and $\text{F}(\cdot)$ is the cdf given in (2). Quantile residuals can detect a lack of fit in models, and their distribution is approximately normal, with a mean equal to zero and unit variance. The model provides a good fit if the residual indices plot shows no trend or standard.

According to the time series prediction theory of an ARMA (Brockwell and Davis, 2009) model, the conditional median predictions of a BXII-ARMA model can be obtained as follows. Let h_0 be the forecast horizon, we assume that the values of the covariate x_t are available or obtainable for $t = n+1\ \ldots, n+h_0$. To obtain the estimates $\hat{\mu}_{m+1}, \ldots, \hat{\mu}_n$ for the conditional median μ_t considering the CMLE $\hat{\boldsymbol{\delta}}$ we need to recompose the error term $\{r_t\}_{t=1}^n$, which will be denoted by $\hat{r}_t$. We start by defining $\hat{r}_t = \mathbb{E}(r_t)$, for $t \in \{1, \ldots, m\}$, which is usually equal to 0. Starting at $t = m+1$, we sequentially define

$$\hat{\mu}_t = g^{-1}\left\{\hat{\alpha} + x_t^\top\hat{\boldsymbol{\beta}} + \sum_{i=1}^{p}\hat{\phi}_i\left[g(y_{t-i}) - x_{t-i}^\top\hat{\boldsymbol{\beta}}\right] + \sum_{j=1}^{q}\hat{\theta}_j\hat{r}_{t-j}\right\},$$

where $\hat{r}_t = g(y_t) - g(\hat{\mu}_t)$ for $t \in \{m+1, \ldots, n\}$. The predict values of the μ_{n+h}, where $h = 1, 2, \ldots, h_0$, are defined sequentially by

$$\hat{\mu}_{n+h} = g^{-1}\left\{\hat{\alpha} + x_{n+h}^{\top}\hat{\boldsymbol{\beta}} + \sum_{i=1}^{p}\hat{\phi}_i\left[g(y_{n+h-i}) - x_{n+h-i}^{\top}\hat{\boldsymbol{\beta}}\right] + \sum_{j=1}^{q}\hat{\theta}_j\hat{r}_{n+h-j}\right\},$$

where $\hat{r}_t = 0$, for $t > n$ and

$$g(y_t) = \begin{cases} g(\hat{\mu}_t) & \text{if} \quad t > n, \\ g(y_t) & \text{if} \quad t \leq n. \end{cases}$$

To compare the predicted values of the BXII-ARMA model with other known models, some accuracy measures are also defined in this section. Among them, we can mention the Mean Square Error (MSE), Mean Absolute Percentage Error (MAPE), and Mean Absolute Scaled Error (MASE) measures. For more information on univariate time series accuracy measures, see Hyndman and Koehler (2006). These measurements are obtained through the difference between the observed and predicted values. Therefore, the lower these measures, the better the model has performance, so the model that performs better in the accuracy measures is the model that best represents the series in question. The MSE, MAPE, and MASE measures can be expressed as

$$\text{MSE} = \frac{1}{h_0}\sum_{h=1}^{h_0}(y_h - \hat{\mu}_h)^2,$$

$$\text{MAPE} = \frac{1}{h_0}\sum_{h=1}^{h_0}\frac{|y_h - \hat{\mu}_h|}{|y_h|} \quad \text{and}$$

$$\text{MASE} = \frac{1}{h_0}\sum_{h=1}^{h_0}\frac{|y_h - \hat{\mu}_h|}{\frac{1}{h-1}\sum_{h=2}^{h_0}|y_h - y_{h-1}|},$$

respectively, where y_h are the observed values, and $\hat{\mu}_h$ are the predicted values for the forecast horizon ($h = 1, \ldots, h_0$). Lower values of these measures indicate more accurate predictions.

5. Monte Carlo simulation

In this section, a Monte Carlo study is carried out to evaluate the performance of CMLEs of the BXII-ARMA model. $R = 10{,}000$ replicas and sample sizes n are considered, with $n \in \{70, 150, 300, 500\}$. The value of $\tau = 0.5$ is fixed so that the parameter μ represents the median of the distribution. The simulation study is implemented in *software* `R` (Team, 2019), and we use the `optim()` function to maximize the log-likelihood with the BFGS algorithm for numerical optimization.

The mean of the estimates, the relative bias (RB%), and the mean squared error (MSE) are used as a measure to evaluate the performance of the CMLEs. The coverage rate (CR95%) of the confidence intervals is estimated, considering a nominal confidence level of 95% and using the inverse of Fisher's information matrix to obtain the standard errors.

In Table 1, we have the results for the Monte Carlo simulation considering different structures of the BXII-ARMA model, namely: ARMA(1,1), ARMA(1,0) or AR(1) and ARMA(0,1) or MA(1). We observe that the means of the estimates approximate the true values of the parameters across all the different data generation mechanisms. In all structures of the BXII-ARMA model, the RB% decreases rapidly as the sample increases, and considering a sample size of 500, all RB% results are less than 1%. All CMLEs are approximately equal to zero, even in samples of size 70; hence they are considered consistent. For a sample size of 500, the CR95%. In general, all results presented are adequate. For more results considering other structures of the BXII-ARMA model, we recommend that the reader consult Appendix A.

Table 1: Numerical evidence on the performance of the CMLEs for the BXII-ARMA(p, q) model under different data-generating mechanisms setting $\tau = 0.5$.

