STATISTICS IN TRANSITION new series, September 2026
Vol. 27, No. 3, pp. xx–XX, https://doi.org/10.59139/stattrans-2026-xxx
Received – xx.xx.xxxx; accepted – xx.xx.xxxx

Modeling exchange rate volatility using EM algorithm with hybrid GARCH models

Ali Salman Habeeb[1]

Abstract

The most important basic tools for modeling and analyzing econometric models are linear and nonlinear time series analyses, with the aim of formulating an accurate probabilistic model that expresses the behavior of the phenomenon over a specific time period. In this study, we construct a three-step hybrid probability model that includes model identification, model estimation, and diagnosis checking using the Box-Jenkins approach. We explored the performance of two important tools for estimating unknown parameters of the proposed model. The first is Maximum Likelihood (ML), which has a wide reputation in the statistical literature, and is provided in $Eviews10$ software. Furthermore, we compared it with the second proposed numerical method, namely Expectation-Maximization (EM algorithm), which requires creating a special MATLAB code to implement its performance. Moreover, we compared two methods using information criteria (AIC, BIC, HQ). In order to show the effectiveness of the hybrid ARCH-GARCH models, we used them to measure the volatility of the weekly exchange rate of the Central Bank of Iraq's sales of the US dollar against the Iraqi dinar from the first week of 2021 to the last week of 2023. We concluded that the appropriate model is the hybrid $MA(1)-GARCH(1,1)$ model, which is represented by the Moving Average $MA(1)$ of the average equation and the $GARCH(1,1)$ model of the conditional variance equation.

Key words: time series analysis, exchange rate volatility, maximum likelihood, expectation-maximization(EM algorithm), generalized autoregressive conditional heteroscedasticity.

1. Introduction

The analysis of financial time series is one of the important topics in the field of financial econometrics because of its reflection the various developments of quantitative economics. Therefore, it has attracted the interest of many researchers in this field because it focuses on the evaluation of assets and their volatility, as well as the series of stock returns over longer periods of time that cannot be directly observed. Thus, the importance of statistical methods appears in analyzing real phenomena related to time, including time series analysis as a random uncertain process subject to the laws of probability and uncertainty, which requires us to reveal the mathematical structure of both types (linear and nonlinear) inherent in the data through statistical analysis or modeling and identifying the appropriate model. That phenomenon is defined through diagnosis and estimation, and then it is used in the issue of predicting the future values of the phenomenon under research. In this paper, we

[1]Department of Statistics, University of Sumer, Dhi Qar Governorate, Al-Rifai, Iraq.
E-mail: asalhabeeb@gmail.com. ORCID: https://orcid.org/0000-0003-0037-8292.

considered one of the most difficult topics facing decision-makers in the field of financial policy making in Iraq, which is identifying the best econometric models to modeling the phenomenon of financial fluctuations over a relatively long period of time extending from the first week of January 2021 to the last week of January 2023, with 156 weekly observations (five trading days a week). In addition, what are the appropriate monetary policies in the event of random financial shocks in the Iraqi stock market? One of the first nonlinear models that dealt with the issue of volatility was the autoregressive model with the condition of heterogeneity of variance (ARCH), presented by researcher Engle (1982) in an article that studied the autoregressive model of the first order and the unconditional mean to be zero. Draper and Joha (1982) the presented research based on the nonlinearity test for residual series, while Bollerslev (1986) developed the ARCH model to become GARCH, and Cochrane (1988) investigated a measure of the continuity of volatility in Gross National Product (GNP) based on variation and its differences in the long run and studied a group of models, perhaps the most important of which is the random walk model. Nelson (1991) used the GARCH models to model the relationship between conditional variance and asset risk. These models have at least some major defects when pricing assets, perhaps the most important of which is that random shocks do not continue in GARCH models due to stationarity measurement standards. Hall and Yao (2003) showed that asset aggregation and risk calculations rely heavily on volatility models and provided a comparison of forecast performance using a group of ARCH models and working on daily exchange rate data against the dollar and the importance of specific and correct distribution of asset returns and not focusing only on daily volatility. Awartani and Corradi (2005) studied the relative predictive ability of different types of GARCH models, with a focus on asymmetric models. Pairwise comparisons were made with different models compared with the $GARCH(1,1)$ model; the same finding applies to different longer forecast horizons, although the predictive superiority of asymmetric models is not as striking as in the one-step-ahead comparison. So and Yu (2006) investigated the efficacy of seven distinct GARCH model specifications, incorporating various market indicators and foreign exchange rate to evaluate Value at Risk (VaR) and according to the RiskMetrics framework across multiple error distribution levels, they demonstrated that these models exhibit asymmetric error distributions. Furthermore, their analysis confirmed that daily exchange rate data are characterized by heavy-tailed distributions and long-memory properties. Ghahramani and Thavaneswaran (2008) confirmed that financial returns are often designed as a time series, especially the GARCH model, and the error distribution is often determined from parametric distributions, which was not justified. They confirmed in the research a specific method for diagnosing the identity of GARCH models for a number of data sets. On the other hand, Fan and Xiu (2014)) proposed that the non-Gaussian maximum likelihood estimator be used frequently in GARCH models; with heavy-tailed distributions, the estimator is biased and inconsistent in the case of the parametric likelihood family. They provided an unknown scale parameter to the identification for consistency, and they proposed a three-step quasi-maximum likelihood procedure, particularly when likelihood functions have heavy tails. Kononovicius and Ruseckas (2015) used the normal distribution in modeling volatility even if the return distribution is asymmetric and has excessive kurtosis. Klein and Walther (2016) assumed a log-normal distribution of oil price returns in order to compare different types of ARCH models in terms of accuracy of forecasting volatility. Ewing and Malik (2017) found that there is a problem with the issue of asymmetry and excessive kurtosis in oil revenues, which are slightly low when the model calculates the structural breaks. Cerqueti et al. (2020) proposed the application of non-Gaussian GARCH models of volatility in a new and comprehensive study about the cryptocurrency market, evaluating the forecasting performance for three of the most important cryptocurrencies (Bitcoin, Ethereum, and Litecoin) in terms of market capitalization, the results demonstrated the effectiveness of the prediction performance. Massadikov (2021) used the VAR-GARCH model to determine the interaction between oil prices and stock markets in terms of return and volatility in the financial markets of some developing countries. It was demonstrated through the use of this model that it is important in discovering and interpreting the reflection of risks for some economies of oil-producing countries. Habeeb et al. (2021) discussed the method of estimating smoother splines for the Additive Autoregressive Model, as it is one of the most important numerical methods used in curve fitting of the regression function. Tashpulatov (2022) in his paper presents volatility modeling for wholesale electricity market prices in England and Wales over a determined period by the AR-ARCH model with four flexible distributions that account for asymmetry, heavy tails, excess kurtosis, and a peak higher than in normal distribution, observed in the data or model residuals; it applied generalized deviation, and that distribution is suitable to study this case. Chifurita and Chinhamn (2022) dealt with the daily returns of financial market variables, which often have empirical distributions that are heavy or semi heavy tailed, and they also proved that the estimation value-at-risk depends highly on the distributional characteristics of the stock returns. The aim of the study is to investigate the relative performance of the error distributions of the GARCH model, such as the generalized hyperbolic skew Student-t and Pearson $type-IV$ distributions, as this study recommends that the $ARMA(1,1)EGARCH(1,1)$ model be used in the estimation of the VaR for the daily returns from the FTSE/JSE growth index (J280). Ampadu et al. (2024) presented a study to exhibit volatility clusters and effects on the importance of error distribution to improve the performance of the GARCH model through simulation experiments of five different types of error distributions. The study showed that the Skewed Generalized Error Distribution (SGED) is the most efficient and consistent for modeling daily return volatility through the use of different criteria. Antony and Prabakarant (2024) considered the spatial autoregressive model, which is applied to state-wise $COVID19$ confirmed rates. This article aims to build a stochastic model that describes the financial volatilities of the Iraqi dinar exchange rate against the weekly US dollar in Iraq for the period (2021-2023) in the Iraqi stock market using the $GARCH(1,1)$ model, and in the presence of skewed distributions for the error term, particularly those exhibiting heavy tails, deploying a robust mathematical model is essential for phenomenon study. In this paper, we provide the theoretical background by describing the theoretical basis of the weak stationarity of the model and building a theoretical analysis of the ARCH model, paying particular attention to the organization and some important tests in this context. Then we studied the $GARCH(1,1)$ model and discussed the estimation process. Finally, we reach some brief conclusions, including estimating the proposed model.

