APPLICATION NOTE

Estimation of the parameters of symmetric stable ARMA and ARMA–GARCH models

Aastha M. Sathe and N. S. Upadhye
Department of Mathematics, Indian Institute of Technology Madras, Chennai, India

ABSTRACT
In this article, we first propose the modified Hannan–Rissanen Method for estimating the parameters of autoregressive moving average (ARMA) process with symmetric stable noise and symmetric stable generalized autoregressive conditional heteroskedastic (GARCH) noise. Next, we propose the modified empirical characteristic function method for the estimation of GARCH parameters with symmetric stable noise. Further, we show the efficiency, accuracy and simplicity of our methods with Monte-Carlo simulation. Finally, we apply our proposed methods to model the financial data.

ARTICLE HISTORY
Received 2 December 2020
Accepted 5 May 2021

KEYWORDS
ARMA–GARCH models; stable distributions; parameter estimation; simulation

1. Introduction

In finance and econometrics, important information about past market movements is modelled through the conditional distribution of return data series using the autoregressive moving average (ARMA) models. However, in these models, the conditional distribution is assumed to be homoskedastic which poses challenges in the modelling and analysis of returns with time-varying volatility and volatility clustering, commonly observed phenomena. This led to the development of autoregressive conditional heteroskedastic (ARCH) and generalized ARCH (GARCH) models introduced by Engle and colleagues [4] and Bollerslev [1], respectively. In empirical finance, the most important and widely used model is the combination of ARMA with GARCH, referred to as an ARMA–GARCH model (see [19,20]).

In practice, the noise term of ARMA, GARCH and ARMA–GARCH models is assumed to be normal or Student-\(t\) distribution, where the normal distribution is known for the desirable property of stability but fails to capture the heavy-tails of the data. On the other hand, the Student-\(t\) distribution allows for heavier tails, but, when compared with the normal, lacks the desirable property of stability. These flaws, therefore, motivate us to use the family of stable distributions as a model for unconditional, conditional homoskedastic and conditional heteroskedastic return series distribution. Some of the significant and attractive features of stable distributions, apart from stability are heavy-tails, leptokurtic...
shape, domains of attraction and skewness. For more details on stable distributions, see [16].

For the parameter estimation of the above-mentioned models, only a few estimation techniques are available. The details on the different estimation techniques used for such models, see [3,5,11,12]. Therefore, we believe that this article is an important contribution to the literature available on these models.

In this article, we develop a method for estimating the parameters of ARMA model with symmetric stable noise and symmetric stable GARCH noise by modifying the Hannan–Rissanen Method (HR) [8]. For the estimation of the symmetric stable GARCH parameters, we use the model discussed in [15] as opposed to the standard GARCH model with symmetric stable noise and tempered stable noise used by Calzolari et al. [3] and Feng and Shi [5], respectively. The estimation technique used in [3] is the indirect inference approach (requires Student’s-t distributed noise, that assumes the existence of the second moment as the auxiliary model) while Feng and Shi use the maximum likelihood estimation technique. We propose the modified empirical characteristic function method (MECF), which is simple and direct and is based on the method discussed in [14]. The efficiency and effectiveness of the two proposed methods are validated through Monte Carlo simulation. For comparative analysis, we compare the modified Hannan–Rissanen (MHR) method with the HR method along with three specific M-estimators introduced by Cadler and Davis [2], namely the least absolute deviation (LAD) estimator; the least-squares (LS) estimator, and the maximum likelihood (MLE) estimator for ARMA with symmetric stable noise and heavy-tailed noise (finite variance). Amongst the three M-estimators, LAD is an excellent choice for estimating ARMA parameters with stable noise due to its robustness, computational simplicity and good asymptotic properties. Through simulations, we observed that there is not much decline in the performance of the MHR method when the noise is finite variance. Also, the simulated tables showcase that the method works at par with the LAD estimator both in terms of accuracy and computational simplicity.

On the other hand, for the estimation of ARMA–GARCH with symmetric stable noise, we compare our proposed methods with the numerical implementation of maximum likelihood technique discussed by Wuertz et al. [18]. Both the proposed methods are efficient for $\alpha \geq 1.5$.

The paper is organized as follows. Section 2 gives a brief introduction to the stable distributions along with the necessary definitions and notations. Section 3 discusses the two new methods proposed for the estimation of ARMA, GARCH and ARMA–GARCH parameters. Section 4 deals with simulations and comparative analysis of the proposed methods. Section 5 discusses an application of our proposed methods to the financial data. Finally, Section 6 gives some concluding remarks on the proposed methods.

2. Preliminaries and notations

2.1. Stable distributions

These distributions form a rich class of heavy-tailed distributions, introduced by Lévy [10], in his study on the Generalized Central Limit Theorem. Each distribution, in this class, is characterized by four parameters, namely $\alpha$, $\beta$, $\sigma$, and $\delta$, which, respectively, denote the index of stability, skewness, scale and shift of the distribution. Their respective ranges are
given by $\alpha \in (0,2], \beta \in [-1,1], \sigma > 0$ and $\delta \in \mathbb{R}$. For more details, see [17]. In this paper, we deal with symmetric $\alpha$-stable distributions, denoted by $S\alpha S$. We say that $X \sim S\alpha S$ if and only if $\beta = \delta = 0$. The characteristic function (cf) representation of $X \sim S\alpha S$ is given by

$$\phi(t) = \exp\{-(\sigma |t|)^{\alpha}\}. \quad (1)$$

Note that when $\alpha = 2$, $\beta = 0$ and $\delta = 0$, the distribution of $X$ is Gaussian, i.e. $X \sim N(0,2\sigma^2)$.

Next, we discuss some important definitions and notations required for parameter estimation of ARMA, GARCH and ARMA–GARCH models with $S\alpha S$ noise, $\alpha > 1$. The assumption $\alpha > 1$ is required for the mean of $S\alpha S$ to be finite. However, the assumption $\alpha > 1$ is not very restrictive in the context of financial modelling, because most financial time series of interest appear to have finite means. $\phi = (\phi_1, \ldots, \phi_p)$ and $\theta = (\theta_1, \ldots, \theta_q)$ denote the parameter vectors used in the ARMA model.

Let the time series $\{X_t : t \in \mathbb{Z}\}$ be a (mean-corrected) ARMA($p, q$) process with $S\alpha S$ noise, denoted by $S\alpha S$-ARMA($p, q$), if

$$X_t - \sum_{i=1}^{p} \phi_i X_{t-i} = \sum_{j=1}^{q} \theta_j \epsilon_{t-j} + \epsilon_t, \quad (2)$$

let $\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p$ is the autoregressive (AR) polynomial, $\theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q$ is the moving average (MA) polynomial, $p$ denotes the order of autoregression, $q$ denotes the order of moving average and $\{\epsilon_t\}$ denotes the noise sequence of iid random variables with $S\alpha S$ distribution with (cf) given in (1) for $\alpha \in (1,2]$ and $\sigma > 0$. Further, we assume that in the unit disc $\{z \in \mathbb{C} : |z| \leq 1\}$, $\phi(z)$ and $\theta(z)$ have no common zeros. For more details on ARMA($p, q$) process with $S\alpha S$ noise, (see [16, pp. 376–380]).