Model	n	Measures	$\alpha = 1.0$	$\phi_1 = 0.5$	$\theta_1 = 0.2$	$c = 0.5$
BXII-ARMA(1,1)	70	Mean	1.0316	0.4715	0.2211	0.5304
		RB%	−3.1554	5.6982	−10.5693	−6.0748
		MSE	0.3483	0.0214	0.0136	0.0070
		CR95%	0.9416	0.9474	0.9004	0.9529
	150	Mean	1.0126	0.4835	0.2097	0.5129
		RB%	−1.2645	3.3099	−4.8603	−2.5894
		MSE	0.1548	0.0102	0.0108	0.0024
		CR95%	0.9466	0.9505	0.9172	0.9526
	300	Mean	1.0036	0.4897	0.2054	0.5062
		RB%	−0.3569	2.0508	−2.6901	−1.2483
		MSE	0.0748	0.0048	0.0042	0.0011
		CR95%	0.9454	0.9528	0.9265	0.9432
	500	Mean	1.0000	0.4931	0.2024	0.5039
		RB%	−0.0043	1.3785	−1.1828	−0.7810
		MSE	0.0437	0.0039	0.0032	0.0007
		CR95%	0.9434	0.9497	0.9311	0.9480
BXII-ARMA(1,0)	70	Mean	1.0191	0.4791	–	0.5246
		RB%	−1.9136	4.1827	–	−4.9223
		MSE	0.2179	0.0107	–	0.0069
		CR95%	0.9462	0.9389	–	0.9584
	150	Mean	1.0070	0.4894	–	0.5107
		RB%	−0.7047	2.1216	–	−2.1333
		MSE	0.0946	0.0045	–	0.0025
		CR95%	0.9494	0.9457	–	0.9521
	300	Mean	1.0040	0.4952	–	0.5056
		RB%	−0.3957	0.9609	–	−1.1275
		MSE	0.0473	0.0021	–	0.0011
		CR95%	0.9475	0.9446	–	0.9554
	500	Mean	1.0012	0.4972	–	0.5033
		RB%	−0.1239	0.5639	–	−0.6519
		MSE	0.0280	0.0012	–	0.0006
		CR95%	0.9480	0.9464	–	0.9496
BXII-ARMA(0,1)	70	Mean	0.9861	–	0.1925	0.5170
		RB%	1.3889	–	3.7607	−3.4029
		MSE	0.2386	–	0.0112	0.0049
		CR95%	0.9393	–	0.9255	0.9527
	150	Mean	0.9929	–	0.1968	0.5076
		RB%	0.7138	–	1.6172	−1.5136
		MSE	0.1100	–	0.0043	0.0019
		CR95%	0.9449	–	0.9370	0.9545
	300	Mean	0.9964	–	0.1987	0.5041
		RB%	0.3582	–	0.6608	−0.8234
		MSE	0.0543	–	0.0020	0.0009
		CR95%	0.9491	–	0.9447	0.9526
	500	Mean	0.9967	–	0.1993	0.5023
		RB%	0.3300	–	0.3270	−0.4567
		MSE	0.0327	–	0.0012	0.0005
		CR95%	0.9480	–	0.9463	0.9502

6. Application

In this context, the BXII-ARMA model is applied to two sets of real data, one related to the financial market and the other to meteorological data. Time series models have already been used to forecast financial market data. In the stock market, time series analysis is essential to improve accuracy when trading, becoming a strategy to stay one step ahead of your competitors. Jeong and Lee (2019) used the ARMA-GARCH model with the incorporation of hyperbolic tangent functions and used it to forecast the daily closing index of the S&P500 from January 1950 to December 2018. Korbi and Lleshaj (2020) used the ARMA model to predict the value of Albania finance lease over the period 2008 to 2020. Koh et al. (2020) forecast the opening prices of Maxis Berhad shares from January 2018 to December 2019 based on the previous 96 months. In the studies of Shah et al. (2022) a proposal was made to aggregate the ARIMA model with the Long Short-Term Memory model (LSTM) to obtain more adequate coefficients, with this, the proposal was also compared with other conventional models applied to the data of the Indian stock market. Lee and Park (2022) also applied the ARIMA and Recurrent Neural Network (RNN) models to predict the trading volume of houses for the real estate market.

Forecast studies have also been explored in the meteorological field and mainly with wind speed data, which help in the creation of new policies in the economic market of the electric grid, in addition, the forecasts provide analyzes to improve the use of wind energy. (Wang et al., 2018) used the ARMA model to forecast wind energy in the short term, based on historical data from a wind farm. Wu et al. (2022) we developed a Temporal Fusion Transformer (TFT) model and applied it to 8 different wind speed datasets, which obtained results that outperformed other methodologies. Chen et al. (2022) presented a new forecasting model for wind speed based on the combination of LSTM and neural network. Elsaraiti et al. (2019) used the ARIMA model to predict wind speed on historical data from the Chester region of Nova Scotia, Canada. In the proposal of the RARMA model, Bayer et al. (2020) studied trends in ocean winds and their monthly average speed from Yenagoa, Nigeria. Huang and Gu (2019) analyzed the time-varying standard deviation of the wind speed series of Typhoon Chan-Hom as it passed over the sea east of Xanguai by combining the ARMA-GARCH model and the first-order difference GARCH method. Wind speed is an important parameter in meteorological studies, which encompasses studies of atmospheric systems, ocean-atmospheric mechanisms, and especially for wind energy applications (Bayer et al., 2020).

Furthermore, the ARMA (Box and Jenkins, 1970), GARMA (Benjamin et al., 2003), and RARMA Bayer et al. (2020) models are also fitted for comparative purposes. The function `arima()` is used to the ARMA fit. The implementation of the GARMA and RARMA models is similar to the BXII-ARMA model and is available in the PTSR package using the `ptsr.fit()` function. For more information, see Prass et al. (2022). The application is implemented in the `R` software. Thus, the ARMA, GARMA, RARMA, and BXII-ARMA models were adjusted to choose the model that best represents the behavior of the series studied. The best fit of each model class is selected through the AIC, BIC, and HQ measures. After, the best fit of each class, the model with the best prediction is chosen. The last observations of the series were removed to obtain the MSE, MAPE, and MASE measures to choose the most accurate model. The forecast is estimated by one-step-ahead, updated by the actual value.

6.1. Time series model for finance data

This section presents an empirical application in data sets related to economics. The first data set referred to the trading volume of *Banco* Bradesco S.A. (BBD) stocks and was collected from the Yahoo Finance website (`https://finance.yahoo.com/most-active`), which provides up-to-date financial news, international market data, including stock quotes, financial reports and original content. The information corresponds from February 14, 2022, to February 10, 2023, totaling 250 observations. The trading volume corresponds to the number of stocks bought and sold in a day, and the standard unit of these assets is given in U.S. Dollars (US$). Since trading volumes are usually large-scale numbers, we divided the series by one hundred million to better visualize the results. The last thirty observations of the series were removed to obtain accuracy measures for choosing the best model.