2. Methodology

Most financial studies deal with a series of asset returns instead of price returns, but in this research, we will deal with price returns (Bala, 2013). If we assume that pt represents the simple price for one period of time (selling price at closing) over a time period (such as a day, a week, or a month), and pt1 represents the simple price for a previous period of time, then the return is $(r_t = \frac{p_t}{p_{t-1}})$. So, the returns can be represented through a logarithmic transformation of the ratio of the return values for the current period relative to the last period. In this paper, we consider the mathematical formula as follows: $y_t = \log(\frac{x_t}{x_{t-1}}) = \log(x_t) - \log(x_{t-1})$,

where y_t: represents the weekly returns of the exchange rate, x_t is the exchange rate of the current period, and x_{t-1} is the exchange rate of the previous period. Volatility represents the standard deviation of returns (symbolized by σ_t)). It measures the price behavior of returns and is sometimes considered to be the accompanying risk.

The mathematical formula adopted to calculate it is: $\sigma_t^2 = \sqrt{\frac{1}{n-1}\sum_{t=1}^{n}(y_t - \mu)^2)}$, where: μ represents the arithmetic mean, and n represents the number of observations in the sample. When representing volatility in returns by the *GARCH* model, there are two distinctive characteristics of these models, the conditional mean and the conditional variance. $E(y_t|y_{t-1},\ldots,y_{t-p}) = \mu_t \quad and, Var(y_t|y_{t-1},\ldots,y_{t-p}) = E(y_t^2|y_{t-1}^2,\ldots) = \sigma_t^2$.

2.1. Definition for Model and Stationary of the Series

In some articles for studying and analyzing the time series, the series $\{y_t\}_{t=1}^{n}$ must be stationary, and it has two types, The first is Strict stationarity, which assumes that the time series is $\{y_t, t \in \mathscr{Z}\}$, where $\mathscr{Z}$ is the set of integers; we say that it is strictly stationary in the distribution function for the time period (t) and has the same distribution for the subsequent time period $(t+k)$, and this type is called strongly stationary and can be expressed when the following condition holds: $F(y_{t_1},\ldots,y_{t_n}) = F(y_{t_1+k},\ldots,y_{t_n+k})$. The joint probability distribution for the variables is the same as for the variables $(y_{t_1+k},\ldots,y_{t_n+k})$. The second type is weak stationarity if we have the following: $i)E(y_t) = \mu, \quad ii)E(y_t^2) < \infty; \quad iii)\gamma_y(s,t) = \gamma_y(s+k,t+k) = Cov(y_t, y_{t+k}).\forall t \in \mathscr{Z}$, where k is integer numbers, this means $\gamma(k), \forall k = 0,1,2,\ldots$. In addition, if $k = 0$, then the covariance of the time series is the variance $(i.e. \gamma(0) = \sigma_t^2)$. The autocorrelation function measures the linear dependence between $y_t \, and \, y_{(t+k)}$. Likewise, the autocorrelation function is given according to the following mathematical formula: $\rho_k = \frac{Cov(y_t, y_{t+k})}{Var(y_t)} = \frac{\gamma(k)}{\gamma(0)}$, where k is an integer, this means $\rho_k, \forall k = 0,1,2,\ldots$, and it is called the autocorrelation function. In addition, if $k = 0$ then $\rho_0 = 1$. Also, the autocorrelation function is independent of the time series measurements. The most important test that has wide application among researchers is the Augmented Dickey-Fuller Approach (ADF), which is one of the unit root tests in the search for the stationarity of time series, and the expanded form of this test works to estimate the first-order autoregressive model:

$$\begin{aligned} &\Delta y_t = \phi y_{t-1} + \sum_{j=1}^{p-1} \alpha_j^* \Delta y_{t-j} + \varepsilon_t, \\ withlinearTrend \qquad &\Delta y_t = \delta + \phi y_{t-1} + \sum_{j=1}^{p-1} \alpha_j^* \Delta y_{t-j} + \varepsilon_t; \end{aligned} \tag{1}$$

where $\phi = -\alpha(1), \quad \alpha_j^* = -(\alpha_{j+1}, \ldots, \alpha_p); \qquad \Delta y_t = y_t - y_{t-1}$. Therefore, we test the significance of the above model coefficient, which in turn means testing the significance of the unit root (ϕ) through the following hypothesis: $H_0 : \phi = 0 \quad Vs. \quad H_1 : \phi < 1$. This condition is verified for the ARCH(p) model when the following is true: $i)E(y_t^2 < \infty, ii)\sum_{i=1}^{p} \alpha_i < 1$. The roots lie outside the unit circle. When we are interested in the ARCH(1) model, the condition for weak stationary is $(\alpha_1 < 1)$. As for the GARCH(p,q) model, weak stationary is achieved by verifying the following (Andersen et al. 2009) $i)Ey_t^2 < \infty, ii)\sum_{i=1}^{p} \alpha_i + \sum_{i=1}^{p} \beta_j < 1$. The roots lie outside the unit circle. When we are interested in the GARCH(1,1) model (Jafari et al. 2007), the condition for weak stationary is $(\alpha_1 + \beta_1 < 1)$. The basic idea is to find a model that can sense the effects of unpredictable volatility that are the main cause of changes (in more detail, e.g. Alberg et al. 2008; Angabini and Wasiuzzaman 2011 and Engle 2001). And uncertainty here indicates the inability of the model to predict the future that will affect the phenomenon's behavior, and that addressing this idea requires a stochastic model that can sense changes over time. And y_t is a random variable that has a conditional mean and conditional variance defined on the set of information $\mathscr{F}_{t-1}$ (generated from the $\sigma-$field) defined through the random process $\{y_t; t \in \mathscr{Z}\}$ and it has weak stationarity. The general form of the model is of the order ARCH(p) for positive integer $(p \geq 1)$, the random process has $E(y_t|\mathscr{F}_{t-1}) = 0 \quad and \quad E(y_t^2|\mathscr{F}_{t-1}) = \sigma_t^2$ (see more details in, e.g. Draper and John (1982); and So and Yu (2006)).

$$y_t = \mu + e_t \implies e_t = \varepsilon_t \sqrt{\sigma_t^2}; \tag{2}$$

where ε_t is the random error series, which is independent and identically distributed with zero mean and specified variance (σ_t^2), and is denoted by the process $\{\varepsilon_t; t \in \mathscr{Z}\}$, and y_t return series. It can be shown that the conditional distribution of $y_t|\mathscr{F}_{t-1} \sim N(0, \sigma_t^2)$, where σ_t^2 is the conditional variance of a time series for different time periods.

$$\begin{aligned} Var(y_t|y_{t-1}, \ldots, y_{t-p}) = E(y_t^2|y_{t-1}, \ldots) - E^2(y_t|y_{t-1}, \ldots) = \sigma_t^2 \\ \sigma_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \cdots + \alpha_p e_{t-p}^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i e_{t-i}^2; \end{aligned} \tag{3}$$

where $\alpha_0 > 0, \alpha_i \geq 0; \forall i = 1, 2, \ldots, p$.