Let the time series $\{X_t : t \in \mathbb{Z}\}$ be a (mean-corrected) GARCH($p, q$) process with $S\alpha S$ noise denoted by $S\alpha S$-GARCH($p, q$), if

$$X_t = \sigma_t' \epsilon_t, \quad \sigma_t' = c + \sum_{i=1}^{q} a_i |X_{t-i}| + \sum_{j=1}^{p} b_j \sigma_{t-j}', \quad (3)$$

where $c > 0$, $a_i, i = 1, \ldots, q$ and $b_j, j = 1 \ldots p$ are non-negative constants and $\{\epsilon_t\}$ denotes the noise sequence of iid random variables with $S\alpha S$ distribution with (cf) given in (1).

We say that the time series $\{X_t : t \in \mathbb{Z}\}$ is a (mean-corrected) ARMA($p_A, q_A$) process with $S\alpha S$-GARCH($p_G, q_G$) noise, denoted by ARMA($p_A, q_A$)-$S\alpha S$-GARCH($p_G, q_G$), if

$$X_t = \sum_{i=1}^{p_A} \phi_i X_{t-i} + \sum_{j=1}^{q_A} \theta_j \epsilon_{t-j} + \epsilon_t, \quad \sigma_t' = c + \sum_{i=1}^{q_G} a_i |\epsilon_{t-i}| + \sum_{j=1}^{p_G} b_j \sigma_{t-j}', \quad (4)$$

where the autoregressive (AR) polynomial $\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^{p_A}$ and the moving average (MA) polynomial $\theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^{q_A}$ have no common roots in the unit disc $\{z \in \mathbb{C} : |z| \leq 1\}, c > 0, a_i, i = 1, \ldots, q_G$ and $b_j, j = 1 \ldots p_G$ are non-negative constants and $\{\epsilon_t\}$ denotes the noise sequence of iid random variables with $S\alpha S$ distribution with (cf) given in (1) for $\alpha \in (1,2]$ and $\sigma > 0$. 
2.2. Measures of dependence for stable processes

Note now that due to undefined covariance when \( \alpha < 2 \), the classical autocovariance function cannot be considered as a tool for developing methods of parameter estimation of the processes defined in (2) and (4). In such a case, alternative measures of dependence should be used. A few known choices are: normalized autocovariation, autocodifference and fractional lower order covariance (FLOC). For details, see [7, Subsection 2.1]. Amongst these choices, for the estimation of the processes defined in (2) and (4), we prefer to use normalized autocovariation due to its simple formulation and computational efficiency.

The covariation of \( X \) on \( Y \) can be expressed as

\[
CV(X, Y) = p \frac{\mathbb{E}(XY^{p-1})}{\mathbb{E}|Y|^p} \sigma^g_Y, \quad 1 \leq p < \alpha,
\]

(5)

where \( \sigma_Y \) is the scale parameter of the random variable \( Y \) and the signed power is defined as \( y^{(p-1)} = |y^{p-1}| \text{sign}(y), y \in \mathbb{C} \). Thus, the autocovariation function of the stationary \( S\alpha S \) process \( \{X_t\} \) with \( \alpha > 1 \) for lag \( k, k = 0, \pm 1, \pm 2 \ldots \) is defined as the covariation of random variables \( X_t \) and \( X_{t-k} \) with \( p = 1 \).

In our proposed method of estimating the parameters of the processes defined in (2) and (4), we make use of the normalized autocovariation. The normalized autocovariation (see [6]) for \( S\alpha S \) process \( \{X_t\} \) with lag \( k, k = 0, \pm 1, \pm 2 \ldots \) is given by

\[
NCV(X_t, X_{t-k}) = \frac{\mathbb{E}(X_t \text{sign}(X_{t-k}))}{\mathbb{E}|X_{t-k}|} \frac{CV(X_t, X_{t-k})}{\sigma_{X_{t-k}}^g}
\]

(6)

and its estimator for a sample \( x_1, \ldots, x_N \) being a realization of a stationary process \( \{X_t\} \) is given by

\[
\hat{NCV}(X_t, X_{t-k}) = \frac{\sum_{t=l}^{r} x_t \text{sign}(x_{t-k})}{\sum_{t=1}^{N} |x_t|},
\]

(7)

where \( l = \max(1, 1 + k) \) and \( r = \min(N, N + k) \). Refer [6] to study the main properties of the estimator.

3. Parameter estimation

We propose two methods for the estimation of the parameters \( \phi, \theta, c, a, b, a_G \) and \( b_G \) for the processes defined in (2), (3) and (4), namely the modified Hannan–Rissanen method (MHR) and the modified empirical characteristic function method (MECF).

3.1. Modified Hannan–Rissanen method (MHR)

The classical Hannan–Rissanen method [8] uses linear regression to establish estimates for the parameters and the white noise variance of an ARMA(\( p, q \)) process. However, to establish estimates for processes with infinite variance, we modify the Hannan–Rissanen method [8]. Let \( \{X_t\} \) be defined as in (2) or (4). Firstly, we fit a high order AR(\( a \)) model to \( \{X_t\} \) (with \( a > \max(p, q) \)) using the modified Yule-Walker estimation method, for \( S\alpha S \) autoregressive models with \( \alpha \in (1, 2] \), see [9] as follows:
Let \( \{X_t\} \) be an AR(\( a \)) process with \( a > \max(p, q) \) given by

\[
X_t - \phi_1 X_{t-1} - \cdots - \phi_a X_{t-a} = \epsilon_t,
\]

where \( \{\epsilon_t\} \) constitutes sample of an iid \( \alpha \) random variables with \( \alpha > 1 \).

Multiply both sides of (8) by the vector \( S = [S_{t-1}, \ldots, S_{t-a}] \) where \( S_t = \text{sign}(X_t) \) and take the expectation to obtain \( a \) equations of the form

\[
E X_t S_{t-j} - \sum_{i=1}^{a} \phi_i E X_{t-i} S_{t-j} = E \epsilon_t = 0, \quad j = 1, \ldots, a.
\]

Divide the \( j \)th equation of (9) by \( E |X_{t-j}| \) to obtain

\[
\frac{E X_t S_{t-j}}{E |X_{t-j}|} - \sum_{i=1}^{a} \phi_i \frac{E X_{t-i} S_{t-j}}{E |X_{t-j}|} = 0, \quad j = 1, \ldots, a.
\]

Applying normalized autocovariation as given in (6) to (10), we obtain the following matrix form:

\[
\lambda = \Lambda \Phi,
\]

where \( \lambda \) and \( \Phi \) are vectors of length \( a \) defined as

\[
\lambda = [\text{NCV}(X_t, X_{t-1}), \ldots, \text{NCV}(X_t, X_{t-a})]' \quad \Phi = [\phi_1, \ldots, \phi_a]'
\]

and the matrix \( \Lambda \) has size \( a \times a \) with the \((i,j)\)th elements given by

\[
\Lambda(i,j) = \text{NCV}(X_t, X_{t-i+j}).
\]

Finally, in order to estimate the values of the parameter vector \( \Phi \), the matrix \( \Lambda \), \( \hat{\Lambda} \) should be nonsingular and make use of the sample normalized autocovariation \( \hat{\text{NCV}} \). Thus,

\[
\hat{\Phi} = \hat{\Lambda}^{-1} \hat{\lambda}.
\]

Next, using the estimated coefficients of AR(\( a \)) model, we compute the estimated residuals. Finally, we use least absolute regression to obtain the estimates of the parameters of the considered given model.