Table 2: Descriptive analysis of BBD trading volume data

Minimum	Median	Mean	Maximum	Variance
0.1261	0.3317	0.3556	1.4854	0.0245

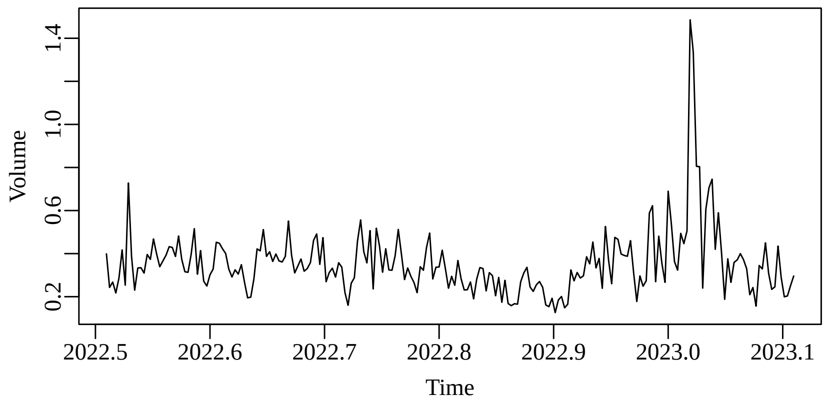
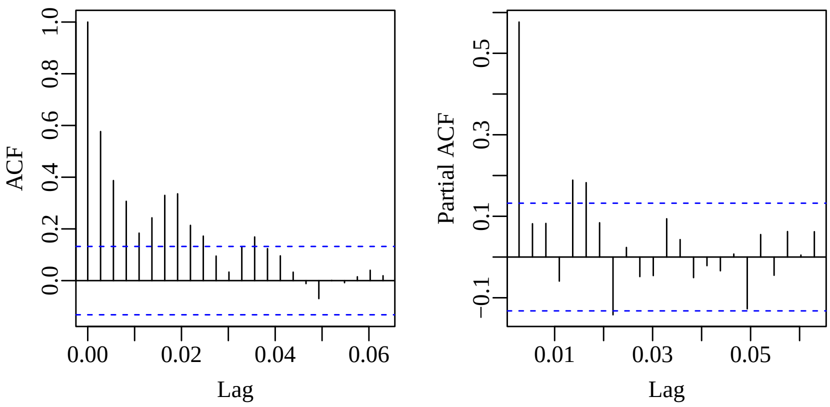


Figure 2: Time series plot of BBD trading volume

Table 2 presents some descriptive statistics of the series in question. On the original scale of the series, the average BBD trading volume is \$35.56 million and the median volume corresponds to \$33.17 million. The minimum and maximum trading volume is \$12.61 and \$148.54 million respectively. The change in trading volume is 0.02. Figure 2 shows the graphs of the BBD trading volume series. Figure 3 shows the autocorrelation function (ACF) and partial autocorrelation (PACF) plots of the BBD trading volume. A positive correlation is observed in the ACF (Figure 3(a)), and the vast majority of observations are outside the confidence interval (CI). In the PACF, shown in Figure 3(b), almost all observations are within the CI. The series was identified as stationary through the unit root tests, Phillips-Perron (PP) and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) with p-values equal to 0.01 and 0.1, respectively.



(a) ACF (b) PACF

Figure 3: ACF and PACF of BBD trading volume

Table 3 shows the fit of the models with the parameter estimates, the standard errors (SE), and the p-values of each estimate. The structure for the BBD trading volume is given by one autoregressive coefficient and four moving average filters, namely ARMA(1,4), GARMA(1,4), RARMA(1,4), and BXII-ARMA(1,4). For the ARMA model, only the intercept was significant. For the GARMA and RARMA models, all coefficients were significant, except for the autoregressive ϕ_1 and the moving average coefficient θ_4. All coefficients of the BXII-ARMA model were significant at the 5% significance level. Table 4 presents the adequacy measures adopted as methodologies for comparing forecast performance between the different best-fitted models in each class. Therefore, the MSE, MAPE, and MASE measurements are calculated for all models, considering the thirty-day out-of-sample estimates with the last thirty observations taken from the series. Therefore, the best results of the accurate measurements are given by the BXII-ARMA model for the data sets.

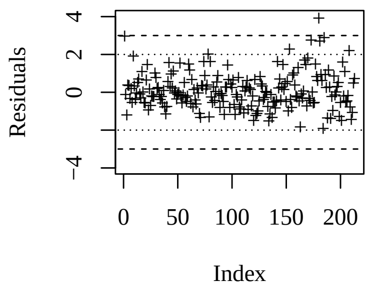
Table 3: ARMA, GARMA, RARMA, and BXII-ARMA adjustments in the time series.

ARMA(1,4)				GARMA(1,4)			
Coef.	Estimate	SE	p-value	Coef.	Estimate	SE	p-value
Int.	0.3558	0.0189	0.0000	α	−1.0887	0.1188	0.0000
ϕ_1	−0.2537	2.4091	0.9161	ϕ_1	−0.0280	0.0971	0.7730
θ_1	0.8051	2.4058	0.7379	θ_1	1.4578	0.1434	0.0000
θ_2	0.5003	1.3250	0.7057	θ_2	1.0380	0.2281	0.0000
θ_3	0.4213	0.8648	0.6261	θ_3	1.0413	0.1716	0.0000
θ_4	0.0881	0.7846	0.9106	θ_4	−0.0078	0.1746	0.9640
–	–	–	–	φ	11.5539	1.0861	0.0000
RARMA(1,4)				**BXII-ARMA(1,4)**			
Coef.	Estimate	SE	p-value	Coef.	Estimate	SE	p-value
α	−1.3357	0.1799	0.0000	α	−0.6286	0.0880	0.0000
ϕ_1	0.3683	0.5293	0.4866	ϕ_7	0.3398	0.0663	0.0000
θ_1	1.2574	0.3980	0.0016	θ_1	0.3705	0.0541	0.0000
θ_2	1.0528	0.4046	0.0093	θ_2	0.2454	0.0651	0.0002
θ_3	1.1475	0.3057	0.0002	θ_3	0.2801	0.0611	0.0000
θ_4	−0.0185	0.3283	0.9549	θ_6	0.2361	0.0626	0.0002
–	–	–	–	c	3.7013	0.1806	0.0000

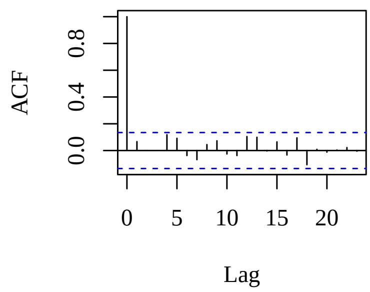
Table 4: Forecasting performance comparison among different best-fitted models in each class.

Model	MSE	MAPE	MASE
ARMA	0.0555	0.4233	0.8224
GARMA	0.0567	0.4129	0.8215
RARMA	0.0676	0.3834	0.8679
BXII-ARMA	0.0071	0.1491	0.3020

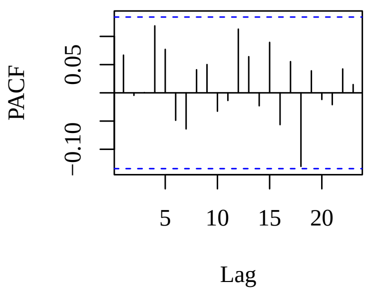
Figure 4 shows the residuals BXII-MA(1,4) plots to the BBD trading volume. In Figure 4(a), the quantile residuals show the absence of the tendency, and their behavior is similar to white noise. Figures 4(b) and 4(c), have the ACF and PACF of the residuals that confirm that they are similar to white noise. In Figure 4(d), it is concluded that the quantile residuals follow a standard normal distribution. The result of the Ljung-Box test showed a p-value of 0.5321, so the quantile residuals are not autocorrelated. Figure 5 provides a plot of the actual and adjusted values. Figure 5 shows the observed and adjusted values of the BBD trading volume. All analyzed graphs show that BXII-ARMA models adjusted to BBD trading volume can be used to make step-ahead predictions of the considered sample.