2.2. Generalized Autoregressive Conditional Heteroskedicity GARCH(p,q) Model

According to the conditional variance relationship in (3) the current variance is also dependent upon the lag variables of the conditional variance series. $\sigma^2_{t-1},\ldots,\sigma^2_{t-q}$. In addition, the squares of the time lags for the process $\{y_t; t \in \mathscr{X}\}_{t=1}^{T}$ become a general model called $GARCH(p,q)$, where $(p \geq 0, q \geq 1)$. (see, Bollerslev (1986); Kononovicius and Ruseckas (2015) and Lundbergh and Teräsvirta (2002)).

$$\begin{aligned} & y_t = \mu + e_t \implies e_t = \varepsilon_t\sqrt{\sigma_t^2}\,. \\ and, \quad & \sigma_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \cdots + \alpha_p e_{t-p}^2 + \beta_1\sigma_{t-1}^2 + \cdots + \beta_p\sigma_{t-q}^2, \\ & \sigma_t^2 = \alpha_0 + \sum_{i=1}^{p}\alpha_i e_{t-i}^2 + \sum_{j=1}^{q}\beta_j\sigma_{t-j}^2. \end{aligned} \tag{4}$$

where $\alpha_0 > 0 \quad ,\alpha_i \geq 0; \quad \beta_j \geq 0 \quad \forall i,j = 1,2,\ldots,p,q$; respectively. The model parameters to be estimated, and we can write the relationship in (4), (Andersen et al. (2009) and Ding (2021)) as follows:

$$y_t = \mu + e_t \implies e_t = \varepsilon_t\sqrt{\sigma_t^2} \quad and,\, \sigma_t^2 \quad = \alpha_0 + \sum_{i=1}^{p}\alpha_i e_{t-i}^2 + \sum_{j=1}^{q}\beta_j\sigma_{t-j}^2. \tag{5}$$

3. Testing the model

In the literature of time series analysis, volatility represents the standard deviation of currency sales returns in dollars (Mapa 2004). It measures the price behavior of the currency and is sometimes expressed as the current exchange rate. In order to verify that the model is free of any problems that negatively affect the accuracy of the model (see, e.g. Bangura et al. (2021) and Draper and Joha (1982)), a set of tests can be used in this regard Lagrange Multiplier test is one of the most important tests used to detect that the model contains the problem of heteroscedasticity, Draper and Joha (1982). To test the existence of the heteroskedasticity problem, the coefficient of determination is calculated, which depends on the regression of the squares of the residuals for the current period with respect to the squares of the residuals for the previous period, and its statistic for this test is

$$ARCH - test = T \times R^2 \sim \chi^2_{(p)}; \tag{6}$$

where is the number of model parameters, T is the size of the sample, and R is the coefficient of determination. We calculate the test statistic; it follows a chi-square distribution with degrees of freedom (p). The process $\varepsilon_t^2 = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \cdots + \alpha_p\varepsilon_{t-p}^2$. As for the hypothesis that we are testing, $\quad H_0 : \alpha_1 = \cdots = \alpha_k = 0 \quad Vs. \quad H_1 : \alpha_k \neq 0, \quad \forall k = 1,\ldots,K.$

4. Theoretical distributions of the model residuals

Let y_t denote a real-valued discrete-time stochastic process, and let $\mathscr{F}_t$ denote the $\sigma-algebra$ generated by the history time series. The process $\{y_t; t \in \mathscr{Z}\}$ is represented for time series, then y_t follows a GARCH(p,q) model, and this model can also be written in (5) as follows:

$$\begin{aligned} y_t &= \mu + e_t, \qquad (Mean\ Equation) \\ and,\ e_t = \varepsilon_t \sigma_t &\implies \sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i e_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2; \quad (Variance\ Equation) \end{aligned} \tag{7}$$

where $\alpha_0 > 0, \quad \alpha_i \geq 0, \quad \beta_j \geq 0, \quad \forall\ i = 1,2,\ldots,p\ \&\ j = 1,2,\ldots,q$. The model in (7) will be stationary if the following condition is met $\left[\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1\right]$.

Here if the process is white noise (i.e. $\varepsilon_t \sim N(0,1)$), we have the following:

$$\begin{aligned} E(y_t|\mathscr{F}_{t-1}) &= E(\sigma_t\varepsilon_t) = E(E(\sigma_t\varepsilon_t|\mathscr{F}_{t-1})) = E(\sigma_t E(\varepsilon_t|\mathscr{F}_{t-1})) = E(\sigma_t E(\varepsilon_t)) = 0, \\ Var(y_t|\mathscr{F}_{t-1}) &= E(y_t^2|\mathscr{F}_{t-1}) - E^2(y_t|\mathscr{F}_{t-1}) = E(\sigma_t^2\varepsilon_t^2|\mathscr{F}_{t-1}) = E(\sigma_t^2 E(\varepsilon_t^2)) = \sigma_t^2. \end{aligned}$$

Then, $y_t \sim D(0,\sigma_t^2)$, $Cov(y_t y_s) = 0$ for $t \neq s \Rightarrow \sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j < 1$. Furthermore,

$$\begin{aligned} E(\sigma_t^2|\mathscr{F}_{t-1}) &= E(\alpha_0|\mathscr{F}_{t-1}) + \sum_{i=1}^{q} \alpha_i E(y_{t-i}^2|\mathscr{F}_{t-i}) + \sum_{j=1}^{p} \beta_j E(\sigma_{t-j}^2|\mathscr{F}_{t-j}), \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^{q} \alpha_i \sigma_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 = \frac{\alpha_0}{1 - \sum_{i=1}^{q} \alpha_i - \sum_{j=1}^{p} \beta_j}. \end{aligned} \tag{8}$$

This implies $E[y_t|\mathscr{F}_{t-1}] = 0$ and $E[y_t^2|\mathscr{F}_{t-1}] = \sigma_t^2 \implies y_t|\mathscr{F}_{t-1} \sim D(0,\sigma_t^2)$.

For the $GARCH(p,q)$ model, which is defined in (9), assuming that ε_t is the disturbance term, it has three different types of distributions called heavy-tailed distributions (see more details in, e.g. Hall and Yao (2003), Preminger et al., (2017), and Chen and Liu (2013)), which can be given according to the following cases:

Case I The Gaussian distribution: If ε_t is a random variable with probability density function as follows: $f(\varepsilon_t;\sigma_t^2) = \frac{1}{\sqrt{2\pi\sigma_t^2}} e^{-\frac{1}{2}\frac{\varepsilon_t^2}{\sigma_t^2}}$ where $\varepsilon_t \in \mathscr{R}, \mu = 0; \sigma_t \in \mathscr{R}^+$ is the scale parameter, denoted by $\varepsilon_t \sim N(0,\sigma_t^2)$.

Case II The Student's t distribution: If ε_t is a random variable, the probability density function of ε_t, is defined as (see, e.g. Adubisi et al., (2022) and Awalludin et al., (2018)): $f(\varepsilon_t;\sigma_t^2,v) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)\sigma_t\sqrt{\pi(v-2)}} \left[1 + \frac{(\varepsilon_t)^2}{(v-2)}\right]^{-\left(\frac{v+1}{2}\right)}$ where $\varepsilon_t \in \mathscr{R}, \mu = 0; \sigma_t \in \mathscr{R}^+$ is the scale parameter, and $2 < v < \infty$ is the positive integer number (degrees of freedom) and $\Gamma(.)$ is the Euler gamma function.