The algorithm is as follows:

1. **Step 1.** Assume (8) holds and estimate the values of parameter vector \( \Phi \) using (11).
2. **Step 2.** Using the obtained estimated coefficients \( \hat{\phi}_{a1}, \ldots, \hat{\phi}_{aa} \), compute the estimated residuals from

\[
\hat{\epsilon}_t = X_t - \hat{\phi}_{a1} X_{t-1} - \cdots - \hat{\phi}_{aa} X_{t-a}, \quad t = a + 1, \ldots, n.
\]
3. **Step 3.** Next, we estimate the vector of parameters, \( \beta = (\phi, \theta) \) by least absolute deviation regression of \( X_t \) onto \( (X_{t-1}, \ldots, X_{t-p}, \hat{\epsilon}_{t-1}, \ldots, \hat{\epsilon}_{t-q}), t = a + 1 + q, \ldots, n \).
by minimizing $S(\beta)$ with respect to $\beta$ where,

$$S(\beta) = \sum_{t=a+1+q}^{n} |X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} - \theta_1 \hat{\epsilon}_{t-1} - \cdots - \theta_q \hat{\epsilon}_{t-q}|.$$  \hspace{1cm} (12)

**Remark:** The above-proposed algorithm cannot be used if $\Lambda$ and $\hat{\Lambda}$ are singular.

### 3.2. Modified empirical characteristic function method (MECF)

We employ this method to obtain the estimates of $a$, $b$, and $c$ for processes defined in (3) and (4). The proposed method is based on the method discussed in [14]. For simplicity, we shall discuss the method for the process defined in (3) (see [13], for details), for the case $p = 1$ and $q = 1$ relevant in empirical finance. The method can be extended to higher orders of $p$ and $q$ in a similar spirit. For $p = 1$ and $q = 1$, we get

$$X_t = \sigma'_t \epsilon_t, \quad \sigma'_t = c + a_1 |X_{t-1}| + b_1 \sigma'_{t-1},$$  \hspace{1cm} (13)

where $c > 0$, $a_1$ and $b_1$ are non-negative satisfying the stationary condition $\lambda a_1 + b_1 < 1$, where $\lambda = \mathbb{E} |\epsilon_1|$ (refer [15] for different values of $\lambda$) and $\{\epsilon_t\}$ denotes the noise sequence of iid random variables with $S\alpha S$ distribution, $\alpha > 1$. The parameter estimates for (13) are obtained by minimizing the function $f$ over $c$, $a_1$, and $b_1$ defined as

$$f(c, a_1, b_1) = \sum_{j=1}^{n} |\psi_{\text{theoretical}}(\hat{\epsilon}_j) - \psi_{\text{empirical}}(\hat{\epsilon}_j)|,$$

where $\hat{\epsilon}_j = \frac{X_j}{c + a_1 |X_{j-1}| + b_1 \sigma'_{j-1}}$, $\psi_{\text{theoretical}}(\hat{\epsilon}_j) = \exp(-|\hat{\epsilon}_j|^\hat{\alpha})$, $\psi_{\text{empirical}}(\hat{\epsilon}_j) = \frac{1}{n} \sum_{j=1}^{n} \cos(\hat{\epsilon}_j Y_j)$ where $Y_1, \ldots, Y_n$ are iid random variables with $S\alpha S$ distribution and $\hat{\alpha}$ is the estimate of $\alpha$ obtained from $\hat{\epsilon}_j$ under the assumption that it is $S\alpha S$ distributed using the hybrid method discussed in [17]. We make use of the function ‘nlminb’ available in ‘stats’ package of R for the minimization of the function $f$ over $c$, $a_1$, and $b_1$.

### 4. Simulation and comparative analysis

In this section, we investigate and compare the performance of the proposed MHR method with the HR method available in ‘ismr’ package in R and the three M-estimators, namely least absolute deviation (LAD); the least-squares (LS) and the maximum likelihood estimation (MLE), introduced by Cadler and Davis [2] developed for the estimation of the parameters of ARMA models with heavy-tailed (Students’s-t) or $S\alpha S$ noise. Note that MHR uses LAD to obtain the parameter estimates with slight modification in (12). For the estimation of ARMA–GARCH models with $S\alpha S$ noise, we compare our proposed method (MHR+MECF) with the maximum likelihood estimation (MLE) discussed in [18] available in ‘GEVStableGarch’ package in R.

Next, we consider two models (M1) and (M2) to check the relative performance of the proposed estimator (MHR) in comparison to different estimators when the noise sequence $\{\epsilon_t\}$ is assumed to be $S\alpha S$, Student’s-t or Gaussian through Monte Carlo simulations, namely
Further, we consider (M3) and (M4) models for the comparative analysis of MLE with our proposed (MHR+MECF) method, both, developed for ARMA models with SαS-GARCH noise.

(M3) MA(1)-SαS-GARCH(1,1): \( X_t = \theta_1 \epsilon_{t-1} + \epsilon_t' \), \( \epsilon_t' = \sigma_t' \epsilon_t \) with \( \sigma_t' = c + a_1 |X_{t-1}| + b_1 \sigma_{t-1} \).

(M4) ARMA(1,1)-SαS-GARCH(1,1): \( X_t = \phi_1 X_{t-1} + \theta_1 \epsilon_{t-1} + \epsilon_t' \), \( \epsilon_t' = \sigma_t' \epsilon_t \) with \( \sigma_t' = c + a_1 |X_{t-1}| + b_1 \sigma_{t-1} \).