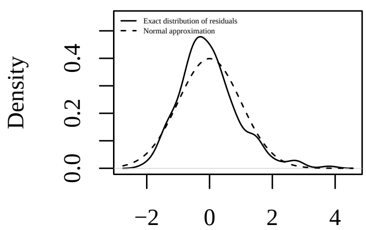
(a) Residuals



(b) ACF Residual



(c) PACF Residual



(d) Density of quantile residuals

Figure 4: Residual diagnostic plots of the fitted BXII-ARMA(1,4) model for the trading volume of BBD.

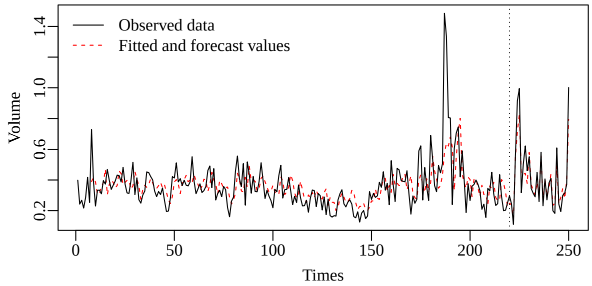


Figure 5: Observed and adjusted values of the BXII-ARMA model

6.2. Time series model for meteorological data

This section presents a real data application related to meteorology. The second data set is the average monthly wind speed for Yenagoa, Nigeria. It is reported in Amadi (2018). The series corresponds to January 2013 to December 2017, totaling 60 observations. The last twelve months of the series were removed to obtain accuracy measures for choosing the best model. The wind speed measurement is given in meters per second (m/s), and the series is divided by ten for better visualization of the results.

Table 5: Descriptive analysis of Yenagoa wind speed data

Minimum	Median	Mean	Maximum	Variance
0.0614	0.1589	0.1786	0.3392	0.0064

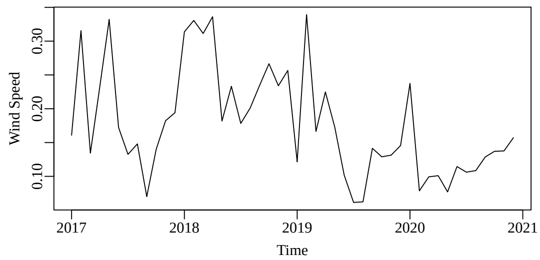
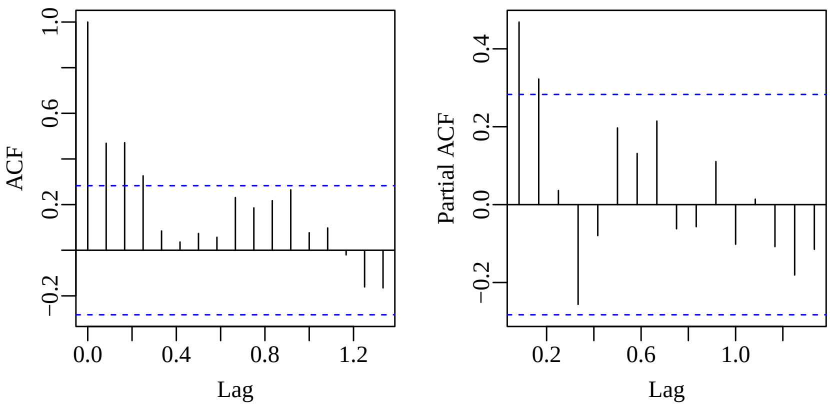


Figure 6: Time series plot of Yenagoa wind speed

Table 5 presents some descriptive statistics of the series in question. The Yenagoa wind speed data, the series average is 1.79 m/s. The median velocity corresponds to 1.59 m/s. The minimum and maximum wind speed is 0.61 m/s and 3.39 m/s, respectively, and the variance is 0.06 m/s. Figure 6 shows the graphs of the Yenagoa wind speed series. Figure 7 brings the ACF and PACF of the Calabar Wind Speed. The ACF (Figure 7(a)) shows a positive correlation until lag 15. From lag 16 onwards, we perceive a negative correlation, and it is noticed that some observations are outside the CI. In the sample PACF, Figure 7(b), only one observation is outside the CI. The series was identified as stationary through the PP and KPSS tests with p-values equal to 0.01 and 0.55, respectively.



(a) ACF (b) PACF

Figure 7: ACF and PACF of Yenagoa wind speed

Table 6 shows the structure of the models adjusted for the Yenagoa wind speed corresponds to one autoregressive coefficient and two moving average filters, that is, ARMA(1,2), GARMA(1,2), RARMA(1,2) and BXII-ARMA(1,2). All coefficients are significant at the 5% significance level, except the θ_1 of the BXII-ARMA(1,2) model adjusted, which was significant at 10%. Table 7 presents the adequacy measures adopted as methodologies for comparing forecast performance between the different best-fitted models in each class. Therefore, the MSE, MAPE, and MASE measurements are calculated for all models, considering the twelve-month out-of-sample estimates with the last twelve observations that were taken from the series. Therefore, the best results of the accurate measurements are given by the BXII-ARMA model for the data sets.

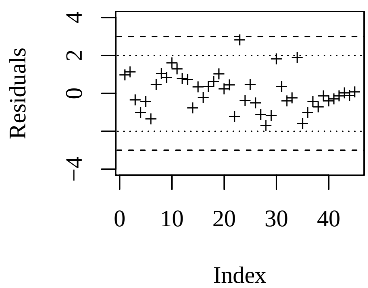
Table 6: ARMA, GARMA, RARMA, and BXII-ARMA adjustments in the time series.

ARMA(1,2)				GARMA(1,2)			
Coef.	Estimate	SE	p-value	Coef.	Estimate	SE	p-value
Int.	0.5738	0.1633	0.0004	α	−0.7345	0.2637	0.0053
ϕ_1	−0.3515	0.1654	0.0336	ϕ_1	0.5567	0.1407	0.0001
θ_1	0.4817	0.1545	0.0018	θ_1	−1.5956	0.7105	0.0247
θ_2	0.1797	0.0232	0.0000	θ_2	2.9015	0.7833	0.0002
–	–	–	–	φ	8.3038	1.6621	0.0000
RARMA(1,2)				**BXII-ARMA(1,2)**			
Coef.	Estimate	SE	p-value	Coef.	Estimate	SE	p-value
α	0.2149	−2.3516	0.0000	α	−1.1128	0.2482	0.0000
ϕ_1	3.0436	1.1281	0.0070	ϕ_3	0.3384	0.1279	0.0081
θ_1	0.9512	−2.1907	0.0213	θ_1	0.3060	0.1605	0.0565
θ_2	3.4043	1.0684	0.0014	θ_2	0.3997	0.1136	0.0004
–	–	–	–	c	3.2151	0.3643	0.0000

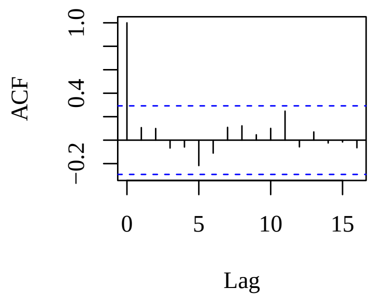
Table 7: Forecasting performance comparison among different best-fitted models in each class.