Case III The Generalized Error Distribution (GED): If ε_t is a random variable, and it has probability density function as follows: $f(\varepsilon_t;\sigma_t,v)=\frac{v}{\lambda\sigma_t 2^{(1+1/v)}\Gamma(1/v)}.\exp\left(-\frac{1}{2}\left|\frac{\varepsilon_t}{\lambda}\right|^{v}\right)$. where $\varepsilon_t\in\mathscr{R},\mu=0,\sigma_t\in\mathscr{R}^{+}\quad\lambda=\frac{1}{2^{\frac{1}{v}}}.\left[\frac{\Gamma(1/v)}{\Gamma(3/v)}\right]^{\frac{1}{2}}$, and $v\in\mathbb{N}$ is the shape parameter.
Therefore, it is necessary to know the type of distribution of the error term, which varies according to the distributions, as the variance is not constant in this type of model, which has heavy tails, and by plotting the probability density function for financial volatility data, we can identify these important distributions when analyzing time series . We have shown

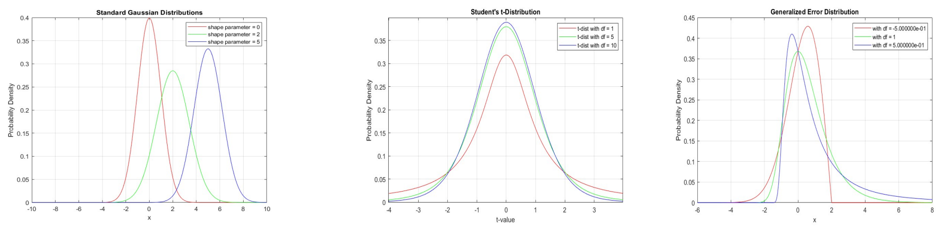

Figure 1. Probability density function for: (a) Plot of the Gaussian distribution with different shape parameters; (b) Plot of the Student's t-distribution with different degrees of freedom; (c) Plot of the GED with different skewed values

in (1) the three cases of the different distributions: the Gaussian distribution for different shape parameters, the Student's t-distribution with different degrees of freedom, and the generalized error distribution with heavy tails.

5. Estimation of the Model

One of the methods commonly used in the literature to estimate model parameters is the maximum likelihood method because it has good and desirable theoretical properties when interested in estimating the parameters of this type of the (ARCH or GARCH) model (in more detail, e.g. Huang et al., (2008) and Rzgar and Kale (2007)), which represents financial volatility, and we can express it by the standard deviation (variance) of the random error term (see more details in, e.g. Augustyniak (2014) and Bera and Higgins (1993)). Therefore, it is necessary to know the type of distribution of the error term, which varies according to the distributions, as the variance is not constant in this type of model, which has heavy tails, and by plotting the probability density function for financial volatility data we can identify these important distributions, in particularly when analyzing time series. We assume that the random error term has a normal distribution, and then the error term is white noise at a mean of zero and a specified variance, and this case is considered optimal (PzCruz et al., (2003)). However, this assumption may be inaccurate and thus affect the accuracy of choosing the appropriate model for the analysis. Therefore, the distribution of the residuals must be tested for the appropriate distribution (Gao et al., (2012)).

5.1. Maximum Likelihood Method

The maximum likelihood method is considered one of the more common traditional approaches used in the statistical literature for estimating model parameters because it has good theoretical properties and gives the best fit for the candidate model. We are interested in estimating the unknown parameters of the ARCH and GARCH models (see, e.g. Huang et al., (2008) and Ruzgar and Kale (2007)). This model represents financial volatility, which is expressed by the standard deviation (variance) of the random error term (in more detail, e.g. Augustyniak (2014); Bera and Higgins (1993) and Horv and Kokoszka (2003)). Furthermore, we used this method as a basic tool to deal with the GARCH(p,q) model. We assume that the residuals of the model have the probability distributions mentioned in section 4.

Let $\{e_t,\, t = 1,2,\cdots,T\}$ be a sequence of random residual variables from a distribution that has a probability density function $f(e_t,\theta)$, where θ is the vector of parameters of the model. The $f(\theta|e_t)$ represents the conditional function of the GARCH(p,q) model, and it has an associated probability distribution function. Therefore, the likelihood function is formally defined as: $\ell(\theta|e_t) = \sum_{t=1}^{T} f(\theta|e_t)$. In other words, the likelihood function is often replaced by its natural logarithm, and we have the log-likelihood function given by $L(\theta) = \ln\sum_{t=1}^{T}(\ell(\theta|e_t))$. Therefore, this approach can be applied to the three cases mentioned in Section 4; we obtain the following:

Case I The log likelihood function of the $GARCH(p,q)$ model if the residual has the Gaussian distribution:

$$L(\theta) = \frac{-T}{2} ln(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\left(ln(\sigma_t^2) + (\frac{\varepsilon_t^2}{\sigma_t^2})\right). \tag{9}$$

Case II The log likelihood function of the $GARCH(p,q)$ model if the residual has the Student's t distribution:

$$\begin{aligned} L(\theta) = T&\left(ln\Gamma\left(\frac{v+1}{2}\right) - ln\Gamma\left(\frac{v}{2}\right) - \frac{1}{2}ln(\pi(v-2))\right) \\ &-\frac{1}{2}\sum_{t=1}^{T}\left(ln(\sigma_t^2) + (1+v)ln(1+\frac{\varepsilon_t^2}{v-2})\right). \end{aligned} \tag{10}$$

Case III The log likelihood function of the $GARCH(p,q)$ model if the residual has the Generalized Error Distribution (GED):

$$L(\theta) = \sum_{t=1}^{T}\left(ln(\frac{v}{\lambda}) - \frac{1}{2} \mid \frac{\varepsilon_t}{\lambda} \mid^{v} - (1+v^{-1})ln(2) - ln\Gamma\left(\frac{1}{v}\right) - \frac{1}{2}ln(\sigma_t^2)\right) \tag{11}$$

where $\lambda = 2^{\frac{-1}{v}}\Gamma\left(\frac{1}{v}\right)^{\frac{1}{2}}\Gamma\left(\frac{3}{v}\right)^{\frac{-1}{2}}$, and T is sample size.

Then the maximum likelihood estimator (MLE) for the parameters is found determining the value that maximize $L(\theta|e_t)$, which is:

$$\hat{\theta}_{MLE} = \underbrace{\textbf{arg max}}_{\theta\in\Theta} \Big(L(\theta)\Big). \tag{12}$$

5.2. Expectation-Maximization Method

The EM algorithm is a general method as an iterative approach used to solve maximum likelihood estimation problems with the presence of incomplete data. Let y denote the incomplete data (observed), and the probability density $f_Y(y;\phi)$ indexed by the parameter vector $\phi \in \Omega \subseteq \mathscr{R}^2$ and. Let ε denote the complete data (see, e.g. Feder and Weinstein (1988) and Zhang et al., (2021)).

$$f_\varepsilon(\varepsilon_t;\phi) = f_{\varepsilon|Y=y}(\varepsilon_t;\phi)\cdot f_Y(y;\phi) \tag{13}$$

where $f_\varepsilon(\varepsilon_t;\phi)$ is the probability density associated with ε, and $f_{\varepsilon|Y=y}(\varepsilon_t;\phi)$ is the conditional probability density of ε given $Y=y$. Taking the logarithm for both sides of (15), we obtain: $\log f_Y(y;\phi) = \log f_\varepsilon(\varepsilon_t;\phi) - \log f_{\varepsilon|Y=y}(\varepsilon_t;\phi)$, The conditional expectation given $Y=y$ at a parameter value ϕ

$$E\left\{\log f_Y(y;\phi)\right\} = E\left\{\log f_\varepsilon(\varepsilon_t;\phi)|Y=y;\phi\right\} - E\left\{\log f_{\varepsilon_t|Y=y}(\varepsilon_t;\phi|Y=y;\phi)\right\} \tag{14}$$