When noise sequence \( \{\epsilon_t\} \) is considered to be iid SαS, \( \alpha \in (1, 2) \) and \( \beta, \sigma \) and \( \delta \) are fixed to 0, 1, 0, respectively, and is generated using the function ‘rstable’ in the ‘stabledist’ package of R. For different values of \( n, \alpha \in [1.5, 2) \), \( \theta_1 \in (0, 0.5] \) and \( \phi_1 \in (0, 1) \), simulations are run where 1000 realizations of each model are generated. Similar procedure is followed when \( \{\epsilon_t\} \) is assumed to be Student’s-\( t \) and Gaussian and generated using the function ‘rt’.

| (\( \theta_1 \), \( n \), \( \alpha_{\text{stable}} \)) | Method | SαS noise | Gaussian noise | Student’s-\( t \) noise |
|----------------|--------|---------|--------------|----------------|
| (0.1, 100, 1.5) | HR     | 0.079 (0.128) | 0.085 (0.117) | 0.094 (0.125) |
|                 | MHR    | 0.091 (0.078) | 0.076 (0.130) | 0.101 (0.134) |
|                 | LAD    | 0.092 (0.077) | 0.089 (0.126) | 0.098 (0.127) |
|                 | LS     | 0.087 (0.086) | 0.086 (0.108) | 0.095 (0.110) |
|                 | MLE    | 0.104 (0.075) | 0.099 (0.111) | 0.106 (0.122) |
| (0.1, 400, 1.5) | HR     | 0.101 (0.088) | 0.095 (0.048) | 0.106 (0.061) |
|                 | MHR    | 0.096 (0.032) | 0.093 (0.060) | 0.109 (0.062) |
|                 | LAD    | 0.099 (0.070) | 0.090 (0.100) | 0.110 (0.061) |
|                 | LS     | 0.100 (0.086) | 0.091 (0.093) | 0.109 (0.050) |
|                 | MLE    | 0.099 (0.063) | 0.092 (0.102) | 0.105 (0.067) |
| (0.5, 200, 1.65) | HR     | 0.499 (0.078) | 0.493 (0.046) | 0.494 (0.077) |
|                 | MHR    | 0.476 (0.072) | 0.483 (0.090) | 0.481 (0.075) |
|                 | LAD    | 0.487 (0.121) | 0.494 (0.139) | 0.496 (0.072) |
|                 | LS     | 0.505 (0.190) | 0.492 (0.123) | 0.497 (0.070) |
|                 | MLE    | 0.490 (0.122) | 0.507 (0.135) | 0.498 (0.073) |
| (0.5, 500, 1.65) | HR     | 0.497 (0.039) | 0.501 (0.046) | 0.498 (0.053) |
|                 | MHR    | 0.485 (0.031) | 0.489 (0.056) | 0.489 (0.051) |
|                 | LAD    | 0.498 (0.039) | 0.492 (0.059) | 0.500 (0.050) |
|                 | LS     | 0.493 (0.050) | 0.499 (0.050) | 0.500 (0.050) |
|                 | MLE    | 0.500 (0.033) | 0.504 (0.044) | 0.500 (0.049) |
| (0.3, 300, 1.85) | HR     | 0.298 (0.057) | 0.296 (0.063) | 0.302 (0.059) |
|                 | MHR    | 0.290 (0.058) | 0.300 (0.074) | 0.300 (0.059) |
|                 | LAD    | 0.294 (0.089) | 0.300 (0.124) | 0.305 (0.060) |
|                 | LS     | 0.300 (0.095) | 0.294 (0.111) | 0.303 (0.055) |
|                 | MLE    | 0.295 (0.084) | 0.300 (0.121) | 0.306 (0.059) |
| (0.3, 700, 1.85) | HR     | 0.291 (0.040) | 0.309 (0.037) | 0.299 (0.038) |
|                 | MHR    | 0.284 (0.037) | 0.306 (0.043) | 0.293 (0.041) |
|                 | LAD    | 0.279 (0.064) | 0.320 (0.076) | 0.295 (0.040) |
|                 | LS     | 0.280 (0.076) | 0.321 (0.072) | 0.299 (0.037) |
|                 | MLE    | 0.279 (0.066) | 0.323 (0.077) | 0.298 (0.041) |

Note: Here, \( \alpha_{\text{stable}} \) values are used only for SαS noise.
Table 2. Mean and RMSE (in parentheses) of 1000 estimates of $\phi_1$ and $\theta_1$ for M2 model for different values of $\alpha$ and $n$.

| $(\phi_1, \theta_1, n, \alpha_{\text{stable}})$ | Method | $\text{S\alpha S noise}$ | | $\text{Gaussian noise}$ | | $\text{Student’s-t noise}$ | |
|---|---|---|---|---|---|---|
| | Mean ($\hat{\phi}_1, \hat{\theta}_1$) | RMSE ($\hat{\phi}_1, \hat{\theta}_1$) | Mean ($\hat{\phi}_1, \hat{\theta}_1$) | RMSE ($\hat{\phi}_1, \hat{\theta}_1$) | Mean ($\hat{\phi}_1, \hat{\theta}_1$) | RMSE ($\hat{\phi}_1, \hat{\theta}_1$) | |
| $(0.4, 0.1, 100, 1.5)$ | HR | (0.410, 0.097) | (0.152, 0.216) | (0.418, 0.060) | (0.146, 0.190) | (0.429, 0.034) | (0.178, 0.223) | |
| | MHR | (0.389, 0.106) | (0.118, 0.140) | (0.387, 0.096) | (0.265, 0.297) | (0.390, 0.092) | (0.178, 0.223) | |
| | LAD | (0.389, 0.106) | (0.145, 0.142) | (0.370, 0.123) | (0.201, 0.227) | (0.373, 0.111) | (0.237, 0.247) | |
| | LS | (0.389, 0.106) | (0.192, 0.182) | (0.362, 0.122) | (0.172, 0.179) | (0.364, 0.110) | (0.222, 0.241) | |
| | MLE | (0.389, 0.106) | (0.117, 0.137) | (0.409, 0.079) | (0.205, 0.227) | (0.398, 0.100) | (0.233, 0.255) | |
| $(0.4, 0.1, 400, 1.5)$ | HR | (0.402, 0.096) | (0.056, 0.137) | (0.399, 0.085) | (0.161, 0.113) | (0.404, 0.084) | (0.090, 0.096) | |
| | MHR | (0.396, 0.103) | (0.053, 0.057) | (0.377, 0.105) | (0.202, 0.144) | (0.383, 0.106) | (0.097, 0.103) | |
| | LAD | (0.381, 0.122) | (0.047, 0.051) | (0.361, 0.109) | (0.288, 0.256) | (0.379, 0.113) | (0.099, 0.104) | |
| | LS | (0.383, 0.122) | (0.063, 0.073) | (0.361, 0.111) | (0.272, 0.231) | (0.378, 0.114) | (0.102, 0.103) | |
| | MLE | (0.398, 0.110) | (0.069, 0.075) | (0.377, 0.098) | (0.270, 0.242) | (0.381, 0.114) | (0.099, 0.101) | |
| $(0.2, 0.4, 100, 1.7)$ | HR | (0.272, 0.301) | (0.174, 0.206) | (0.284, 0.304) | (0.176, 0.214) | (0.309, 0.294) | (0.176, 0.220) | |
| | MHR | (0.195, 0.382) | (0.193, 0.222) | (0.186, 0.385) | (0.241, 0.237) | (0.226, 0.363) | (0.230, 0.256) | |
| | LAD | (0.181, 0.422) | (0.168, 0.181) | (0.185, 0.387) | (0.252, 0.255) | (0.180, 0.419) | (0.226, 0.235) | |
| | LS | (0.164, 0.431) | (0.183, 0.179) | (0.195, 0.382) | (0.221, 0.213) | (0.156, 0.420) | (0.182, 0.184) | |
| | MLE | (0.174, 0.436) | (0.164, 0.177) | (0.214, 0.374) | (0.216, 0.210) | (0.161, 0.430) | (0.225, 0.228) | |
| $(0.2, 0.4, 600, 1.7)$ | HR | (0.257, 0.349) | (0.072, 0.070) | (0.231, 0.366) | (0.087, 0.083) | (0.234, 0.367) | (0.087, 0.086) | |
| | MHR | (0.231, 0.377) | (0.077, 0.075) | (0.201, 0.396) | (0.100, 0.095) | (0.205, 0.389) | (0.085, 0.081) | |
| | LAD | (0.209, 0.400) | (0.071, 0.066) | (0.192, 0.405) | (0.095, 0.085) | (0.198, 0.401) | (0.078, 0.077) | |
| | LS | (0.222, 0.376) | (0.072, 0.069) | (0.197, 0.400) | (0.075, 0.072) | (0.198, 0.403) | (0.079, 0.075) | |
| | MLE | (0.235, 0.385) | (0.063, 0.057) | (0.197, 0.400) | (0.089, 0.083) | (0.206, 0.392) | (0.087, 0.086) | |