Model	MSE	MAPE	MASE
ARMA	0.0036	0.4728	1.3922
GARMA	0.0042	0.5285	1.5398
RARMA	0.0028	0.3499	1.1174
BXII-ARMA	0.0008	0.1809	0.5698

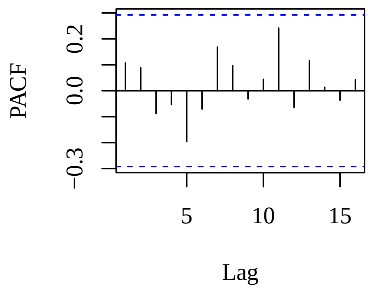
In Figure 8, the diagnostic analysis of the residuals of the BXII-ARMA(1,2) model adjusted to the Yenagoa wind speed. In Figure 8(a), the quantile residuals are arranged randomly, and their behavior is similar to white noise. Figures 8(b) and 8(c) show the residual ACF and PACF, where all lags are inside the IC. In Figure 8(d), it is concluded that the quantile residuals follow a standard normal distribution. The result of the Ljung-Box test showed a *p*-value of 0.7412, so the quantile residuals are not autocorrelated. Figure 9 provides a plot of the actual and adjusted values. Figure 9, we have the observed and adjusted values of the Yenagoa wind speed. All analyzed graphs show that BXII-ARMA models adjusted to Yenagoa wind speed can be used to make step-ahead predictions of the considered sample.



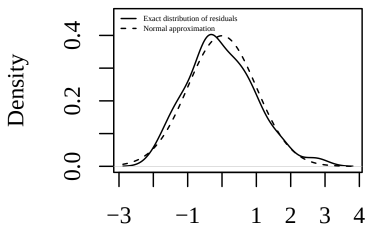
(a) Residuals



(b) ACF Residual



(c) PACF Residual



(d) Density of quantile residuals

Figure 8: Residual diagnostic plots of the fitted BXII-ARMA(1,2) model for the Yenagoa wind speed.

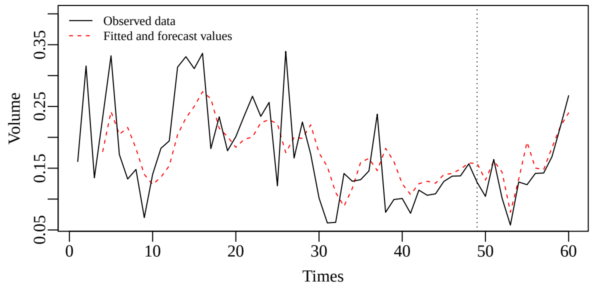


Figure 9: Observed and adjusted values of the BXII-ARMA model

7. Concluding remarks

This study proposes an autoregressive moving averages model based on a quantile parameterization of the Burr XII distribution. We carried out an economic and meteorological application using the trading volume of *Banco* Bradesco stock and the Yenagoa wind speed to verify the adequacy of the proposed model. The proposed model is compared with other models of time series to verify its goodness of fit to the dataset used. We show some results for the mathematical properties of the proposed model, such as conditional moments and probability density, cumulative distribution, and conditional quantile functions. In addition, we estimated the model's parameters by the maximum likelihood method and derived Fisher's conditional information matrix. The Monte Carlo results under different data generation mechanisms show that the maximum likelihood method is very efficient in estimating the parameters of the BXII-ARMA model. We also discuss diagnostic tools for the BXII-ARMA model predictions and model selection criteria. In the application, two data sets were considered. The first refers to the BBD trading volume and BXII-ARMA(1,4) was fitted, with one autoregressive coefficient and four moving average filters with positive effects. The second application was performed with wind speed data from Yenagoa, Nigeria, and a BXII-ARMA(1,2) model was fitted, with one autoregressive coefficient and two moving average filters with positive effects. The BXII-ARMA models fitted in both applications and obtaining predicted values close to the real values of the series also outperform the ARMA and GARMA models in the MSE, MAPE, and MASE measurements, being a useful and flexible alternative for the adjustment of non-negative time series and with asymmetric patterns. Thus, the BXII-ARMA model is suitable to satisfactorily capture the dynamics of *Banco* Bradesco's trading volume data and the Yenagoa wind speed. It should be noted that the BXII-ARMA is a new methodology whose applicability can be extended to other topics.

Appendix A.

Let Y_t be a random variable with conditional pdf in (1), using the results of Watkins (1997), we know that

$$
\begin{aligned}
\mathrm{E}\left[\log(1+Y_t^c)|\mathcal{F}_{t-1}\right] &= -h_t^{-1},\\
\mathrm{E}\left[\frac{Y_t^c\log(Y_t)}{1+Y_t^c}|\mathcal{F}_{t-1}\right] &= \frac{1-\gamma-\psi\,(-h_t)}{c\,(1-h_t)} \quad \text{and} \\
\mathrm{E}\left[\frac{Y_t^c\log^2(Y_t)}{(1+Y_t^c)^2}|\mathcal{F}_{t-1}\right] &= \frac{\frac{\pi^2}{6}+\gamma^2-2\gamma+2(\gamma-1)\psi\,(1-h_t)+\psi^2\,(1-h_t)+\psi'\,(1-h_t)}{c^2\,(1-h_t)\,(2-h_t)\,(-h_t)^{-1}}.
\end{aligned}
\tag{A.1}
$$

Table A.8: Further numerical evidence on the performance of the CMLEs for the BXII-ARMA(p, q) model under different data-generating mechanisms setting $\tau = 0.25$.