The algorithm starts with an arbitrary initial value $\phi(0)$, and $\phi(j)$ denotes the current estimate of ϕ after j iterations of the algorithm. The iteration cycle of the EM algorithm can be described. Therefore, we can express this algorithm through two steps, namely the E-step and the M-step, which are described below for the GARCH model fitting. Hence, the Φ function of the algorithm (Popovici and Dumitrescu 2007) will be as follows:

E-step: In the first step, we Assume that $\Phi\left(\phi|\phi^{(m)}\right)$ is the function for unknown parameters, and the vector $\Phi = \left(\mu,\theta,\alpha_0,\alpha_1,\beta_1\right)$ denotes the parameters of the model:

$$\begin{aligned}\Phi\left(\phi|\phi^{(m)}\right) &= E_{\varepsilon_t|Y=y,\phi^{(m)}}\left[\log\Pi_{t=1}^n f_\varepsilon(\varepsilon_t;\phi)\right] \quad = E_{\varepsilon|Y,\phi^{(m)}}\left[\sum_{t=1}^n \log f_\varepsilon(\varepsilon_t;\phi\right] \\ &= \sum_{t=1}^n E_{\varepsilon|Y,\phi^{(m)}}\left[\log f_\varepsilon(\varepsilon_t;\phi)\right]\end{aligned} \tag{15}$$

M-step: The second step of the EM algorithm is that for maximizing the expectation. We perform the following: we use $(m+1)$ cycle iteration for from $\phi^{(m+1)}$, which is maximization for

$$\phi^{(m+1)} = \underbrace{\textbf{arg max}}_{\phi\in\Phi}\left(\Phi(\phi|\phi^{(m)})\right). \tag{16}$$

6. Application

To provide an overview of the dataset, and the purpose of getting a general idea for it. We present Table 1 which contains the descriptive statistics summary of the weekly exchange rate volatility of the Central Bank of Iraq's official sales of the US dollar against the Iraqi dinar. The observation period extends from the first week of 2021 to the last week of 2023. Accordingly, the indicators of the phenomenon under study are as follows.

Table 1. Summary of the descriptive statistics for weekly returns

Des. Stat.	Mean	Std.Dev.	Min	Max	Median	Ske.	Kur.	JB
Exchange Rate	1493	46.38	1408	1720	1480	1.915	7.367	219.36

It is noted from Table 1 that the average exchange rate is (1493) dinars per dollar in the currency auction and that the lowest price at which the dollar was sold against the dinar was the 121st week of May 2023, when it reached (1408) dinars per dollar. The highest selling price was in the 108th week of January 2023, when it reached 1720 dinars per dollar. The value of the median is 1480 dinars per dollar, and this indicates the rate of sale during the period studied, without being affected by outliers or extreme values. It is known that the value of the median is a robust indicator of the trend of outliers and extreme values, and the median is considered an appropriate average for measuring volatility in the exchange rate. The standard deviation was 46.38, and this indicates that the data has a high dispersion during the measured period. We also note that the skewness coefficient is far from zero, and it is positive. This indicates that the data has a distribution skewed to the right, in addition to the kurtosis coefficient, whose peak is much higher than 3. This indicates that it has a high peak that is higher than the peak of the normal distribution (2 a), which represents the time series that displays the behavior of the phenomenon of weekly volatilities of the exchange rate of the Iraqi dinar against the dollar during the time period (from the first week of 2021 until the last week of 2023) according to the sales tables published on the Central Bank of Iraq website. From Table 1, it can be noted that the value shown by the Jarque-Bera statistic tests the normal distribution of series data, which proved that the data does not follow a normal distribution. In addition, (2 b) is graphically observed on the $(q-q)$ plot; it shows strong deviation from normal distribution.

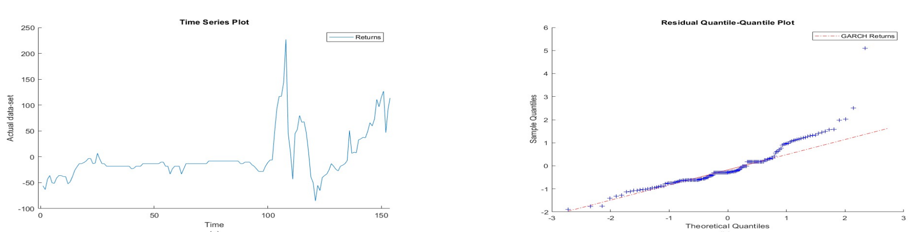

Figure 2. (a) The plot of the weekly returns; (b) The normal Q-Q of the weekly returns

Table 2. The stationary tests

Tests		Exchange Rate series			First Difference for series		
		t-Statistic	**Prob.**	**Test Critical Values**	**t-Statistic**	**Prob.**	**Test Critical Values**
I. Augmented Dickey-Fuller (ADF))	None	0.611	0.847	-2.58	-14.545	<0.01	-2.58
	Constant	-3.216	0.021	-3.47	-14.531	<0.01	-3.47
	Constant + Linear Trend	-4.112	0.008	-4.018	-14.482	<0.01	-4.018
II. Phillips-Perron (PP)	None	0.912	0.903	-2.579	-15.487	<0.01	-2.58
	Constant	-3.216	0.021	-3.47	-15.588	<0.01	-3.47
	Constant + Linear Trend	-3.991	0.011	-4.018	-15.569	<0.01	-4.019

6.1. The Box-Jenkins method for Analyzing and modeling

The time series of returns is nonstationary on average. In addition, there is severe volatility after 2023, as noted in (2.a). These are considered financial shocks (and the reason for this may be due to the entry of dollar sales into the international monitoring platform for currency sales or resulting from severe geopolitical changes that coincided with that period). And, this can be seen through the probability value of (p-value=0.9404); it is greater than the level of significance $(\alpha = 0.01)$, which was calculated by the MATLAB code program. In addition, it is possible to test the stability of the time series by using the Augmented Dickey-Fuller (ADF) test, as in Table 2, which shows the results of this test, which were calculated using the $EViews10$ program. The hypothesis test were evaluated under three cases determined : an intercept, an intercept with a linear trend, and the absence of both. The calculated results of these tests are presented in Table 2 below.

6.2. Augmented Dickey-Fuller Test

In the case of the two tests (Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP)), we see that the calculated value of t-statistics is less than the value of t-theoretical of this test except for the constant with linear trend in ADF, which is not less than the value of t-theoretical, and in this case, we see that this value is very critical, but the PP test supports the decision-maker, and hence the series is not stable in the level and stable in the first difference. Table 2 shows for the $(ADFandPP)$ tests that the probability values using the program $(EViews10)$ were $(P-values = 0.8473, 0.021,$ and $0.008 \simeq 0.01$ for each one from: none, constant, and constant with linear trend, respectively), which is greater than the level of significance $(\alpha = 0.01)$, and therefore the null hypothesis cannot be rejected in (6), and the series has a unit root; it is nonstationary, and the correction is through finding the first difference of the series data. In addition, the graphical representation between the original series data and the logarithmic transformation of the data can be compared in order to find out whether nonlinear transformations as shown in (3) can be used for the purpose of treating instability in the behavior of the phenomenon.

In order to prove this, we conduct the Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) tests for the exchange rate series and the first difference for the series according to what was found in the literature. We note from Table 2 for the Augmented Dickey-Fuller and Phillips-Perron tests that the probability value was (P-value = 0.0000), which is less than the level of significance $(\alpha = 0.01)$ for the first difference, and therefore, the null hypothesis cannot be accepted in (6) and that the series does not have a unit root. The time series is stable on average, and therefore the first difference series can be adopted in the analysis and model building. And to investigate the stationarity of the series and avoid the unit root problem in the exchange rate data, the Phillipa-Perron test was utilized. Additionally, this test was employed to corroborate the findings through the assessment of structural breaks within the series.