Note: Here, $\alpha_{\text{stable}}$ values are used only for SoS noise.

and ‘rnorm’ in the ‘stats’ package of R. Tables 1, 2, 3 and 4 showcase the performance of different estimators for models (M1), (M2), (M3) and (M4), respectively, in terms of mean and root mean squared error (RMSE).

From Table 1, we observe that the MHR method works efficiently when the noise is SoS in comparison to other estimators and slightly less efficient than HR if the noise is either standard Gaussian or standard Student’s-t.

From Table 2, we observe that HR method estimates $\phi_1 = 0.2$ and $\theta_1 = 0.4$ inaccurately for different values of $n$ and noise sequence. The MHR method works almost at par with other estimators when the noise is SoS and standard Student’s-t and slightly less efficient when the noise is Gaussian.

Thus, in general, we observe that the finite variance noise does not cause a great decline in the performance of MHR method. Also, amongst the three M-estimators, LAD is an excellent choice for estimating ARMA parameters with heavy-tailed noise, both, in terms of computational simplicity and accuracy.

From Tables 3 and 4, we observe that for different values of $n$, the (MHR+MECF) method is almost at par with the (MLE) method in obtaining the estimate of $(\theta_1, c, a_1, b_1)$ and $(\theta_1, \phi_1, c, a_1, b_1)$ for model (M3) and (M4), respectively, when the noise is SoS and Gaussian. The estimate of $\alpha$ for both the models is obtained from the SoS noise sequence $\{\epsilon_t\}$ using the hybrid method.

We observe that for all the four models (M1), (M2), (M3) and (M4), the proposed methods are at par with other estimators in terms of the robustness and accuracy of the estimates.
Table 3. Mean and RMSE (in parentheses) of 1000 estimates of $\theta_1$, $c$, $a_1$ and $b_1$ for M3 model for different values of $\alpha$ and $n$.

| $(n, \theta_1, c, a_1, b_1, \alpha_{\text{stable}})$ | Method     | $S\alpha$ noise |                          |                          | $S\alpha$ noise |                          |                          |
|-----------------------------------------------|------------|------------------|--------------------------|--------------------------|------------------|--------------------------|--------------------------|
| $(300, 0.1, 0.001, 0.1, 0.6, 1.75)$          | MHR+MECF  | (0.066, 0.0008, 0.097, 0.468) | (0.078, 0.0007, 0.068, 0.274) | (0.097, 0.0006, 0.091, 0.374) | (0.060, 0.0007, 0.086, 0.356) |
|                                              | MLE        | (0.076, 0.0013, 0.100, 0.508) | (0.070, 0.0008, 0.051, 0.293) | (0.104, 0.0011, 0.104, 0.581) | (0.061, 0.0007, 0.049, 0.275) |
| $(500, 0.1, 0.001, 0.1, 0.6, 1.75)$          | MHR+MECF  | (0.099, 0.0006, 0.064, 0.393) | (0.066, 0.0008, 0.064, 0.352) | (0.086, 0.0005, 0.075, 0.419) | (0.058, 0.0007, 0.069, 0.329) |
|                                              | MLE        | (0.096, 0.0012, 0.105, 0.532) | (0.047, 0.0007, 0.040, 0.191) | (0.094, 0.0016, 0.106, 0.466) | (0.042, 0.0008, 0.039, 0.266) |
| $(200, 0.3, 0.01, 0.05, 0.8, 1.85)$          | MHR+MECF  | (0.282, 0.005, 0.091, 0.589) | (0.094, 0.007, 0.125, 0.424) | (0.303, 0.006, 0.045, 0.596) | (0.104, 0.006, 0.053, 0.412) |
|                                              | MLE        | (0.277, 0.023, 0.079, 0.540) | (0.068, 0.022, 0.053, 0.387) | (0.308, 0.026, 0.067, 0.517) | (0.082, 0.027, 0.048, 0.443) |
| $(700, 0.3, 0.01, 0.05, 0.8, 1.85)$          | MHR+MECF  | (0.310, 0.007, 0.048, 0.658) | (0.058, 0.006, 0.028, 0.347) | (0.277, 0.005, 0.063, 0.542) | (0.051, 0.007, 0.119, 0.460) |
|                                              | MLE        | (0.313, 0.014, 0.068, 0.702) | (0.052, 0.010, 0.134, 0.342) | (0.283, 0.020, 0.061, 0.635) | (0.041, 0.019, 0.029, 0.304) |
| $(100, 0.5, 0.05, 0.3, 0.1, 1.65)$          | MHR+MECF  | (0.452, 0.045, 0.260, 0.137) | (0.196, 0.040, 0.171, 0.208) | (0.498, 0.034, 0.254, 0.08) | (0.172, 0.034, 0.172, 0.071) |
|                                              | MLE        | (0.523, 0.043, 0.213, 0.279) | (0.117, 0.025, 0.178, 0.394) | (0.500, 0.046, 0.288, 0.130) | (0.093, 0.016, 0.105, 0.236) |
| $(600, 0.5, 0.05, 0.3, 0.1, 1.65)$          | MHR+MECF  | (0.455, 0.036, 0.245, 0.076) | (0.084, 0.047, 0.185, 0.082) | (0.480, 0.022, 0.222, 0.102) | (0.064, 0.040, 0.183, 0.130) |
|                                              | MLE        | (0.496, 0.051, 0.261, 0.128) | (0.022, 0.009, 0.050, 0.125) | (0.497, 0.049, 0.203, 0.097) | (0.040, 0.005, 0.037, 0.081) |

Note: Here, $\alpha_{\text{stable}}$ values are used only for $S\alpha$ noise.
Table 4. Mean and RMSE (in parentheses) of 1000 estimates of $\theta_1$, $\phi_1$, $c$, $a_1$ and $b_1$ for M4 model for different values of $\alpha$ and $n$.