Model	n	Measure	$\alpha = -1$	$\phi_1 = 0.1$	$\phi_2 = -0.4$	$\theta_1 = 0.3$	$\theta_2 = 0.2$	$c = 3$
BXII-ARMA(2,2)	70	Mean	−1.0337	0.1584	0.2974	0.2396	0.2639	3.1746
		RB%	−3.3750	−58.4001	25.6505	20.1368	−31.9363	-5.8196
		MSE	0.1156	0.0938	0.0815	0.0978	0.0451	0.1242
		CR%	0.9027	0.8494	0.8344	0.8358	0.8410	0.9227
	150	Mean	−1.0093	0.1225	0.3618	0.2770	0.2216	3.0740
		RB%	−0.9349	−22.4711	9.5411	7.6753	−10.7921	−2.4655
		MSE	0.0400	0.0284	0.0242	0.0296	0.0147	0.0428
		CR%	0.9313	0.9193	0.9178	0.9140	0.9088	0.9375
	300	Mean	−1.0040	0.1091	0.3838	0.2907	0.2092	3.0358
		RB%	−0.4033	−9.1181	4.0566	3.1098	−4.5933	−1.1945
		MSE	0.0177	0.0103	0.0088	0.0110	0.0062	0.0186
		CR%	0.9439	0.9416	0.9389	0.9380	0.9403	0.9442
	500	Mean	−1.0026	0.1043	0.3911	0.2958	0.2042	3.0212
		RB%	−0.2639	−4.2792	2.2146	1.3957	−2.1049	−0.7066
		MSE	0.0104	0.0054	0.0048	0.0058	0.0035	0.0107
		CR%	0.9432	0.9446	0.9452	0.9426	0.9458	0.9464
BXII-ARMA(2,0)	70	Mean	−1.0588	0.0876	0.3693	–	–	3.1125
		RB%	−5.8830	12.3973	7.6849	–	–	−3.7513
		MSE	0.0442	0.0092	0.0097	–	–	0.1021
		CR%	0.9491	0.9404	0.9343	–	–	0.9438
	150	Mean	−1.0241	0.0954	0.3858	–	–	3.0582
		RB%	−2.4074	4.6068	3.5468	–	–	−1.9385
		MSE	0.0182	0.0038	0.0040	–	–	0.0409
		CR%	0.9466	0.9482	0.9413	–	–	0.9460
	300	Mean	−1.0074	0.0989	0.3958	–	–	3.0149
		RB%	−0.7370	1.0980	1.0607	–	–	−0.4957
		MSE	0.0049	0.0011	0.0011	–	–	0.0105
		CR%	0.9491	0.9472	0.9476	–	–	0.9485
	500	Mean	−1.0129	0.0985	0.3922	–	–	3.0232
		RB%	−1.2914	1.4892	1.9532	–	–	−0.7750
		MSE	0.0085	0.0018	0.0019	–	–	0.0182
		CR%	0.9476	0.9469	0.9426	–	–	0.9444
BXII-ARMA(0,2)	70	Mean	−0.9795	–	–	0.2937	0.1926	3.1049
		RB%	2.0530	–	–	2.1141	3.6910	−3.4961
		MSE	0.0124	–	–	0.0099	0.0112	0.0931
		CR%	0.9190	–	–	0.9369	0.9261	0.9441
	150	Mean	−0.9929	–	–	0.2983	0.1973	3.0414
		RB%	0.7078	–	–	0.5642	1.3636	−1.3816
		MSE	0.0056	–	–	0.0040	0.0042	0.0363
		CR%	0.9367	–	–	0.9439	0.9413	0.9483
	300	Mean	−0.9953	–	–	0.2983	0.1987	3.0228
		RB%	0.4721	–	–	0.5805	0.6378	−0.7611
		MSE	0.0028	–	–	0.0019	0.0019	0.0172
		CR%	0.9449	–	–	0.9467	0.9457	0.9458
	500	Mean	−0.9968	–	–	0.2989	0.1993	3.0147
		RB%	0.3227	–	–	0.3659	0.3741	−0.4889
		MSE	0.0017	–	–	0.0011	0.0011	0.0102
		CR%	0.9428	–	–	0.9502	0.9487	0.9467

Table A.9: Further numerical evidence on the performance of the CMLEs for the BXII-ARMA(p, q) model under different data-generating mechanisms setting $\tau = 0.5$.

Model	n	Measure	$\alpha = -1$	$\phi_1 = 0.1$	$\phi_2 = -0.4$	$\theta_1 = 0.3$	$\theta_2 = 0.2$	$c = 3$
BXII-ARMA(2,2)	70	Mean	−1.1362	0.1320	0.3060	0.2651	0.2683	3.1767
		RB%	−13.6190	−32.0296	23.5113	11.6226	−34.1643	−5.8915
		MSE	0.2348	0.0822	0.0709	0.0877	0.0446	0.1308
		CR%	0.9040	0.8560	0.8498	0.8417	0.8409	0.9296
	150	Mean	−1.0481	0.1144	0.3634	0.2855	0.2237	3.0752
		RB%	−4.8117	−14.4078	9.1426	4.8431	−11.8274	−2.5066
		MSE	0.0733	0.0280	0.0239	0.0291	0.0154	0.0456
		CR%	0.9384	0.9165	0.9130	0.9134	0.9063	0.9426
	300	Mean	−1.0233	0.1060	0.3835	0.2940	0.2109	3.0388
		RB%	−2.3293	−5.9751	4.1337	2.0138	−5.4489	−1.2936
		MSE	0.0327	0.0102	0.0090	0.0108	0.0066	0.0206
		CR%	0.9454	0.9409	0.9376	0.9394	0.9293	0.9414
	500	Mean	−1.0104	0.1035	0.3920	0.2966	0.2056	3.0230
		RB%	−1.0414	−3.4815	1.9888	1.1232	−2.8004	−0.7665
		MSE	0.0185	0.0055	0.0048	0.0058	0.0037	0.0116
		CR%	0.9448	0.9443	0.9451	0.9425	0.9410	0.9476
BXII-ARMA(2,0)	70	Mean	−1.0858	0.0899	0.3699	−	−	3.1176
		RB%	−8.5769	10.1215	7.5321	−	−	−3.9205
		MSE	0.0765	0.0090	0.0096	−	−	0.1048
		CR%	0.9412	0.9415	0.9350	−	−	0.9475
	150	Mean	−1.0389	0.0960	0.3860	−	−	3.0550
		RB%	−3.8871	3.9711	3.4914	−	−	−1.8337
		MSE	0.0311	0.0038	0.0039	−	−	0.0438
		CR%	0.9475	0.9483	0.9449	−	−	0.9438
	300	Mean	−1.0172	0.0983	0.3935	−	−	3.0278
		RB%	−1.7205	1.6738	1.6131	−	−	−0.9257
		MSE	0.0145	0.0018	0.0018	−	−	0.0197
		CR%	0.9476	0.9459	0.9436	−	−	0.9446
	500	Mean	−1.0111	0.0989	0.3959	−	−	3.0148
		RB%	−1.1070	1.0531	1.0172	−	−	−0.4923
		MSE	0.0085	0.0010	0.0011	−	−	0.0111
		CR%	0.9480	0.9531	0.9476	−	−	0.9499
BXII-ARMA(0,2)	70	Mean	−1.0013	−	−	0.3004	0.2017	3.1068
		RB%	−0.1261	−	−	−0.1435	−0.8616	−3.5598
		MSE	0.0055	−	−	0.0113	0.0124	0.0938
		CR%	0.9333	−	−	0.9337	0.9221	0.9450
	150	Mean	−1.0003	−	−	0.3000	0.2010	3.0476
		RB%	−0.0289	−	−	0.0093	−0.5052	−1.5870
		MSE	0.0025	−	−	0.0044	0.0047	0.0363
		CR%	0.9455	−	−	0.9455	0.9378	0.9494
	300	Mean	−1.0006	−	−	0.3005	0.2007	3.0240
		RB%	−0.0592	−	−	−0.1635	−0.3351	−0.8002
		MSE	0.0013	−	−	0.0021	0.0021	0.0169
		CR%	0.9473	−	−	0.9449	0.9436	0.9514
	500	Mean	−1.0002	−	−	0.3001	0.2001	3.0147
		RB%	−0.0186	−	−	−0.0243	−0.0453	−0.4892
		MSE	0.0008	−	−	0.0012	0.0012	0.0100
		CR%	0.9464	−	−	0.9446	0.9479	0.9510