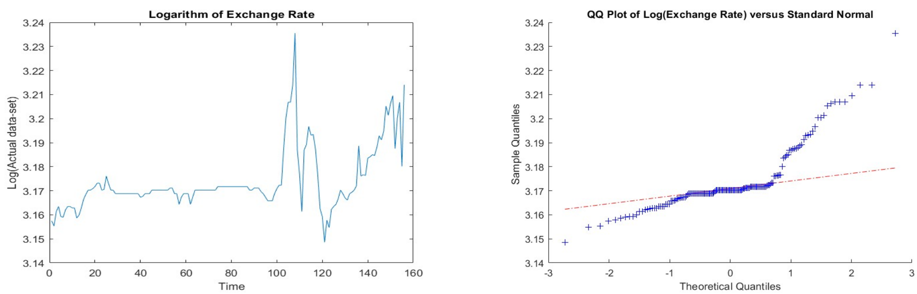

Figure 3. The plot of the logarithm of the weekly returns and the plot of the normal Q-Q of the logarithm of the weekly returns

6.3. Identification of the model

We used the sample ACF and PACF plots to suggest an appropriate parsimonious subclass order for the model, based on the data employed. Hence, we use the sample ACF and PACF plots in (5); it has an exponential decay pattern for PACF, and it cuts off after lag(1) for ACF. Therefore, the proposed appropriate model can be the MA(1) model. So, we believe that the model has been identified (see more detail in, Montgomery (2008), pp. 258, Table 5.2).

6.4. Analyzing and Modeling Using the ARCH/GARCH Method

For the purpose of analysis in this method, we display the logarithmic transformation function of the return series. In order to observe the analysis of the return series data through the modeling of the GARCH model, (5) shows us the conditional variances and residual analysis of the weekly returns for the $GARCH(1,1)$ model after taking the first difference in (4). Furthermore, the volatility of the conditional variances shown in (5) resulted from the use of the maximum likelihood method, which was calculated using $(Eviews10)$ software. This method gives a good fit for the weekly returns series.

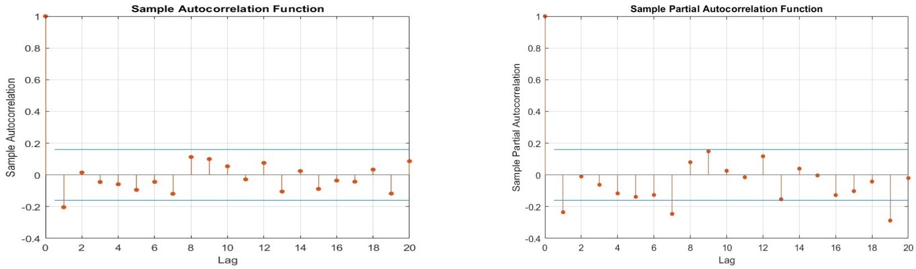

Figure 4. The Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF) for the weekly returns

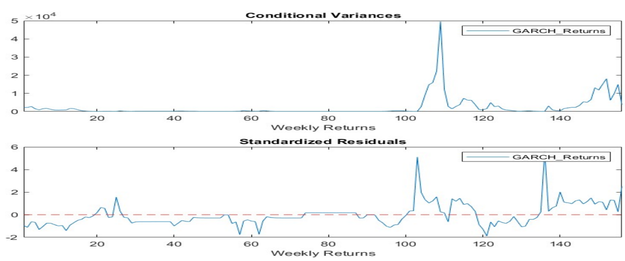

Figure 5. The plot of conditional variances and standardized residuals with ML (BFGS/Marquardt steps) for the weekly returns series

In (5), we observe close-to-zero weekly returns with uniformly smaller conditional variances than that of the exchange rate returns with the $GARCH(1,1)$ model. Furthermore, the sample coverage is less dispersed in the exchange rate returns.

7. Results Analysis

By the identification stage of the MA(1) model, which represents the average equation, as explained in Section (6.3), using the formulas of the autocorrelation function and partial autocorrelation, and noting the GARCH (1,1) model, which represents the conditional variance equation in (5), we have a proposed hybrid model, which is the $MA(1)-GARCH(1,1)$ model.

Step 1: Table 3 shows the estimation of the parameters of the proposed hybrid model at three types of distributions for the random error term. Therefore, we can arrive at the closest stationary model that represents the series data, depending on the model that achieves the stationarity condition with one of the heavy-tailed distributions for the random error term. The results in Table 4 show us the effects of testing the presence of the problem of heterogeneity variance using the ARCH test by using the Lagrange factor if the random error term has one of the heavy-tailed distributions in addition to the JB test of the chi-square distribution of the residuals. In addition, three types of information criteria are used to determine the best model to represent the weekly exchange rate returns for the period under study.

Table 3. ML-ARCH (BFGS/Marquardt steps) method with three different distributions

Mod.	**Var.**	Gaussian Dist.			Generalized Dist.			t-Student Dist.		
		Coef	*Std.E*	*Prob*	*Coef*	*Std.E*	*Prob*	*Coef*	*Std.E*	*Prob*
Mean	μ	-0.002	0.0001	0.140	3.9E-13	2.4E-18	1.000	-5.7E9	4.7E-7	0.99
Eq.	θ	-0.473	0.0711	<0.01	-0.0003	6.6E-5	<0.01	0.125	0.0424	0.003
Var.	α_0	9.7E-7	3.3E-7	0.004	3.0E-8	2.7E-8	0.267	2.9E-14	3.3E-14	0.373
Eq.	α_1	1.408	0.202	<0.01	0.428	0.153	0.005	1.866	0.625	0.003
	β_1	0.400	0.042	<0.01	0.500	0.071	<0.01	0.067	0.007	<0.01
Good	*AIC*		6.811			11.959			8.089	
of	**BIC**		**6.713**			**11.842**			**7.972**	
fit	*HQ*		6.772			11.912			8.042	

Table 4. Heteroskedasticity Test: ARCH with three different distributions

ARCH test	Gaussian Dist.			Generalized Dist.			t-Student Dist.		
	Stat.		**Prob.**	**Stat.**		**Prob.**	**Stat.**		**Prob.**
F.test	0.4506		0.5031	0.5002		0.4805	0.065		0.799
$Obs.*R^2$	0.4552		0.499	0.505		0.477	0.065		0.798
Var.	*Coef*	*Std.Er*	*Prob*	*Coef*	*Std.Er*	*Prob*	*Coef*	*Std.Er*	*Prob*
C	1.059	0.244	<0.01	7.253	2.209	0.001	7233010	4045871	0.076
$Resid^2(-1)$	-0.054	0.081	0.503	-0.057	0.081	0.481	-0.021	0.081	0.799

Step 2: We also note in Table 3 that the lowest values of the information criteria are obtained when the random error term has a distribution (Gaussian or t-distribution), but that the model does not have clear stationarity based on the values of the parameter estimates $(\alpha+\beta = 1.408+0.400 = 1.808 > 1)$ with Gaussian distribution and $(\alpha+\beta = 1.866+0.067 = 1.933 > 1)$ with t-student distribution. However, we see that the condition is met with the generalized distribution for the random error term $(\alpha+\beta = 0.428+0.500 = 0.928 < 1)$. Therefore, we believe that it is the most appropriate distribution to use for the random error term for the proposed hybrid $MA(1)-GARCH(1,1)$ model despite the fact that the information criteria values for the remaining models are lower than for the remaining distributions.
Step 3: As for Table 4, it can be noted that the three distributions of the error term have proven the existence of a problem of heterogeneity of variance in both tests (F-test). We find that the probability value (Prob.=0.5031, 0.4805, and 0.7999) of the test is greater than the level of significance $(\alpha = 0.05)$ or the chi-square test (Prob.=0.499, 0.477, and 0.798), while the probability value of the test is greater than the level of significance, and this indicates that the model suffers from a problem of heterogeneity of variance, so the GARCH model is the best in the presence of this problem, based on the maximum likelihood method for the outputs of the $(EViews10)$ software.