| $(n, \theta_1, \phi_1, c, a_1, b_1, \alpha_{\text{stable}})$ | Method       | SαS noise          | Gaussian noise       |
|------------------------------------------------------------|--------------|---------------------|----------------------|
|                                                           | Mean ($\hat{\theta}_1, \hat{\phi}_1, \hat{c}, \hat{a}_1, \hat{b}_1$) | RMSE ($\hat{\theta}_1, \hat{\phi}_1, \hat{c}, \hat{a}_1, \hat{b}_1$) | Mean ($\hat{\theta}_1, \hat{\phi}_1, \hat{c}, \hat{a}_1, \hat{b}_1$) | RMSE ($\hat{\theta}_1, \hat{\phi}_1, \hat{c}, \hat{a}_1, \hat{b}_1$) |
| (200, 0.2, 0.9, 1, 0.1, 0.5, 1.9) | MHR+MECF     | (0.201, 0.887, 1.025, 0.070, 0.391) | (0.100, 0.0296, 0.473, 0.068, 0.245) | (0.175, 0.884, 0.894, 0.094, 0.316) | (0.109, 0.042, 0.425, 0.124, 0.299) |
| MLE                                                        | (0.230, 0.884, 1.515, 0.100, 0.293) | (0.081, 0.033, 0.991, 0.076, 0.384) | (0.188, 0.883, 1.451, 0.088, 0.311) | (0.081, 0.040, 0.912, 0.068, 0.373) |
| (600, 0.2, 0.9, 1, 0.1, 0.5, 1.9) | MHR+MECF     | (0.168, 0.899, 1.011, 0.063, 0.366) | (0.061, 0.017, 0.420, 0.061, 0.253) | (0.213, 0.879, 1.075, 0.071, 0.384) | (0.062, 0.036, 0.322, 0.053, 0.244) |
| MLE                                                        | (0.184, 0.901, 1.383, 0.108, 0.342) | (0.045, 0.015, 0.659, 0.033, 0.262) | (0.207, 0.886, 1.232, 0.108, 0.390) | (0.043, 0.030, 0.520, 0.037, 0.227) |
| (100, 0.1, 0.7, 0.001, 0.1, 0.3, 1.75) | MHR+MECF     | (0.087, 0.668, 0.0007, 0.078, 0.202) | (0.144, 0.114, 0.0008, 0.071, 0.174) | (0.094, 0.658, 0.0006, 0.080, 0.188) | (0.143, 0.117, 0.0008, 0.082, 0.186) |
| MLE                                                        | (0.090, 0.679, 0.0010, 0.086, 0.307) | (0.124, 0.089, 0.0006, 0.091, 0.370) | (0.096, 0.679, 0.0009, 0.096, 0.339) | (0.124, 0.088, 0.0006, 0.098, 0.383) |
| (500, 0.1, 0.7, 0.001, 0.1, 0.3, 1.75) | MHR+MECF     | (0.113, 0.686, 0.0006, 0.052, 0.171) | (0.043, 0.054, 0.009, 0.069, 0.200) | (0.090, 0.692, 0.0006, 0.066, 0.203) | (0.068, 0.046, 0.0008, 0.063, 0.172) |
| MLE                                                        | (0.114, 0.681, 0.001, 0.103, 0.218) | (0.037, 0.061, 0.008, 0.030, 0.232) | (0.093, 0.685, 0.001, 0.094, 0.298) | (0.062, 0.036, 0.0033, 0.037, 0.246) |
| (300, 0.2, 0.6, 0.01, 0.01, 0.6, 1.85) | MHR+MECF     | (0.220, 0.575, 0.005, 0.022, 0.453) | (0.098, 0.077, 0.007, 0.046, 0.303) | (0.219, 0.575, 0.006, 0.012, 0.481) | (0.081, 0.109, 0.007, 0.015, 0.280) |
| MLE                                                        | (0.589, 0.209, 0.015, 0.019, 0.351) | (0.079, 0.065, 0.011, 0.033, 0.462) | (0.207, 0.592, 0.016, 0.015, 0.344) | (0.078, 0.065, 0.012, 0.026, 0.481) |
| (500, 0.2, 0.6, 0.01, 0.01, 0.6, 1.85) | MHR+MECF     | (0.190, 0.601, 0.008, 0.029, 0.546) | (0.074, 0.051, 0.006, 0.057, 0.196) | (0.194, 0.591, 0.005, 0.010, 0.358) | (0.072, 0.061, 0.007, 0.018, 0.392) |
| MLE                                                        | (0.183, 0.607, 0.017, 0.006, 0.310) | (0.055, 0.044, 0.012, 0.016, 0.500) | (0.215, 0.586, 0.012, 0.015, 0.488) | (0.057, 0.052, 0.011, 0.026, 0.454) |

Note: Here, $\alpha_{\text{stable}}$ values are used only for SαS noise.
Table 5. Comparison of CPU time through different methods.

| Model (true values) | Method | CPU time (in sec) |
|---------------------|--------|------------------|
| M1 \((n = 500, \theta_1 = 0.4, \alpha = 1.7)\) | LAD    | 1.27             |
|                     | MLE    | 12.04            |
|                     | LS     | 1.20             |
|                     | MHR    | 0.30             |
|                     | HR     | 1.17             |
| M2 \((n = 400, \theta_1 = 0.4, \phi_1 = 0.2, \alpha = 1.5)\) | LAD    | 1.37             |
|                     | MLE    | 20.00            |
|                     | LS     | 1.42             |
|                     | MHR    | 0.37             |
|                     | HR     | 0.80             |
| M3 \((n = 500, \theta_1 = 0.3, c = 0.001, a_1 = 0.1, b_1 = 0.3, \alpha = 1.65)\) | MHR+MECF | 0.65         |
|                     | MLE    | 2.15             |
| M4 \((n = 600, \theta_1 = 0.1, \phi_1 = 0.7, c = 0.001, a_1 = 0.1, b_1 = 0.3, \alpha = 1.75)\) | MHR+MECF | 0.89         |
|                     | MLE    | 2.45             |

4.1. Computational Burden

Table 5 showcases the time taken (in sec) to estimate the ARMA and ARMA–GARCH parameters with SαS noise by different methods. We observe that our proposed MHR and MHR+MECF method are computationally efficient in comparison to the M-estimators, HR method and MLE method, respectively. Also, the estimation done via MLE is computationally complex due to lack of a closed form expression of stable density function.

5. Application to financial data

We have considered two datasets for the empirical analysis of different GARCH models, namely

(1) The International Business Machines Corporation (IBM) obtained from Yahoo Finance for the period 19 January 2000–19 March 2005, comprising of 1297 daily log returns values of the adjusted closing price. The skewness is \(-0.030\) and kurtosis is 5.81.

(2) The dataset sp500dge (available in ‘fGarch’ package in R) that contains daily returns from the SP500 index as used in the paper of Ding et al. [4]. The sample period is from 3 January 1928 to 30 August 1991 for a total of 17,055 observations with kurtosis and skewness of 22.41 and \(-0.48\), respectively.

To determine whether the considered time series data is stationary, we implement the Augmented Dickey-Fuller Test (ADF) ‘adf.test’ available in the ‘series’ package in R. The p-value obtained is 0.01, which confirms the stationarity of the data.