Table A.10: Further numerical evidence on the performance of the CMLEs for the BXII-ARMA(p, q) model under different data-generating mechanisms setting $\tau = 0.75$.

Model	n	Measure	$\alpha = -1$	$\phi_1 = 0.1$	$\phi_2 = -0.4$	$\theta_1 = 0.3$	$\theta_2 = 0.2$	$c = 3$
BXII-ARMA(2,2)	70	Mean	−1.2169	0.1514	0.2770	0.2432	0.2847	3.1800
		RB%	−21.6892	−51.4123	30.7596	18.9194	−42.3303	−5.9999
		MSE	0.3836	0.1062	0.0867	0.1116	0.0472	0.1358
		CR%	0.8933	0.8282	0.8169	0.8156	0.8218	0.9286
	150	Mean	−1.0917	0.1146	0.3557	0.2833	0.2291	3.0791
		RB%	−9.1724	−14.6417	11.0761	5.5716	−14.5421	−2.6371
		MSE	0.1101	0.0286	0.0252	0.0301	0.0154	0.0477
		CR%	0.9430	0.9116	0.9062	0.9109	0.9012	0.9385
	300	Mean	−1.0425	0.1037	0.3825	0.2952	0.2118	3.0379
		RB%	−4.2543	−3.6934	4.3859	1.6032	−5.9016	−1.2632
		MSE	0.0484	0.0103	0.0093	0.0108	0.0066	0.0209
		CR%	0.9498	0.9413	0.9396	0.9413	0.9332	0.9434
	500	Mean	−1.0238	0.1040	0.3885	0.2957	0.2071	3.0233
		RB%	−2.3820	−3.9875	2.8852	1.4260	−3.5693	−0.7755
		MSE	0.0280	0.0058	0.0051	0.0060	0.0037	0.0118
		CR%	0.9435	0.9427	0.9453	0.9445	0.9397	0.9484
BXII-ARMA(2,0)	70	Mean	−1.1102	0.0902	0.3698	–	–	3.1144
		RB%	−11.0249	9.7660	7.5470	–	–	−3.8124
		MSE	0.1091	0.0092	0.0092	–	–	0.1032
		CR%	0.9346	0.9365	0.9352	–	–	0.9457
	150	Mean	−1.0483	0.0955	0.3870	–	–	3.0535
		RB%	−4.8255	4.4873	3.2494	–	–	−1.7834
		MSE	0.0437	0.0037	0.0038	–	–	0.0417
		CR%	0.9475	0.9483	0.9468	–	–	0.9507
	300	Mean	−1.0219	0.0978	0.3944	–	–	3.0293
		RB%	−2.1880	2.2273	1.4054	–	–	−0.9767
		MSE	0.0206	0.0018	0.0018	–	–	0.0198
		CR%	0.9470	0.9496	0.9477	–	–	0.9482
	500	Mean	−1.0136	0.0994	0.3957	–	–	3.0151
		RB%	−1.3602	0.5858	1.0836	–	–	−0.5030
		MSE	0.0120	0.0011	0.0011	–	–	0.0113
		CR%	0.9505	0.9469	0.9476	–	–	0.9490
BXII-ARMA(0,2)	70	Mean	−1.0162	–	–	0.2946	0.1974	3.1146
		RB%	−1.6198	–	–	1.7945	1.2868	−3.8205
		MSE	0.0066	–	–	0.0111	0.0123	0.1000
		CR%	0.9173	–	–	0.9277	0.9218	0.9430
	150	Mean	−1.0066	–	–	0.2999	0.2005	3.0496
		RB%	−0.6626	–	–	0.0428	−0.2747	−1.6519
		MSE	0.0028	–	–	0.0044	0.0046	0.0384
		CR%	0.9391	–	–	0.9437	0.9379	0.9489
	300	Mean	−1.0034	–	–	0.2992	0.1995	3.0259
		RB%	−0.3414	–	–	0.2688	0.2738	−0.8643
		MSE	0.0014	–	–	0.0021	0.0021	0.0183
		CR%	0.9407	–	–	0.9489	0.9483	0.9492
	*500	Mean	−1.0021	–	–	0.2999	0.1997	3.0152
		RB%	−0.2051	–	–	0.0232	0.1322	−0.5059
		MSE	0.0008	–	–	0.0012	0.0012	0.0108
		CR%	0.9480	–	–	0.9494	0.9491	0.9458

References

Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle. Springer , 267–281.

Amadi, H.N., 2018. Wind energy potential assessment of coastal states in south-south nigeria based on the Weibull distribution model. European Journal of Electrical Engineering and Computer Science 2, 1–6.

Araújo, F.J.M., Guerra, R.R., Peña-Ramírez, F.A., 2022. The Burr XII quantile regression for salary-performance models with applications in the sports economy. Computational and Applied Mathematics In press, In press.

Bayer, F.M., Bayer, D.M., Marinoni, A., Gamba, P., 2020. A novel rayleigh dynamical model for remote sensing data interpretation. IEEE Transactions on Geoscience and Remote Sensing 58, 4989–4999.

Bayer, F.M., Bayer, D.M., Pumi, G., 2017. Kumaraswamy autoregressive moving average models for double bounded environmental data. Journal of Hydrology 555, 385–396.

Benjamin, M.A., Rigby, R.A., Stasinopoulos, D.M., 2003. Generalized autoregressive moving average models. Journal of the American Statistical Association 98, 214–223.

Bhatti, F.A., Hamedani, G.G., Korkmaz, M.Ç., Sheng, W., Ali, A., 2021. On the Burr XII-moment exponential distribution. Plos one 16, 1–21.