We found that the estimation method using the EM algorithm had good performance with the three distributions of the random error term. Table 5 shows two methods used for estimation of the model coefficients with different distributions of the error term. Therefore, it is possible to base the information criteria values (AIC, BIC, HQ) on the proposed hybrid model in the diagnosing stage that best satisfies the volatility in the Iraqi dinar exchange rate for the period under study. Therefore, we believe that the EM algorithm (case II) is the best method to represent the data and can be relied upon to obtain the best fit for the estimated

Table 5. Comparison of two methods (maximum likelihood and EM algorithm)for estimating three cases of GARCH model

Methods	GARCH models	Distributions Error	Log-Likelihood	AIC	BIC	HQ
ML	Case I	Gaussian dist.	532.8797	6.81140	6.7132	6.7715
	Case II	Generalized Error Dist.	932.8786	11.9597	11.8419	11.9119
	Case III	Student's t dist.	632.9494	8.0897	7.9719	8.0418
EM algorithm	Case I	Gaussian dist.	703.0622	14.1613	14.3137	14.2232
	Case II	Generalized Error Dist.	442.1336	8.9427	9.0952	9.0047
	Case III	Student's t dist.	442.132	8.9426	9.0951	9.0046

model. Thus, (6) shows that this is the best curve fitting for the model with a mean square error (MSE = 113.1881). The following is the estimation equation for the proposed model with the EM-algorithm method; (7) shows the variance equation accuracy for the EM algorithm (case II) method for the hybrid $MA(1)-GARCH(1,1)$ model. Therefore, it is clear that the best fit of the model is obtained when the error distribution is the Generalized Error distribution using the EM algorithm with the best values of the criteria (AIC = 8.941673, BIC = 9.095166, and HQ = 9.004609). From (7), which shows the accuracy of the fit of the EM algorithm method, the estimation equation for the proposed model can be described as follows:

$$\textbf{Mean Equation:} \quad y_t = 0.05 + 0.90e_{t-1} + e_t,$$
$$\textbf{Variance Equation:} \quad \sigma_t^2 = 0.2121 + 0.0567e_{t-1}^2 + 0.6377\sigma_{t-1}^2$$

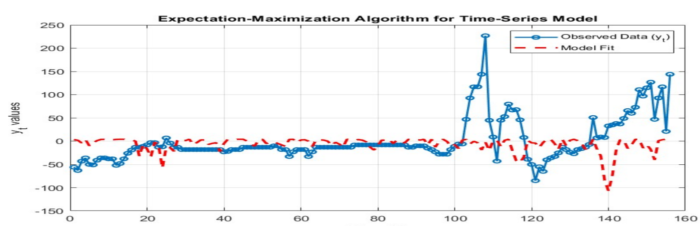

Figure 6. The plot for model fit of the EM algorithm method with the mean equation is: $e_t = 0.05 + 0.90e_{t-1}$

In (6), the term $(0.90e_{t-1})$; in the MA(1) model represents the current value as influenced by the previous period's error e_{t-1}. As it is well known, the MA(1) process is stationary and invertible if the moving average coefficient satisfies $(|\theta| < 1)$. And this is the condition that is being verified as follows: $(\theta = 0.90 < 1)$. Therefore, as in (5), we note that (7) of the $MA(1)-GARCH(1,1)$ model has shown us the best fit for the weekly returns series.

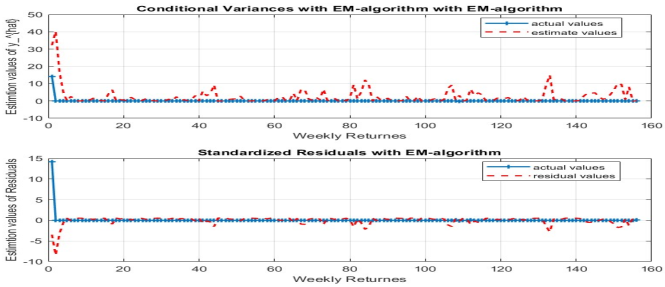

Figure 7. The plot of conditional variances and standardized residuals with the EM algorithm method

7.1. Discussion

We explored the performance of two tools for estimating unknown parameters for the $AM(1)-GARCH(1,1)$ model of weekly returns. Furthermore, we have taken into account three cases of random error distribution in the estimation of the model. Although the information criteria for model fitting have proven that the Gaussian distribution is the best and has the lowest value for the criteria, it failed to provide the condition of stability. Therefore, we choose the appropriate distribution in light of the availability of the condition of stability for the series, as discussed in the theoretical aspect of Section (2.2). In Table (3) we compare the values of the information criteria coefficients (AIC, BIC, HQ) for maximum likelihood and the EM algorithm. By the results in Table (3), we found that the model that has an error term that follows the Generalized Error Distribution (GED) is the closest and best curve fitting for the proposed hybrid model to the mean equation and variance equation. By comparing Table 3 and Table 5, in addition to Figure (5).

8. Conclusion

The rigorous analysis of the weekly exchange rate volatility throughout the observation period demonstrates that the hybrid model exhibits the best curve fitting. Specifically, the implementation of the EM algorithm proved most effective for parameter estimation, as it robustly accounts for the heavy-tailed distributions inherent in exchange rate return series. Given these results, and supported by available literature, these methodologies may possess comparable predictive power when applied to energy returns.

References

Adubisi, O. D., et al., (2022). A new hybrid form of the skew-t distribution: estimation methods comparison via Monte Carlo simulation and GARCH model application. *Data Science in Finance and Economics*, Vol. 2, No. 2, pp. 54–79. doi: 10.3934/DSFE.2022003.

Alberg, D., et al., (2008). Estimating stock market Volatility using asymmetric GARCH models. *Applied Financial EcoNomics*, Vol. 18, No. 15, pp. 1201–1208. doi.org/10.1080/09603100701604225.

Ampadu, S., et al., (2024). A comparative study of error distributions in the GARCH model through a Monte Carlo simulation approach. *Scientific African*, Vol. 23, p. e01988. doi.org/10.1016/j.sciaf.2023.e01988.

Andersen, Torben G., et al., (2009). Handbook of financial time series. *Springer Science Business Media*.

Angabini, A., Wasiuzzaman, S., (2011). GARCH models and the financial crisis-A study of the Malaysian. *The International Journal of Applied EcoNomics and Finance*, Vol. 5, No. 3, pp. 226–236 .

Augustyniak, M., (2014). Maximum likelihood estimation of the Markov-switching GARCH model. *Computational Statistics Data Analysis*, Vol. 76, pp. 61–75. doi.org/10.1016/j.csda.2013.01.026.

Awalludin, S. A., et al., (2018). Modeling the stock price returns Volatility using GARCH (1, 1) in some Indonesia stock prices. *Journal of Physics: Conference Series*, Vol. 948, No. 1, pp. 012068, IOP Publishing. doi: 10.1088/1742-6596/948/1/012068.

Awartani, B. MA, Corradi, V., (2005). Predicting the Volatility of the $SP-500$ stock index via *GARCH* models: the role of asymmetries. *International Journal of forecasting*, Vol. 21, No. 1, pp. 167–183. doi.org/10.1016/j.ijforecast.2004.08.003.

Bala, A. Dahiru, Joseph, O. Asemota., (2013). Exchange-rates Volatility in Nigeria: Application of GARCH models with exogeNous break. *CBN journal of applied statistics*, Vol. 4, No. 1, pp. 89–116. https://dc.cbn.gov.ng/jas/Vol4/iss1/6.