We first perform the Anderson Darling test to check if the datasets come from a population with a Gaussian distribution. For both the datasets, the null hypotheses indicating normality are rejected with p-values \(< 0.05\). Also, the returns plot of both the datasets as shown in Figures 1 and 2 exhibit conditional heteroscedasticity. As a result, application of GARCH(1,1) model with non-Gaussian distributions such as Student’s-t and SαS are expected to outperform GARCH(1,1) model with normal distribution. To compare the models, we calculate the log-likelihood and the AIC values. For the two datasets, the
estimates, standard error, $t$ and $p$-values of the mean-corrected GARCH(1,1) models with different distributions along with the log-likelihood and the AIC values are presented in Tables 6–8. We observe that the GARCH(1,1) models with non-Gaussian distributions are all preferred to the GARCH(1,1) model with Gaussian distribution. Also, amongst the non-Gaussian distributions, $S\alpha S$ (MECF) performs the best. Overall, all the parameter estimates are statistically significant in all models.

**Note:** For the estimation of parameters of GARCH(1,1) model with Gaussian and Student’s-$t$, the function garchfit in ‘fGarch’ package in R is used. The estimation technique used is MLE and the following model is considered

$$X_t = \sigma_t^\prime \epsilon_t, \quad \sigma_t^2 = c + a_1 X_{t-1}^2 + b_1 \sigma_{t-1}^2.$$  (14)

On the other hand, for the estimation of GARCH(1,1) model with $S\alpha S$ distribution, the model given in (13) is considered due to infinite variance and finite mean for $\alpha > 1$ and the techniques used for estimation are: MECF and MLE (gsFit function available in ‘GEVStableGarch’ package in R).
Table 6. Empirical results: IBM.

| Parameters | Gaussian (Estimate, S.E.) | (t-value, p-value) | Student’s-t (Estimate, S.E.) | (t-value, p-value) | SU S (MECF) (Estimate, S.E.) | (t-value, p-value) | SU S (MLE) (Estimate, S.E.) | (t-value, p-value) |
|------------|---------------------------|--------------------|-----------------------------|--------------------|-----------------------------|--------------------|-----------------------------|--------------------|
| c          | (7.28e-07, 5.83e-07)      | (1.23, 0.217)      | (6.82e-07, 5.92e-07)        | (1.15, 0.24)       | (0.00015, 8.64e-05)         | (1.74, 0.08)       | (2.96e-05, 4.21e-05)        | (0.7, 0.42)        |
| a₁         | (6.16e-02, 1.47e-02)      | (4.17, 0)          | (5.36e-02, 1.71e-02)        | (3.13, 0.001)      | (5.02e-02, 1.74e-03)        | (28.64, 0)         | (5.38e-02, 1.37e-02)        | (3.9, 0)           |
| b₁         | (9.3e-01, 1.38e-02)       | (67.94, 0)         | (9.45e-01, 1.60e-02)        | (58.8, 0)          | (7.7e-01, 1.36e-02)         | (56.52, 0)         | (9.33e-01, 1.79e-02)        | (52.0, 0)          |

Table 7. Empirical results: SPS00DGE.

| Parameters | Gaussian (Estimate, S.E.) | (t-value, p-value) | Student’s-t (Estimate, S.E.) | (t-value, p-value) | SU S (MECF) (Estimate, S.E.) | (t-value, p-value) | SU S (MLE) (Estimate, S.E.) | (t-value, p-value) |
|------------|---------------------------|--------------------|-----------------------------|--------------------|-----------------------------|--------------------|-----------------------------|--------------------|
| c          | (7.36e-07, 9.22e-08)      | (8.28, 0)          | (7.03e-07, 9.59e-08)        | (7.33, 0)          | (7.29e-05, 4.53e-05)        | (1.61, 0.107)      | (4.07e-05, 6.17e-06)        | (6.59, 0)          |
| a₁         | (8.71e-02, 4.24e-03)      | (20.52, 0)         | (7.52e-02, 4.79e-03)        | (15.70, 0)         | (5.05e-02, 2.54e-03)        | (19.65, 0)         | (4.80e-02, 2.84e-03)        | (17.08, 0)         |
| b₁         | (9.10e-01, 4.19e-03)      | (216.74, 0)        | (9.20e-01, 4.64e-03)        | (198.29, 0)        | (7.58e-01, 7.53e-03)        | (99.52, 0)         | (9.36e-01, 3.78e-03)        | (247.2, 0)         |
Figure 2. Graphical Analysis: SP500DGE, where QQ plot represents the fitted stable quantiles and the empirical quantiles for the residuals, the density plot showcases the densities of various distributions using the fitted parameters and the boxplot showcases $p$-values of KS two-sample test.

Table 8. Log-Likelihood and AIC values.

| Distributions    | IBM       | SP500DGE     |
|------------------|-----------|--------------|
|                  | Log.Lik   | AIC          | Log.Lik   | AIC       |
| Gaussian         | 3370.683  | -6737.36     | 56,653    | -113,302  |
| Student’s-t      | 3405.63   | -6807.26     | 57,321    | -114,368  |
| $S\alpha S$ (MECF) | 3620.89   | -7237.78     | 59,346    | -117,868  |
| $S\alpha S$ (MLE) | 3411.767  | -6819.56     | 58,360    | -116,716  |

In order to check if the residuals can be considered as an independent sample, we make use of the empirical autocovariance function instead of the classical autocovariance (autocorrelation) function for our model. More precisely, if the considered data comes from the independent sample, then the dependence measure (empirical autocovariance) should be equal (or close) to zero for lags greater than zero. From the residuals autocovariation plot in Figures 1 and 2, one can observe that the dependence in the residual series of both the datasets is almost unidentifiable.
Next, we show that the distribution of the residual series is \( \alpha S \) using the Kolmogorov–Smirnov (KS) test as discussed in [9]. The KS test statistics is defined as

\[
D = \sup_x |G_{n_1}(x) - G_{n_2}(x)|,
\]

where \( G_{n_1}(\cdot) \) and \( G_{n_2}(\cdot) \) denote the empirical cumulative distribution function for Sample 1 of size \( n_1 \) and Sample 2 of size \( n_2 \), respectively. The test derived from the KS statistics is called the two-sample KS test. In our case, Sample 1 corresponds to a part of the residuals, i.e. \( (n_1 = n/2) \) points and Sample 2 corresponds to the random sample of size \( n_2 = 100 \), simulated from \( \alpha S \) distribution with estimated parameters corresponding to the remaining part of residuals with \( ((n/2) + 1) \) points, respectively, where \( n \) denotes the total number of observations for both the datasets. In general, the division of the residuals and the size of the sample simulated from \( \alpha S \) can be done randomly. Finally, we obtain 100 \( p \)-values of the test and create a boxplot as shown in Figures 1 and 2. Assuming a 0.05 confidence level, we fail to reject the hypothesis that the samples are from the same distribution.