Box, G.E., Jenkins, G.M., 1970. Time series analysis forecasting and control. Wisconsin Univ Madison Dept of Statistics , 1–574.

Brockwell, P.J., Davis, R.A., 2009. Time series: theory and methods. Springer Science & Business Media , 1–576.

Burr, I.W., 1942. Cumulative frequency functions. The Annals of Mathematical Statistics 13, 215–232.

Chen, G., Tang, B., Zeng, X., Zhou, P., Kang, P., Long, H., 2022. Short-term wind speed forecasting based on long short-term memory and improved bp neural network. International Journal of Electrical Power & Energy Systems 134, 107365.

Cribari-Neto, F., Scher, V.T., Bayer, F.M., 2023. Beta autoregressive moving average model selection with application to modeling and forecasting stored hydroelectric energy. International Journal of Forecasting 39, 98–109.

Davison, A., Hinkley, D., 1997. Bootstrap methods and their application. Cambridge University Press 1, 1–556.

Dhongde, S., Minoiu, C., 2013. Global poverty estimates: A sensitivity analysis. World Development 44, 1–13.

Dunn, P.K., Smyth, G.K., 1996. Randomized quantile residuals. Journal of Computational and Graphical Statistics 5, 236–244.

Elsaraiti, M., Merabet, A., Al-Durra, A., 2019. Time series analysis and forecasting of wind speed data, in: 2019 IEEE Industry Applications Society Annual Meeting, IEEE. pp. 1–5.

Guerra, R.R., Peña-Ramírez, F.A., Cordeiro, G.M., 2021. The weibull burr xii distribution in lifetime and income analysis. Anais da Academia Brasileira de Ciências 93.

Hannan, E.J., Quinn, B.G., 1979. The determination of the order of an autoregression. Journal of the Royal Statistical Society: Series B (Methodological) 41, 190–195.

Huang, Z., Gu, M., 2019. Characterizing nonstationary wind speed using the arma-garch model. Journal of Structural Engineering 145, 04018226.

Hyndman, R.J., Koehler, A.B., 2006. Another look at measures of forecast accuracy. International journal of forecasting 22, 679–688.

Jeong, Y., Lee, S., 2019. Recurrent neural network-adapted nonlinear arma-garch model with application to s&p 500 index data. The Korean Data & Information Science Society 30, 1187–1195.

Koh, W.S., Heng, H.S., Wong, C.Z., Lai, P.L., Dass, C., 2020. Forecasting stock price using arma model. INTI JOURNAL 2020, 1–8.

Korbi, A., Lleshaj, L., 2020. Finance leasing and arma forecasting: Evidence from albania. European Journal of Business and Management Research 5.

Lee, C., Park, K.K.H., 2022. Forecasting trading volume in local housing markets through a time-series model and a deep learning algorithm. Engineering, Construction and Architectural Management 29, 165–178.

Low, V.J.M., Khoo, H.L., Khoo, W.C., 2021. Statistical modelling of bus travel time with burr distribution. ITM Web of Conferences 36, 1–15.

Melchior, C., Zanini, R.R., Guerra, R.R., Rockenbach, D.A., 2021. Forecasting brazilian mortality rates due to occupational accidents using autoregressive moving average approaches. International Journal of Forecasting 37, 825–837.

Muse, A.H., Mwalili, S., Ngesa, O., Almalki, S.J., Abd-Elmougod, G.A., 2021. Bayesian and classical inference for the generalized log-logistic distribution with applications to survival data. Computational intelligence and neuroscience 2021, 1–24.

Pawitan, Y., 2001. In all likelihood: statistical modelling and inference using likelihood. Oxford University Press , 1–528.

Prass, T.S., Carlos, J.H., Taufemback, C.G., Pumi, G., 2022. Positive time series regression models. arXiv preprint arXiv:2201.03667 , 1–11.

Ramires, T.G., Nakamura, L.R., Righetto, A.J., Carvalho, R.J., Vieira, L.A., Pereira, C.A., 2021. Comparison between highly complex location models and gamlss. Entropy 23, 469.

Rasaki, O.O., Ezekiel, O., Olumide, A.L., Adigun, O.A., 2018. On skew generalized extreme value-arma model: An application to average monthly temperature (1901-2016) in nigeria. International Journal of Statistics and Applied Mathematics 3, 20–27.

Ribeiro, T.F., Cordeiro, G.M., Peña-Ramírez, F.A., Guerra, R.R., 2021. A new quantile regression for the COVID-19 mortality rates in the United States. Computational and Applied Mathematics 40, 1–16.

Ribeiro, T.F., Peña-Ramírez, F.A., Guerra, R.R., Cordeiro, G.M., 2022. Another unit burr xii quantile regression model based on the different reparameterization applied to dropout in brazilian undergraduate courses. Plos one 17, e0276695.

Ruszczynski, A., 2011. Nonlinear optimization. Princeton university press 108, 1–764.

Sagrillo, M., Guerra, R.R., Machado, R., Bayer, F.M., 2023. A generalized control chart for anomaly detection in sar imagery. Computers & Industrial Engineering 177, 109030.

Schwarz, G., 1978. Estimating the dimension of a model. The annals of statistics 6, 461–464.

Shah, H., Bhatt, V., Shah, J., 2022. A neoteric technique using arima-lstm for time series analysis on stock market forecasting, in: Mathematical Modeling, Computational Intelligence Techniques and Renewable Energy: Proceedings of the Second International Conference, MMCITRE 2021, Springer. pp. 381–392.

Singh, S.K., Maddala, G.S., 1976. A function for the size distribution of incomes. Econometrica 44, 963– 970.

Tadikamalla, P.R., 1980. A look at the Burr and related distribution. International Statistical Review/Revue Internationale de Statistique 48, 337–344.

Team, R.C., 2019. R: A language and environment for statistical computing. System 2, 1–13.

Thompson, B.S., 2013. Empirical likelihood-based inference for poverty measures with relative poverty lines. Econometric Reviews 32, 513–523.

Wald, A., 1943. Tests of statistical hypotheses concerning several parameters when the number of observations is large. Transactions of the American Mathematical society 54, 426–482.

Wang, J., Zhou, Q., Zhang, X., 2018. Wind power forecasting based on time series arma model, in: IOP Conference Series: Earth and Environmental Science, IOP Publishing. p. 022015.

Watkins, A., 1997. A note on expected Fisher information for the Burr XII distribution. Microelectron. Reliability 37, 1849–1852.

Wu, B., Wang, L., Zeng, Y.R., 2022. Interpretable wind speed prediction with multivariate time series and temporal fusion transformers. Energy 252, 123990.

Zimmer, W.J., Keats, J.B., Wang, F.K., 1998. The Burr XII distribution in reliability analysis. Journal of Quality Technology 30, 386–394.