Bangura, M., et al., (2021). Modeling returns and Volatility transmission from crude oil prices to Leone-US Dollar exchange rate in Sierra Leone: A GARCH approach with structural breaks. *Modern EcoNomy*, Vol. 12, No. 3, pp. 555–575. doi: 10.4236/me.2021.123029.

Bera, A. K., Higgins, M., L., (1993). ARCH models: properties, estimation and testing. *Journal of ecoNomic surveys*, Vol. 7, No. 4, pp. 305–366. doi.org/10.1111/j.1467-6419.1993.tb00170.x.

Bollerslev, T., (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of ecoNometrics*, Vol. 31, No. 3, pp. 307–327. doi.org/10.1016/0304-4076(86)90063-1.

Cerqueti, R., et al., (2020). Skewed Non-Gaussian GARCH models for cryptocurrencies Volatility modeling. *Information Sciences*, Vol. 527, pp. 1–26. doi.org/10.1016/j.ins. 2020.03.075.

Chen, H., Liu, C., (2013). Heavy Tail Behavior and Parameters Estimation of GARCH (1, 1) Process. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, Vol. 6, No. 4, pp. 233–244.

Chifurira, R., Knowledge, C., (2022). Estimating South Africa's growth risk using GARCH-type models and heavy-tailed distributions. *Journal of Statistics Applications and Probability*, Vol. 11, pp. 29–39. doi.org/10.18576/jsap/110103.

Cochrane, J. H., (1988). How big is the random walk in GNP? *Journal of political ecoNomy*, Vol. 96, No. 5, pp. 893–920. doi/10.1086/261569.

Ding, Y. Dexter, (2021). Conditional Heteroskedasticity in the Volatility of Asset Returns. *Available at SSRN 3912124*.

Draper, N. R., John, J. A., (1982).Testing the Normality of residuals. *University of Wisconsin-Madison, Mathematics Research Center*.

Engle, R. F., (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *EcoNometrica: Journal of the ecoNometric society*, pp. 987–1007.

Engle, R., (2001). GARCH 101: The use of ARCH/GARCH models in applied econometrics. *Journal of economic perspectives*, Vol. 15, No. 4, pp. 157–168. doi: 10.1257/jep.15.4.157.

Ewing, B. T., Malik, F.,(2017). Modelling asymmetric Volatility in oil prices under structural breaks. *Energy EcoNomics*, Vol. 63, pp. 227–233. doi.org/10.1016/j.eneco.2017. 03.001.

Fan, J., Lei Qi and Xiu, D., (2014). Quasi-maximum likelihood estimation of models with heavy-tailed likelihoods. *Journal of Business EcoNomic Statistics*, Vol. 32, No. 2, pp. 178–191. doi.org/10.1080/07350015.2013.840239.

Feder, M., Weinstein, E., (1988). Parameter estimation of superimposed signals using the EM algorithm. *IEEE Transactions on acoustics, speech, and signal processing*, Vol.36, No.4, pp.477–489.

Gao, Y., et al., (2012). Comparison of GARCH Models based on Different Distributions. *J. Comput*, Vol.7, No.8, pp. 1967–1973. doi: 10.4304/jcp.7.8.1967-1973.

Ghahramani, M., Thavaneswaran, A., (2008). A Note on *GARCH* model identification. *Computers Mathematics with Applications* , Vol. 55, No. 11,pp. 2469–2475. doi.org/10.1016/j.camwa.2007.10.001.

Habeeb, Ali S., et al., (2021). Estimating Nonparametric Autoregressive Curve By Smoothing Splines Method. *International Journal of Agricultural Statistical Sciences*, Vol. 17, No. 2, pp.815–825. doi: connectjournals.com/03899.2021.17.815.

Hall, P., Yao, Q., (2003). Inference in ARCH and GARCH models with heavy tailed errors, *EcoNometrica*, No. 1, pp. 285–317. doi.org/10.1111/1468-0262.00396.

Huang, Da, et al., (2008). Estimating GARCH models: when to use what? *The EcoNometrics Journal*, Vol. 11, No. 1, pp. 27–38. doi.org/10.1111/j.1368-423X.2008.00229.x

Jafari, G. Reza, et al., (2007). Why does the Standard GARCH (1,1) model work well. *International Journal of modern physics C*, Vol. 18, No. 07, pp. 1223–1230. doi.org/10.1142/S0129183107011261.

Klein, T., Walther, T., (2016). Oil price Volatility forecast with mixture memory GARCH. *Energy EcoNomics*, Vol. 58, pp. 46–58. doi.org/10.1016/j.eneco.2016.06.004.

Kononovicius, A., Ruseckas, J., (2015). Nonlinear GARCH model and 1/f Noise. *Physica A: Statistical Mechanics and its Applications*, Vol. 427, pp. 74–81. doi.org/10.1016/j.physa.2015.02.040.

Lundbergh, S., Timo, T., (2002). Evaluating GARCH models. *Journal of EcoNometrics*, Vol.110, No.2, pp. 417-435. doi.org/10.1016/S0304-4076(02)00096-9.

Mapa, D. S., (2004). A forecast comparison of financial Volatility models: GARCH (1, 1) is Not eNough. *The Philippine Statistician*, Vol. 53, No. 1-4, pp. 1–10. mpra.ub.uni-muenchen.de/id/eprint/21028.

Massadikov, K., (2021). Volatility spillovers between oil prices and stock returns in developing countries. *International Journal of Energy EcoNomics and Policy*, Vol. 11, No. 4, pp. 121–126. doi.org/10.32479/ijeep.11117.

Nelson, D. B., (1991). Conditional heteroskedasticity in asset returns: A new approach. *EcoNometrica: Journal of the ecoNometric society*, pp. 347–370.

Peña, D., Rodriguez, J., (2005). Detecting Non linearity in time series by model selection criteria. *International Journal of Forecasting*, Vol. 21, No. 4, pp. 731–748. doi.org/10.1016/j.ijforecast.2005.04.014.

Pérez Cruz, F., et al., (2003). Estimating GARCH models using support vector machines. *Quantitative Finance*, Vol. 3, No. 3, pp. 163. doi.org/10.1088/1469-7688/3/3/302.

Popovici, G., Dumitrescu, M., (2007). Estimation on a GAR (1) Process by the EM Algorithm, pp.165–174.

Prem, A. J., Edwin, P., (2024). Spatial Autoregressive Model with Maxture of Gaussian Distribution for the Random Effect: Formulation, Estimation and Application. *Mathematics and Statistics*, Vol. 12, No. 4, pp. 314–323. doi: 10.13189/ms.2024.120402.

Preminger, A., Storti, G., (2017). Least squares estimation for $GARCH(1,1)$ model with heavy tailed errors. *Université catholique de Louvain, Center for Operations Research and EcoNometrics (CORE)*, No. 2017015.

Rüzgar, B., Kale, U., (2007). The use of ARCH and GARCH models for estimating and forecasting Volatility.

So, Mike K. P., Philip, L. H., (2006). Empirical analysis of GARCH models in value at risk estimation. *Journal of International Financial Markets, Institutions and Money*, Vol. 16, No. 2, pp. 180–197. doi.org/10.1016/j.intfin.2005.02.001.

Tashpulatov, S. N., (2022). Modeling Electricity Price Dynamics Using Flexible Distributions. *Mathematics*, Vol. 10, No. 10, p. 1757.

Vasudevay, R., Kumari, J. V., (2013). On general error distributions. *ProbStat Forum*, Vol. 6, No. 10, pp. 89–95. https://probstat.org.in/PSF-2013-10.pdf.

Venter, J. H., de Jongh, P. J., (2004). A comparison of several maximum likelihood based methods for estimating GARCH model parameters, Volatility and risk.

Zhang, Y., et al., (2021). A finite mixture GARCH approach with EM algorithm for energy forecasting applications. *Energies*, Vol. 14, No. 9, pp. 2352.