Additionally, the QQ plots (generated using chart.QQPlot function in ‘PerformanceAnalytics’ package in R) as shown in Figures 1 and 2, represent the quantiles for fitted stable distribution with the estimated parameters from the residuals and the empirical quantiles for residuals. We observe the stable distribution seems to provide a good fit to the residuals.

To compare the fitness of the distributions, we also plot the densities of Gaussian, Students’-t, \( \alpha S \)-MECF and \( \alpha S \)-MLE for both the datasets using the fitted parameters. For Student’s-\( t \), the estimated shape (degrees of freedom) obtained are 6.67 and 5.87 for IBM and SP500DGE, respectively. For \( \alpha S \)-MECF, estimated \( \alpha \) values are 1.806 and 1.829, whereas for \( \alpha S \)-MLE, estimated \( \alpha \) values are 1.864 and 1.846 for IBM and SP500DGE, respectively. We observe that the density of the fitted \( \alpha S \) (MECF) and \( \alpha S \) (MLE) are closer to the density of residuals of both the datasets, in comparison to other fitted distributions.

6. Concluding remarks

To conclude, we make the following observations in relation to our proposed methods.

1. In this article, in order to estimate the ARMA coefficients of ARMA\((p, q)\) process with \( \alpha S \) noise as defined in (2) or ARMA\((p_A, q_A)\) process with \( \alpha S \)-GARCH\((p_G, q_G)\) as defined in (4), we use the Modified Yule-Walker method [9] in the Hannan–Rissanen method and obtain the estimates of the parameters using LAD regression.

2. Through simulations, we observed that the proposed MHR method can also be implemented for ARMA processes with finite variance noise. The M-estimators namely, the LAD and MLE perform well when the noise is heavy-tailed (stable). However, the LAD estimator is preferred over the MLE as it is computationally efficient, robust and does not require information (or estimation) of the parameters of the noise distribution. The proposed MHR method also inherits the same advantages of LAD especially for \( \alpha \in [1.5, 2] \), \( \theta_1 \in (0, 0.5] \) and \( \phi_1 \in (0, 1) \). For details on the limitations of LS and MLE estimator, see [2].

3. For the estimation of the GARCH coefficients for GARCH\((p, q)\) process with \( \alpha S \) noise defined in (3) and ARMA\((p_A, q_A)\) process with \( \alpha S \)-GARCH\((p_G, q_G)\) defined in (4), we introduce and discuss MECF for the case \( p = 1 \) and \( q = 1 \) for simplicity. Through
simulations, we observed that MHR+MECF method is computationally efficient and robust in comparison to MLE.

(4) Finally, we give an application of our proposed methods, where the noise distribution of the dataset is considered to be $S\alpha S$-GARCH due to volatility clustering. From Tables 6–8, we observe the estimates are statistically significant and Figures 1 and 2 shows efficient modelling of the financial data using through residual analysis.

**Disclosure statement**

No potential conflict of interest was reported by the author(s).

**Funding**

This work was supported by Indian Institute of Technology Madras [grant number SB20210848MA MHRD008558].

**References**

[1] T. Bollerslev, *Generalized autoregressive conditional heteroskedasticity*, J. Econom. 31 (1986), pp. 307–327.
[2] M. Cadler and R.A. Davis, *Inference for linear processes with stable noise*, in *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*, R.J. Adler, R.E. Feldman, and M.S. Taqqu, eds., Birkhauser Boston Inc., Cambridge, MA, 1998, pp. 159–176.
[3] G. Calzolari, R. Halbleibl and A. Parrini, *Estimating GARCH-type models with symmetric stable innovations: Indirect inference versus maximum likelihood*, Comput. Stat. Data Anal. 76 (2014), pp. 158–171.
[4] Z. Ding, C. Granger and R.F. Engle, *A long memory property of stock market returns and a new model*, J. Empir. Finance 1 (1993), pp. 83–106.
[5] L. Feng and Y. Shi, *A simulation study on the distributions of disturbances in the GARCH model*, Cogent Econ. Finance 5 (2017), p. 1355503. doi: 10.1080/23322039.2017.1355503
[6] C.M Gallagher, *A method for fitting stable autoregressive models using the autocovariation function*, Stat. Probab. Lett. 53 (2001), pp. 381–390.
[7] A. Grzesiek, S. Sundar and A. Wyłmańska, *Fractional lower order covariance-based estimator for bidimensional AR(1) model with stable distribution*, Int. J. Adv. Eng. Sci. Appl. Math. 11 (2019), pp. 217–229. doi: 11.10.1007/s12572-019-00250-9
[8] E.J. Hannan and J. Rissanen, *Recursive estimation of mixed autoregressive-moving order*, Biometrika 69 (1982), pp. 81–94.
[9] P. Kruczek, A. Wyłomanska, M. Teuerle and J. Gajda, *The modified Yule-Walker method for $\alpha$-stable time series models*, Physica A 469 (2017), pp. 588–603.
[10] P. Lévy, *Théorie des erreurs la loi de Gauss et les lois exceptionelles*, Bull. Soc. France 52 (1924), pp. 49–85.
[11] S. Liu and B. Brorsen, *Maximum likelihood estimation of a GARCH-stable model*, J. Appl. Econom. 10 (1995), pp. 273–285.
[12] T. Mikosch, *Parameter estimation for ARMA models with infinite variance innovations*, Ann. Stat. 23 (1995), pp. 305–326.
[13] S. Mittnik and M.S. Paolella, *Prediction of financial downside risk*, in *Handbook of Heavy Tailed Distributions in Finance*, S. Rachev, ed., Elsevier, Amsterdam, 2003.
[14] V. Omelchenko, *Parameter Estimation of the Stable GARCH(1, 1)-Model*, WDS’09 Proceedings of Contributed Papers, Part I, MATFYZPRESS, CZ, 2009, pp. 137–142.
[15] A. Panorska, S. Mittnik and S. Rachev, *Stable GARCH models for financial time series*, Appl. Math. Lett. 8 (1995), pp. 33–37.
[16] G. Samorodnitsky and M.S. Taqqu, *Stable Non-Gaussian Random Processes, Stochastic Models with Infinite Variance*, Chapman and Hall, New York, 1994, 632 pages.

[17] A.M. Sathe and N.S. Upadhye, *Estimation of the parameters of multivariate stable distributions*, Commun. Stat.- Simul. Comput. (2019), pp. 1–18. doi:10.1080/03610918.2020.1784432

[18] D. Wuertz, Y. Chalabi and L. Luksan, *Parameter estimation of ARMA models with GARCH/APARCH errors: An R and SPlus software implementation*, J. Stat. Softw. 55 (2009), pp. 28–33.

[19] M. Zhang and C. Li, *Application of GARCH model in the forecasting of day-ahead electricity prices*, in *Third International Conference on Natural Computation*, IEEE, Haikou, 2007, pp. 99–103.

[20] B. Zhou, D. He and Z. Sun, *Traffic predictability based on ARIMA/GARCH model*, in *Next Generation Internet Design and Engineering*, IEEE, Valencia, 2006, pp. 199–207.